Nonlinear gyrokinetic equation

BY YOUJUN HU

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

Abstract

The nonlinear δf gyrokinetic equation in Frieman-Chen's paper[2] is re-derived in this article, with more details. All formulas are in SI units. Numerical implementation of the gyrokinetic model using the PIC method is also discussed. A gyrokinetic PIC code called `TEK` was developed using the formulas given in this document. TEK has been benchmarked with GENE code in the DIII-D cyclone base case for both ITG-KBM transition and ITG-TEM transition.

1 Introduction

1.1 Gyrokinetic?

Electromagnetic perturbations of frequency lower than the ion cyclotron frequency are widely believed to be more important than high-frequency ones in transporting plasma in tokamaks. (This assumption can be verified numerically when we are able to do a full simulation including both low-frequency and high-frequency perturbations. This kind of verification is not possible at present due to computation costs.)

If we assume that only low-frequency perturbations are present, the Vlasov equation can be simplified. Specifically, symmetry of the particle distribution function in the phase space can be established if we choose suitable phase-space coordinates and split the distribution function in a proper way. The symmetry is along the so-called gyro-angle α in the guiding-center coordinates ($\mathbf{X}$, μ, ε, α). In obtaining the equation for the gyro-angle independent part of the distribution function, we need to average the equation over the gyro-angle α and thus this model is called "gyrokinetic".

In deriving the gyrokinetic equation, the perturbed electromagnetic field is assumed to be known and of low-frequency. To do a kinetic simulation, we need to solve the field equation to obtain the perturbed electromagnetic field. It is still possible that high frequency modes (e.g., compressional Alfven waves and Ω_H modes) appear in a gyrokinetic simulation. If the amplitude of high frequency modes is significantly large, then the simulation is invalid because the gyrokinetic model is invalid in this case.

Our starting point is the Vlasov equation in terms of particle coordinates $(\mathbf{x}, \mathbf{v})$:

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

where $f = f(t, \mathbf{x}, \mathbf{v})$ is the particle distribution function, $\mathbf{x}$ and $\mathbf{v}$ are the location and velocity of particles. The distribution function f depends on 6 phase-space variables $(\mathbf{x}, \mathbf{v})$, besides the time t.

1.2 Guiding-center coordinates: a simple example

Choose Cartesian coordinates (x, y, z) for the configuration space $\mathbf{x}$. Consider a simple case where the electromagnetic field is a time-independent field given by $\mathbf{B} = B_0(x, y, z)\hat{\mathbf{z}}$ and $\mathbf{E} = 0$. Let us examine the Vlasov equation in this case and see whether there is any coordinate system that can simplify the Vlasov equation.

Describe the velocity space using a right-handed cylindrical coordinates $(v_\perp, \alpha, v_\parallel)$, where $v_\parallel = \mathbf{v} \cdot \mathbf{e}_\parallel$, $\mathbf{e}_\parallel = \mathbf{e}_z$ is the unit vector along the magnetic field, α is the azimuthal angle of the perpendicular velocity ($\mathbf{v}_\perp = \mathbf{v} - v_\parallel \mathbf{e}_z$) relative to $\mathbf{e}_x$. Note that these coordinates are defined relative to the local magnetic field, which, in more general cases, may vary in space.

Next, let us express the spatial gradient of f in terms of partial derivatives:

$$\left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\mathbf{v}} = \left.\frac{\partial f}{\partial x}\right|_{\mathbf{v}} \mathbf{e}_x + \left.\frac{\partial f}{\partial y}\right|_{\mathbf{v}} \mathbf{e}_y + \left.\frac{\partial f}{\partial z}\right|_{\mathbf{v}} \mathbf{e}_z. \tag{2}$$

Note that this gradient is taken by holding $\mathbf{v}$ constant (i.e., (v_x, v_y, v_z) constant rather than $(v_\perp, \alpha, v_\parallel)$ constant). So, we need do the following coordinate transform:

$$(x,\ y,\ z,\ v_x,\ v_y,\ v_z) \Rightarrow \left(x' = x,\ y' = y,\ z' = z,\ v_\perp = \sqrt{v_x^2 + v_y^2},\ \alpha = \mathrm{atan2}(v_y,\ v_x),\ v_\parallel = v_z\right) \tag{3}$$

Using chain rule, expression (2) is written as

$$\left.\frac{\partial f}{\partial x}\right|_{\mathbf{v}} = \frac{\partial f}{\partial x'}\frac{\partial x'}{\partial x} + \left.\frac{\partial f}{\partial v_\perp}\frac{\partial v_\perp}{\partial x}\right|_{\mathbf{v}} + \left.\frac{\partial f}{\partial \alpha}\frac{\partial \alpha}{\partial x}\right|_{\mathbf{v}} + \left.\frac{\partial f}{\partial v_\parallel}\frac{\partial v_\parallel}{\partial x}\right|_{\mathbf{v}}. \tag{4}$$

$$\approx \frac{\partial f}{\partial x'}. \tag{5}$$

Note that the partial derivatives

$$\left.\frac{\partial v_\perp}{\partial x}\right|_{\mathbf{v}}, \left.\frac{\partial \alpha}{\partial x}\right|_{\mathbf{v}}, \left.\frac{\partial v_\parallel}{\partial x}\right|_{\mathbf{v}} \tag{6}$$

are generally nonzero since the defintion of $(v_\perp, \alpha, v_\parallel)$ depends on the local magnetic field direction. In our simple case, they are zero since the magnetic field direction is uniform. Similarly, we obtain

$$\left.\frac{\partial f}{\partial y}\right|_{\mathbf{v}} \approx \frac{\partial f}{\partial y'} \tag{7}$$

$$\left.\frac{\partial f}{\partial z}\right|_{\mathbf{v}} = \frac{\partial f}{\partial z'} \tag{8}$$

Then, in $(x', y', z', v_\perp, \alpha, v_\parallel)$ coordinates, the spatial gradient of f is written as

$$\left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\mathbf{v}} \approx \frac{\partial f}{\partial x'}\mathbf{e}_x + \frac{\partial f}{\partial y'}\mathbf{e}_y + \frac{\partial f}{\partial z'}\mathbf{e}_z. \tag{9}$$

Next, consider the gradient in velocity space

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{v}} &= \frac{\partial f}{\partial v_x}\mathbf{e}_x+\frac{\partial f}{\partial v_y}\mathbf{e}_y+\frac{\partial f}{\partial v_z}\mathbf{e}_z\\ &= \frac{\partial f}{\partial v_\perp}\mathbf{e}_1+\frac{\partial f}{\partial v_\parallel}\mathbf{e}_z+\frac{1}{v_\perp}\frac{\partial f}{\partial \alpha}\mathbf{e}_\alpha,\end{aligned} \tag{10}$$

where $\mathbf{e}_1=\mathbf{v}_\perp/|\mathbf{v}_\perp|$, $\mathbf{v}_\perp=\mathbf{v}-(\mathbf{v}\cdot\mathbf{e}_z)\mathbf{e}_z$, and $\mathbf{e}_\alpha=\mathbf{e}_z\times\mathbf{e}_1$. Using Eq. (10), $\frac{q}{m}(\mathbf{v}\times\mathbf{B})\cdot\frac{\partial f}{\partial \mathbf{v}}$ is written as

$$\begin{aligned}\frac{q}{m}(\mathbf{v}\times\mathbf{B})\cdot\frac{\partial f}{\partial \mathbf{v}} &= \frac{q}{m}(v_\perp\mathbf{e}_1\times\mathbf{B})\cdot\left(\frac{\partial f}{\partial v_\perp}\mathbf{e}_1+\frac{\partial f}{\partial v_\parallel}\mathbf{e}_z+\frac{1}{v_\perp}\frac{\partial f}{\partial \alpha}\mathbf{e}_\alpha\right)\\ &= \frac{q}{m}(v_\perp\mathbf{e}_1\times\mathbf{B})\cdot\left(\frac{1}{v_\perp}\frac{\partial f}{\partial \alpha}\mathbf{e}_\alpha\right)\\ &= \frac{Bq}{m}\frac{\partial f}{\partial \alpha}(\mathbf{e}_1\times\mathbf{e}_z)\cdot\mathbf{e}_\alpha\\ &= -\Omega\left(\frac{\partial f}{\partial \alpha}\right),\end{aligned} \tag{11}$$

where $\Omega=Bq/m$ is the gyro-frequency. Then the Vlasov equation (1) is written as

$$\frac{\partial f}{\partial t}+\mathbf{v}\cdot\frac{\partial f}{\partial \mathbf{x}}-\Omega\frac{\partial f}{\partial \alpha}=0, \tag{12}$$

i.e.,

$$\frac{\partial f}{\partial t}+v_\parallel\frac{\partial f}{\partial z'}+v_\perp\cos\alpha\frac{\partial f}{\partial x'}+v_\perp\sin\alpha\frac{\partial f}{\partial y'}-\Omega\frac{\partial f}{\partial \alpha}=0. \tag{13}$$

Define the following coordinates transform (guiding-center transform):

$$\begin{cases} X=x'+\dfrac{v_\perp\sin\alpha}{\Omega},\\ Y=y'-\dfrac{v_\perp\cos\alpha}{\Omega},\\ Z=z',\\ \alpha'=\alpha,\\ v'_\parallel=v_\parallel,\\ v'_\perp=v_\perp,\\ t'=t \end{cases} \tag{14}$$

Next, we express the partial derivatives of f appearing in Eq. (13) in terms of the guiding-center coordinates. Using the chain rule, we obtain

$$\begin{aligned}\frac{\partial f}{\partial \alpha} &= \frac{\partial f}{\partial X}\frac{\partial X}{\partial \alpha}+\frac{\partial f}{\partial Y}\frac{\partial Y}{\partial \alpha}+\frac{\partial f}{\partial Z}\frac{\partial Z}{\partial \alpha}+\frac{\partial f}{\partial \alpha'}\frac{\partial \alpha'}{\partial \alpha}+\frac{\partial f}{\partial v'_\parallel}\frac{\partial v'_\parallel}{\partial \alpha}+\frac{\partial f}{\partial v'_\perp}\frac{\partial v'_\perp}{\partial \alpha}+\frac{\partial f}{\partial t'}\frac{\partial t'}{\partial \alpha}\\ &= \frac{\partial f}{\partial X}\frac{\partial X}{\partial \alpha}+\frac{\partial f}{\partial Y}\frac{\partial Y}{\partial \alpha}+\frac{\partial f}{\partial \alpha'}\\ &= \frac{\partial f}{\partial x'}\frac{v_\perp\cos\alpha}{\Omega}+\frac{\partial f}{\partial y'}\frac{v_\perp\sin\alpha}{\Omega}+\frac{\partial f}{\partial \alpha'}.\end{aligned} \tag{15}$$

Similarly, for other derivatives, we obtain

$$\frac{\partial f}{\partial t}=\frac{\partial f}{\partial t'}, \tag{16}$$

and

$$\begin{aligned}\frac{\partial f}{\partial z'} &= \frac{\partial f}{\partial X}\frac{\partial X}{\partial z'}+\frac{\partial f}{\partial Y}\frac{\partial Y}{\partial z'}+\frac{\partial f}{\partial Z}\frac{\partial Z}{\partial z'}+\frac{\partial f}{\partial \alpha'}\frac{\partial \alpha'}{\partial z'}+\frac{\partial f}{\partial v'_{\parallel}}\frac{\partial v'_{\parallel}}{\partial z'}+\frac{\partial f}{\partial v'_{\perp}}\frac{\partial v'_{\perp}}{\partial z'}+\frac{\partial f}{\partial t'}\frac{\partial t'}{\partial z'}\\ &\approx \frac{\partial f}{\partial X}\times 0+\frac{\partial f}{\partial Y}\times 0+\frac{\partial f}{\partial Z},\end{aligned} \tag{17}$$

where we have assumed that the dependence of X and Y on z' (via Ω) is weak enough to be neglected. And also

$$\begin{aligned}\frac{\partial f}{\partial x'} &= \frac{\partial f}{\partial X}\frac{\partial X}{\partial x'}+\frac{\partial f}{\partial Y}\frac{\partial Y}{\partial x'}+\frac{\partial f}{\partial Z}\frac{\partial Z}{\partial x'}+\frac{\partial f}{\partial \alpha'}\frac{\partial \alpha'}{\partial x'}+\frac{\partial f}{\partial v'_{\parallel}}\frac{\partial v'_{\parallel}}{\partial x'}+\frac{\partial f}{\partial v'_{\perp}}\frac{\partial v'_{\perp}}{\partial x'}+\frac{\partial f}{\partial t'}\frac{\partial t'}{\partial x'}\\ &= \frac{\partial f}{\partial X}\frac{\partial X}{\partial x'}+\frac{\partial f}{\partial Y}\frac{\partial Y}{\partial x'}\\ &\approx \frac{\partial f}{\partial X}\times 1+\frac{\partial f}{\partial Y}\times 0,\end{aligned} \tag{18}$$

$$\frac{\partial f}{\partial y'}\approx\frac{\partial f}{\partial Y} \tag{19}$$

Using the above results, equation (13) in the guiding-center coordinates is written as

$$\frac{\partial f}{\partial t'}+v'_{\parallel}\frac{\partial f}{\partial Z}+v'_{\perp}\cos\alpha'\frac{\partial f}{\partial X}+v'_{\perp}\sin\alpha'\frac{\partial f}{\partial Y}-\Omega\left(\frac{\partial f}{\partial X}\frac{v_{\perp}\cos\alpha}{\Omega}+\frac{\partial f}{\partial Y}\frac{v_{\perp}\sin\alpha}{\Omega}+\frac{\partial f}{\partial \alpha'}\right)=0. \tag{20}$$

Here we find some terms cancel each other, giving

$$\frac{\partial f}{\partial t'}+v'_{\parallel}\frac{\partial f}{\partial Z}-\Omega\left(\frac{\partial f}{\partial \alpha'}\right)=0. \tag{21}$$

Equation (21) is the equation for f in the guiding-center coordinates $(X,Y,Z,v'_{\perp},\alpha',v'_{\parallel})$.

1.2.1 Gyro-averaging

Define gyro-phase averaging operator

$$\langle\ldots\rangle\equiv(2\pi)^{-1}\int_0^{2\pi}(\ldots)d\alpha'. \tag{22}$$

Then averaging Eq. (21) over α', we get

$$\frac{\partial\langle f\rangle}{\partial t'}+v'_{\parallel}\frac{\partial\langle f\rangle}{\partial Z}-(2\pi)^{-1}\int_0^{2\pi}\Omega\left(\frac{\partial f}{\partial \alpha'}\right)d\alpha'=0. \tag{23}$$

As is assumed above, Ω is approximately independent of α' and thus Ω in Eq. (23) can be moved outside of the gyro-angle integration, giving

$$\frac{\partial\langle f\rangle}{\partial t'}+v_\parallel'\frac{\partial\langle f\rangle}{\partial Z}-\Omega(2\pi)^{-1}\int_0^{2\pi}\left(\frac{\partial f}{\partial\alpha'}\right)d\alpha'=0. \tag{24}$$

Peforming the integration, we get

$$\frac{\partial\langle f\rangle}{\partial t'}+v_\parallel'\frac{\partial\langle f\rangle}{\partial Z}-\Omega(2\pi)^{-1}[f(\alpha'=2\pi)-f(\alpha'=0)]=0. \tag{25}$$

Since $\alpha'=2\pi$ and $\alpha'=0$ correspond to the same phase space location, the corresponding values of the distribution function must be equal. Then the above equation reduces to

$$\frac{\partial\langle f\rangle}{\partial t'}+v_\parallel'\frac{\partial\langle f\rangle}{\partial Z}=0, \tag{26}$$

which is an equation for the gyro-angle independent part of the distribution function.

Next let us investigate whether the guiding-center transform can be made use of to simplify the kinetic equation for more general cases where we have a (weakly) non-uniform static magnetic field, plus electromagnetic perturbations of low frequency (and of small amplitude and $k_\parallel\rho_i\ll 1$). And we will include the effect that $\mathbf{B}_0$ depends on α', i.e., grad-B and curvature drift.

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

Next, we define the guiding-center transform and then transform the Vlasov equation from the particle coordinates $(\mathbf{x},\mathbf{v})$ 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 coordinates.

2.1 Guiding-center transformation

In a magnetic field, given a particle location and velocity ($\mathbf{x}$, $\mathbf{v}$), we know how to calculate its guiding-center location $\mathbf{X}$:

$$\mathbf{X}=\mathbf{x}+\mathbf{v}\times\frac{\boldsymbol{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})}, \tag{27}$$

where $\boldsymbol{e}_\parallel=\mathbf{B}_0/B_0$, $\Omega=qB_0/m$, $\mathbf{B}_0=\mathbf{B}_0(\mathbf{x})$ is the equilibrium (macroscopic) magnetic field at the particle position. We will consider Eq. (27) as a transform and call it guiding-center transform. Note that the transform (27) involves both position and velocity of particles.

Given ($\mathbf{x}$, $\mathbf{v}$), it is straightforward to obtain $\mathbf{X}$ by using Eq. (27). On the other hand, the inverse transform (i.e., given ($\mathbf{X}$, $\mathbf{v}$), to find $\mathbf{x}$) is not easy because Ω and $\mathbf{e}_\parallel$ depend on $\mathbf{x}$, which usually requires us to solve a nonlinear equation for its root. Numerically, one can use

$$\mathbf{x}_{n+1}=\mathbf{X}-\mathbf{v}\times\frac{\boldsymbol{e}_\parallel(\mathbf{x}_n)}{\Omega(\mathbf{x}_n)}, \tag{28}$$

as an iteration scheme to compute $\mathbf{x}$, with the initial guess chosen as $\mathbf{x}_0 = \mathbf{X}$. If we stop at the first iteration, then

$$\mathbf{x} \approx \mathbf{X} - \mathbf{v} \times \frac{\boldsymbol{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}. \tag{29}$$

The approximation relation (29) is usually used as the inverse guiding-center transformation in gyrokinetic PIC simulations. This transformation needs to be performed numerically when we deposit markers to grids, or when we calculate the gyro-averaged field to be used in pushing guiding-centers.

[The equilibrium magnetic field we will consider has spatial scale length much larger than the thermal gyro radius ρ. 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 negligible. The difference between equilibrium field values evaluated at $\mathbf{X}$ and $\mathbf{x}$ is neglected in gyrokinetic theory (except in deriving the gradient/curvature drift). 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.]

For later use, define $\boldsymbol{\rho} \equiv -\mathbf{v} \times \mathbf{e}_\parallel/\Omega$, which is the vector gyro-radius pointing from the the guiding-center to the particle position.

2.2 Choosing velocity coordinates

The guiding-center transformations (27) and (29) involve the particle velocity $\mathbf{v}$. It is the cross product between $\mathbf{v}$ and $\boldsymbol{e}_\parallel$ that is actually used. Therefore, only the perpendicular velocity (which is defined by $\mathbf{v}_\perp = \mathbf{v} - \mathbf{v}\cdot\mathbf{e}_\parallel$) enters the transform. A natural choice of coordinates for the perpendicular velocity is $(v_\perp, \alpha)$, where $v_\perp = |\mathbf{v}_\perp|$ and α is the azimuthal angle of the perpendicular velocity in the local perpendicular plane. There are degrees of freedom in choosing one of the two perpendicular basis vectors, with respect to which α is defined. In `GEM` code, one of the perpendicular direction is chosen as the direction perpendicular to the magnetic surface, which is fully determined at each spatial point. (We 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.)

In the following, α will be called "gyro-angle". Note that α is a coordinate of $\mathbf{v}$ space. In particle coordinate $(\mathbf{x}, \mathbf{v})$, by definition, varying α does not affect $\mathbf{x}$. In the guiding-center coordinates $(\mathbf{X}, v_\parallel, v_\perp, \alpha)$, by definition, varying α does not affect $\mathbf{X}$. However, when transformed back to particle coordinates, α affects both the velocity coordinate and the spatial coordinate. [Consider a series of points in terms of guiding-center coordinates $(\mathbf{X}, v_\parallel, v_\perp, \alpha)$ with $(\mathbf{X}, v_\parallel, v_\perp)$ fixed but with α changing. Using the inverse guiding-center transform (29), we know that the above points form a gyro-ring in space, i.e., α influences spatial location, in addtion to velocity.]

The gyro-angle is an important variable we will stick to because we need to directly perform averaging over this variable (with $\mathbf{X}$ fixed) in deriving the gyrokinetic equation. We have multiple options for the remaining velocity coordinates , such as $(v, v_\parallel)$, or $(v, v_\perp)$, or $(v_\parallel, v_\perp)$. In Frieman-Chen's paper, the velocity coordinates other than α are chosen to be (ε, μ) defined by

$$\varepsilon = \varepsilon(v, \mathbf{x}) \equiv \frac{v^2}{2} + \frac{q\Phi_0(\mathbf{x})}{m}, \tag{30}$$

and

$$\mu = \mu(v_\perp, \mathbf{x}) \equiv \frac{v_\perp^2}{2B_0(\mathbf{x})}, \tag{31}$$

where $\Phi_0(\mathbf{x})$ is the equilibrium (macroscopic) electrical potential. Choosing μ as one of the phase space coordinates is nontrivial because it turns out to be a constant of motion. And this choice seems to be important in sucessfully getting the final gyrokinetic equation (I need to check this).

Note that $(\varepsilon, \mu, \alpha)$ is not sufficient in uniquely determining a velocity vector. An additional parameter $\sigma = \text{sign}(v_\parallel)$ 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.

One important thing to note about the above velocity coordinates is that they are defined relative to the local magnetic field. Because the tokamak magnetic field is spatially varying, the above velocity coordinates are also spatially varying for a fixed velocity $\mathbf{v}$. Specifically, the following derivatives are nonzero:

$$\frac{\partial \alpha}{\partial \mathbf{x}}|_{\mathbf{v}}, \frac{\partial v_\perp}{\partial \mathbf{x}}|_{\mathbf{v}}, \frac{\partial v_\parallel}{\partial \mathbf{x}}|_{\mathbf{v}}, \frac{\partial \mu}{\partial \mathbf{x}}|_{\mathbf{v}}. \tag{32}$$

2.3 Summary of the phase-space coordinate transform

The transform from particle variables $(\mathbf{x}, \mathbf{v})$ to guiding-center variables $(\mathbf{X}, \varepsilon, \mu, \alpha, \sigma)$ is given by

$$\begin{cases} \mathbf{X} = \mathbf{x} + \mathbf{v} \times \frac{\boldsymbol{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})} \\ \varepsilon = \frac{|\mathbf{v}|^2}{2} + \frac{q\Phi_0(\mathbf{x})}{m} \\ \mu = \frac{|\mathbf{v} - \mathbf{v} \cdot \mathbf{e}_\parallel(\mathbf{x})|^2}{2B_0(\mathbf{x})} \\ \alpha = (\text{angle between } \mathbf{v}_\perp \text{ and local } \mathbf{e}_\perp) \\ \sigma = \text{sign}(\mathbf{v} \cdot \mathbf{e}_\parallel(\mathbf{x})) \end{cases} \tag{33}$$

As mentioned above, the dependence of the distribution function on σ will be hidden in the following.

2.4 Distribution function in terms of guiding-center variables

Denote the particle distribution function expressed in particle coordinates $(\mathbf{x},\mathbf{v})$ by f_p, and the same distribution expressed in the guiding-center variables $(\mathbf{X},\varepsilon,\mu,\alpha)$ by f_g. Then

$$f_g(\mathbf{X},\varepsilon,\mu,\alpha)=f_p(\mathbf{x},\mathbf{v}), \tag{34}$$

where $(\mathbf{x},\mathbf{v})$ and $(\mathbf{X},\varepsilon,\mu,\alpha)$ are related to each other by the guiding-center transform (29). Equation (34) along the guiding-center transform can be considered as the definition of f_g.

As is conventionally adopted in multi-variables calculus, both f_p and f_g are often denoted by the same symbol, say f. Which set of independent variables are actually assumed is inferred from the context. This is one subtle thing for gyrokinetic theory in particular and for multi-variables calculus in general. (Sometimes, it may be better to use subscript notation on f to identify which coordinates are assumed. 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 F.)

In practice, f_g is often called the guiding-center distribution function whereas f_p is called the particle distribution function. However, they are actually the same distribution function expressed in different variables. The name "guiding-center distribution function" is misleading because it may imply that we can count the number of guiding-centers to obtain this distribution function but this implication is wrong.

2.5 Spatial gradient operator in guiding-center coordinates

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

$$\frac{\partial f_p}{\partial\mathbf{x}}|_{\mathbf{v}}=\frac{\partial\mathbf{X}}{\partial\mathbf{x}}\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\frac{\partial\varepsilon}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\varepsilon}+\frac{\partial\mu}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\mu}+\frac{\partial\alpha}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\alpha}. \tag{35}$$

From the definition of $\mathbf{X}$, Eq. (27), we obtain

$$\frac{\partial\mathbf{X}}{\partial\mathbf{x}}=\mathbf{I}+\mathbf{v}\times\frac{\partial}{\partial\mathbf{x}}\Big(\frac{\boldsymbol{e}_\parallel}{\Omega}\Big), \tag{36}$$

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

$$\frac{\partial\varepsilon}{\partial\mathbf{x}}=-\frac{q}{m}\mathbf{E}_0, \tag{37}$$

where $\mathbf{E}_0=-\partial\Phi_0/\partial\mathbf{x}$. Using the above results, equation (35) is written as

$$\frac{\partial f_p}{\partial\mathbf{x}}|_{\mathbf{v}}=\frac{\partial f_g}{\partial\mathbf{X}}+\left[\mathbf{v}\times\frac{\partial}{\partial\mathbf{x}}\Big(\frac{\boldsymbol{e}_\parallel}{\Omega}\Big)\right]\cdot\frac{\partial f_g}{\partial\mathbf{X}}-\frac{q}{m}\mathbf{E}_0\frac{\partial f_g}{\partial\varepsilon}+\frac{\partial\mu}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\mu}+\frac{\partial\alpha}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\alpha}. \tag{38}$$

As mentioned above, the partial derivative $\partial/\partial\mathbf{x}$ is taken by holding $\mathbf{v}$ constant. Since $\mathbf{B}_0$ is spatially varying, $\mathbf{v}_\perp$ is spatially varying when holding $\mathbf{v}$ constant. Therefore $\frac{\partial\mu}{\partial\mathbf{x}}$ and $\frac{\partial\alpha}{\partial\mathbf{x}}$ are generally nonzero. The explicit expressions of these two derivatives are needed later in the derivation of the gyrokinetic equation and is discussed in Appendix

I. For notation ease, define

$$\boldsymbol{\lambda}_{B1} = \left[\mathbf{v} \times \frac{\partial}{\partial \mathbf{x}}\left(\frac{\boldsymbol{e}_{\parallel}}{\Omega}\right)\right] \cdot \frac{\partial}{\partial \mathbf{X}}, \tag{39}$$

and

$$\boldsymbol{\lambda}_{B2} = \frac{\partial \mu}{\partial \mathbf{x}} \frac{\partial}{\partial \mu} + \frac{\partial \alpha}{\partial \mathbf{x}} \frac{\partial}{\partial \alpha}, \tag{40}$$

then expression (38) is written as

$$\frac{\partial f_p}{\partial \mathbf{x}}|_{\mathbf{v}} = \frac{\partial f_g}{\partial \mathbf{X}} + [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}] f_g - \frac{q}{m} \mathbf{E}_0 \frac{\partial f_g}{\partial \varepsilon}. \tag{41}$$

Note that $\partial / \partial \mathbf{X}$ is a shorthand for

$$\frac{\partial}{\partial \mathbf{X}}|_{\varepsilon, \mu, \alpha}$$

i.e., it is taken by holding (ε, μ, α) constant (rather than holding $\mathbf{v}$ constant). For notation ease, $\partial / \partial \mathbf{X}$ is sometimes denoted by ∇_X or simply ∇.

2.6 Velocity gradient operator in guiding-center coordinates

Next, let us express 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

$$\frac{\partial f_p}{\partial \mathbf{v}}|_{\mathbf{x}} = \frac{\partial \mathbf{X}}{\partial \mathbf{v}} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \varepsilon}{\partial \mathbf{v}} \frac{\partial f_g}{\partial \varepsilon} + \frac{\partial \mu}{\partial \mathbf{v}} \frac{\partial f_g}{\partial \mu} + \frac{\partial \alpha}{\partial \mathbf{v}} \frac{\partial f_g}{\partial \alpha}. \tag{42}$$

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

$$\begin{aligned} \frac{\partial \mathbf{X}}{\partial \mathbf{v}} &= \frac{\partial}{\partial \mathbf{v}}\left(\frac{\boldsymbol{v} \times \boldsymbol{e}_{\parallel}}{\Omega}\right) \\ &= \frac{\partial \mathbf{v}}{\partial \mathbf{v}} \times \frac{\boldsymbol{e}_{\parallel}}{\Omega} \\ &= \mathbf{I} \times \frac{\boldsymbol{e}_{\parallel}}{\Omega}. \end{aligned} \tag{43}$$

From the definition of ε, we obtain

$$\frac{\partial \varepsilon}{\partial \mathbf{v}} = \mathbf{v}, \tag{44}$$

From the definition of μ, we obtain

$$\frac{\partial \mu}{\partial \mathbf{v}} = \frac{\mathbf{v}_{\perp}}{B_0}, \tag{45}$$

From the definition of α, we obtain

$$\frac{\partial \alpha}{\partial \mathbf{v}} = \frac{1}{v_{\perp}}\left(\boldsymbol{e}_{\parallel} \times \frac{\mathbf{v}_{\perp}}{v_{\perp}}\right) = \frac{\boldsymbol{e}_{\alpha}}{v_{\perp}}, \tag{46}$$

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

$$\boldsymbol{e}_{\alpha} = \boldsymbol{e}_{\parallel} \times \frac{\mathbf{v}_{\perp}}{v_{\perp}}. \tag{47}$$

Using the above results, expression (42) is written

$$\frac{\partial f_p}{\partial \mathbf{v}}|_{\mathbf{x}} = \frac{\mathbf{I} \times \boldsymbol{e}_{\parallel}}{\Omega} \cdot \frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v} \frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_{\perp}}{B_0} \frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_{\alpha}}{v_{\perp}} \frac{\partial f_g}{\partial \alpha}. \tag{48}$$

2.7 Time derivatives in guiding-center coordinates

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

$$\frac{\partial f_p}{\partial t}|_{\mathbf{x},\mathbf{v}} = \frac{\partial f_g}{\partial t}|_{\mathbf{X},\mathbf{V}} + \frac{\partial \mathbf{X}}{\partial t}\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \frac{\partial \mathbf{V}}{\partial t}\cdot\frac{\partial f_g}{\partial \mathbf{V}}, \tag{49}$$

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 very slow in the gyrokinetic ordering discussed later. In the following, $\partial\mathbf{X}\,/\,\partial t$ and $\partial\mathbf{V}\,/\,\partial t$ will be dropped, i.e.,

$$\frac{\partial f_p}{\partial t} \approx \frac{\partial f_g}{\partial t}. \tag{50}$$

2.8 Final form of Vlasov equation in guiding-center coordinates

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

$$\begin{aligned} & \frac{\partial f_g}{\partial t} + \mathbf{v}\cdot\left[\frac{\partial f_g}{\partial \mathbf{X}} + [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}]\, f_g - \frac{q}{m}\mathbf{E}_0\frac{\partial f_g}{\partial \varepsilon}\right] \\ +\;& \frac{q}{m}(\mathbf{E} + \mathbf{v}\times\mathbf{B})\cdot\left(\frac{\mathbf{I}\times\boldsymbol{e}_\parallel}{\Omega}\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v}\frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \\ =\;& 0 \end{aligned} \tag{51}$$

Using tensor identity $\mathbf{a}\cdot\mathbf{I}\times\mathbf{b} = \mathbf{a}\times\mathbf{b}$, equation (51) is written as

$$\begin{aligned} & \frac{\partial f_g}{\partial t} + \mathbf{v}\cdot\left[\frac{\partial f_g}{\partial \mathbf{X}} + [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}]\, f_g - \frac{q}{m}\mathbf{E}_0\frac{\partial f_g}{\partial \varepsilon}\right] \\ +\;& \frac{q}{m}(\mathbf{E} + \mathbf{v}\times\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \frac{q}{m}(\mathbf{v}\times\mathbf{B})\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \\ +\;& \frac{q}{m}\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \\ =\;& 0, \end{aligned} \tag{52}$$

This is the Vlasov equation in guiding-center coordinates.

3 δf form of Vlasov equation in guiding-center variables

3.1 Electromagnetic field perturbation

Since the definition of the guiding-center variables $(\mathbf{X}, \varepsilon, \mu, \alpha)$ involves the equilibrium fields $\mathbf{B}_0$ and $\mathbf{E}_0$, to further simplify Eq. (52), we need to separate electromagnetic field into equilibrium and perturbation parts. Writing the electromagnetic field as

$$\mathbf{E} = \mathbf{E}_0 + \delta\mathbf{E} \tag{53}$$

and

$$\mathbf{B} = \mathbf{B}_0 + \delta\mathbf{B}, \tag{54}$$

then substituting these expressions into equation (52) and moving all terms involving the field perturbations to the right-hand side, we obtain

$$\begin{aligned}
&\frac{\partial f_g}{\partial t} + \mathbf{v}\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v}\cdot\left[(\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}) f_g - \frac{q}{m}\mathbf{E}_0\frac{\partial f_g}{\partial \varepsilon}\right] \\
+\ &\frac{q}{m}(\mathbf{E}_0 + \mathbf{v}\times\mathbf{B}_0)\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \frac{q}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \\
+\ &\frac{q}{m}\mathbf{E}_0\cdot\left(\mathbf{v}\frac{\partial f_g}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \\
=\ &\delta R f_g,
\end{aligned} \tag{55}$$

where δR is defined by

$$\begin{aligned}
\delta R \ =\ &-\frac{q}{m}(\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial}{\partial \mathbf{X}} - \frac{q}{m}(\mathbf{v}\times\delta\mathbf{B})\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial}{\partial \alpha}\right) \\
&-\frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial}{\partial \alpha}\right).
\end{aligned} \tag{56}$$

Next, let us simplify the left-hand side of Eq. (55). Note that

$$\begin{aligned}
&\frac{q}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\left(\frac{\mathbf{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \\
=&\frac{q}{m}(\mathbf{v}\times\mathbf{B}_0)\cdot\frac{\boldsymbol{e}_\parallel\times\mathbf{v}_\perp}{v_\perp^2}\frac{\partial f_g}{\partial \alpha} \\
=&-\Omega\frac{\partial f_g}{\partial \alpha}.
\end{aligned} \tag{57}$$

Note that

$$\frac{q}{m}\mathbf{E}_0\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial \mathbf{X}} = c\left(\frac{\mathbf{E}_0\times\boldsymbol{e}_\parallel}{B_0}\right)\cdot\frac{\partial f_g}{\partial \mathbf{X}} = \mathbf{v}_{\mathbf{E}0}\cdot\frac{\partial f_g}{\partial \mathbf{X}}, \tag{58}$$

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 $\mathbf{E}_0\times\mathbf{B}_0$ drift. Further note that

$$\begin{aligned}
\frac{q}{m}\frac{\mathbf{v}\times\mathbf{B}_0}{c}\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial \mathbf{X}} \ &=\ [(\mathbf{v}\times\mathbf{e}_\parallel)\times\boldsymbol{e}_\parallel]\cdot\frac{\partial f_g}{\partial \mathbf{X}} \\
&=\ [v_\parallel\mathbf{e}_\parallel - \mathbf{v}]\cdot\frac{\partial f_g}{\partial \mathbf{X}},
\end{aligned} \tag{59}$$

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

Using Eqs. (58), (59), and (57), the left-hand side of equation (55) is written as

$$\begin{aligned}
&\frac{\partial f_g}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_{\mathbf{E}0})\cdot\frac{\partial f_g}{\partial \mathbf{X}} + \mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}] f_g - \Omega\frac{\partial f_g}{\partial \alpha} \\
+\ &\frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial \alpha}\right) \equiv L_g f_g,
\end{aligned} \tag{60}$$

where L_g is often called the unperturbed Vlasov propagator in guiding-center coordinates $(\mathbf{X}, \varepsilon, \mu, \alpha)$.

[Equation (60), corresponds to Eq. (7) in Frieman-Chen's paper[2]. In Frieman-Chen's equation (7), there is a term

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

where $\mathbf{E}_{\rm mac}$ is a given macroscopic electric field introduced when defining the guiding-center transformation. In my derivation $\mathbf{E}_{\rm mac}$ is chosen to be equal to the equilibrium electric field, and thus the above term is zero.]

Using the above results, Eq. (55) is written as

$$L_g f_g = \delta R f_g, \tag{61}$$

i.e.

$$\begin{aligned}
&\frac{\partial f_g}{\partial t}+(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_{\mathbf{E}0})\cdot\frac{\partial f_g}{\partial\mathbf{X}}-\Omega\frac{\partial f_g}{\partial\alpha}\\
&+\mathbf{v}\cdot\left[\mathbf{v}\times\frac{\partial}{\partial\mathbf{x}}\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}}+\frac{\partial\mu}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\mu}+\frac{\partial\alpha}{\partial\mathbf{x}}\frac{\partial f_g}{\partial\alpha}\right]\\
&+\frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right)\\
&=-\frac{q}{m}(\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}}-\frac{q}{m}(\mathbf{v}\times\delta\mathbf{B})\cdot\left(\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right)\\
&-\frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial f_g}{\partial\varepsilon}+\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right).
\end{aligned}\tag{62}$$

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 (62) is simplified as

$$\begin{aligned}
&\frac{\partial f_g}{\partial t}+v_\parallel\mathbf{e}_\parallel\cdot\frac{\partial f_g}{\partial\mathbf{X}}-\Omega\frac{\partial f_g}{\partial\alpha}\\
&=-\frac{q}{m}(\delta\mathbf{E})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial f_g}{\partial\mathbf{X}} &(63)\\
&-\frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial f_g}{\partial\varepsilon}+\frac{\mathbf{v}_\perp}{B_0}\frac{\partial f_g}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial f_g}{\partial\alpha}\right) &(64)
\end{aligned}$$

If neglecting the $\delta\mathbf{E}$ perturbation, the above equation reduces to

$$\frac{\partial f_g}{\partial t}+v_\parallel\mathbf{e}_\parallel\cdot\frac{\partial f_g}{\partial\mathbf{X}}-\Omega\frac{\partial f_g}{\partial\alpha}=0, \tag{65}$$

which agrees with Eq. (21) discussed in Sec. 1.2.

3.2 Distribution function perturbation

Expand the distribution function f_g as

$$f_g = F_g+\delta F_g, \tag{66}$$

where F_g is assumed to be an equilibrium distribution function, i.e.,

$$L_g F_g = 0. \tag{67}$$

Using Eqs. (66) and (67) in Eq. (61), we obtain an equation for δF_g:

$$L_g \delta F_g = \delta R F_g + \delta R \delta F_g. \tag{68}$$

3.3 Gyrokinetic ordering

To facilitate the simplification of the Vlasov equation in the low-frequency regime, we assume the following orderings (some of which are roughly based on experiment measure of fluctuations responsible for tokamak plasma transport, some of which can be invalid in some interesting cases.) These ordering are often called the standard gyrokinetic orderings.

3.3.1 Assumptions for equilibrium quantities

Define the spatial scale length L_0 of equilibrium quantities by $L_0 \approx F_g / |\nabla_X F_g|$. Assume that L_0 is much larger than the thermal gyro-radius $\rho_i \equiv v_t/\Omega$, i.e., $\lambda \equiv \rho_i / L_0$ is a small parameter, where $v_t = \sqrt{2T/m}$ is the thermal velocity. That is

$$\frac{1}{F_g} \rho_i |\nabla_X F_g| \sim O(\lambda^1), \tag{69}$$

and

$$\frac{1}{B_0} \rho_i |\nabla_X B_0| \sim O(\lambda^1). \tag{70}$$

The equilibrium $\mathbf{E}_0 \times \mathbf{B}_0$ flow, which is given by

$$\mathbf{v}_{\mathbf{E}0} = \mathbf{E}_0 \times \mathbf{e}_\parallel / B_0 = -\nabla \Phi_0 \times \mathbf{e}_\parallel / B_0, \tag{71}$$

is assumed to be weak with

$$\frac{|\mathbf{v}_{E0}|}{v_t} \sim O(\lambda^1). \tag{72}$$

3.3.2 Assumptions for perturbations

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

$$\delta \mathbf{B} = \nabla \times \delta \mathbf{A}, \tag{73}$$

and

$$\delta \mathbf{E} = -\nabla \delta\Phi - \frac{\partial \delta \mathbf{A}}{\partial t}. \tag{74}$$

Most gyrokinetic simulations approximate the vector potential as $\delta\mathbf{A} \approx \delta A_\parallel \mathbf{e}_\parallel$. Then Eq. (73) is writen as

$$\delta\mathbf{B} = \nabla \times (\delta A_\parallel \mathbf{e}_\parallel) \approx \nabla \delta A_\parallel \times \mathbf{e}_\parallel. \tag{75}$$

We assume that the amplitudes of perturbations are small:

$$\frac{\delta F_g}{F_0} \sim \frac{q\delta\Phi}{T} \sim \frac{\delta B}{B_0} \sim O(\lambda^1). \tag{76}$$

Further we assume that the parallel wavelength of perturbations is much longer than ρ_i:

$$\frac{1}{\delta F_g} |\rho_i \mathbf{e}_\parallel \cdot \nabla_X \delta F_g| \sim O(\lambda^1), \tag{77}$$

and pependicular wavelength is of the same oder of ρ_i:

$$\frac{1}{\delta F_g} |\rho_i \nabla_{X\perp} \delta F_g| \sim O(\lambda^0). \tag{78}$$

Equation (75) indicates that $\delta B \sim \delta A_\parallel / \rho_i$. Then the odering in (76) indicates that $v_t \delta A_\parallel$ is of the same order of $\delta\Phi$.

The electrical field perturbation is written as

$$\delta E_\parallel = \delta\mathbf{E} \cdot \mathbf{e}_\parallel = -\mathbf{e}_\parallel \cdot \nabla\delta\Phi - \frac{\partial \delta A_\parallel}{\partial t}, \tag{79}$$

$$\delta\mathbf{E}_\perp = -\nabla_\perp \delta\Phi. \tag{80}$$

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

$$\frac{\delta E_\parallel}{\delta E_\perp} = O(\lambda^1). \tag{81}$$

We assume the perturbations are of low frequency: $\omega / \Omega \sim O(\lambda^1)$, i.e.,

$$\frac{1}{\delta F_g}\frac{1}{\Omega}\frac{\partial \delta F_g}{\partial t} \sim \frac{1}{\delta\Phi}\frac{1}{\Omega}\frac{\partial \delta\Phi}{\partial t} \sim \frac{1}{\delta A_\parallel}\frac{1}{\Omega}\frac{\partial \delta A_\parallel}{\partial t} \sim O(\lambda^1). \tag{82}$$

3.4 Equation for macroscopic distribution function F_g

The evolution of the macroscopic quantity F_g is governed by Eq. (67), i.e.,

$$L_g F_g = 0, \tag{83}$$

where the left-hand side is written as

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

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

$$\frac{\partial F_{g0}}{\partial \alpha} = 0 \tag{84}$$

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

$$v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial F_{g0}}{\partial \mu}\right) = \Omega\frac{\partial F_{g1}}{\partial \alpha}. \tag{85}$$

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

$$\left(\int_0^{2\pi} d\alpha v_\parallel \mathbf{e}_\parallel\right)\cdot\frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m}\frac{\partial F_{g0}}{\partial \mu}\int_0^{2\pi} d\alpha \mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\right) = \int_0^{2\pi} d\alpha \Omega\frac{\partial F_{g1}}{\partial \alpha} \tag{86}$$

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, $\mathbf{e}_\parallel$, B_0 and Ω can be considered to be independent of α. As to $v_\parallel$, we have $v_\parallel = \pm\sqrt{2(\varepsilon - B_0\mu)}$. Since B_0 is considered independent of α, so does $v_\parallel$. Using these results, equation (86) is written

$$v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m}\frac{\partial F_{g0}}{\partial \mu}\int_0^{2\pi} d\alpha \mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\right) = 0. \tag{87}$$

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

$$v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} + \frac{q}{m}\frac{\partial F_{g0}}{\partial \mu}\int_0^{2\pi} d\alpha\left(\frac{-\mathbf{v}_\perp\cdot\nabla\Phi_0}{B_0}\right) = 0, \tag{88}$$

Note that

$$\int_0^{2\pi} d\alpha \frac{1}{B_0}\mathbf{v}_\perp\cdot\nabla_X\Phi_0 \approx 0, \tag{89}$$

where the error is of $O(\lambda^2)\Phi_0$, and thus, accurate to $O(\lambda)$, the last term of equation (88) is zero. Then equation (88) is written as

$$v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} = 0, \tag{90}$$

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

3.5 Equation for δF_g

Using $F_g \approx F_{g0}$, equation (68) is written as

$$L_g \delta F_g = \delta R F_{g0} + \underbrace{\delta R \delta F_g}_{\text{NonlinearTerm} \sim O(\lambda^2)}, \tag{91}$$

where $\delta R \delta F_g$ is a nonlinear term which is 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

$$\delta R F_{g0} = \underbrace{-\frac{q}{m}\left(\delta \mathbf{E} + \frac{\mathbf{v}\times\delta\mathbf{B}}{c}\right)\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial F_{g0}}{\partial \mathbf{X}}}_{O(\lambda^2)} - \underbrace{\frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial F_{g0}}{\partial\varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial F_{g0}}{\partial\mu}\right)}_{O(\lambda^1)}, \tag{92}$$

In obtaining (92), use has been made of $\partial F_{g0}/\partial\alpha = 0$. Another linear term $L_g \delta F_g$ is written as

$$\begin{aligned} L_g \delta F_g &= \frac{\partial \delta F_g}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_{\mathbf{E}0})\cdot\frac{\partial \delta F_g}{\partial \mathbf{X}} + \mathbf{v}\cdot[\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}]\delta F_g - \underbrace{\Omega\frac{\partial \delta F_g}{\partial\alpha}}_{O(\lambda^1)} \\ &+ \frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial \delta F_g}{\partial\mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial \delta F_g}{\partial\alpha}\right), \end{aligned} \tag{93}$$

where $\Omega\partial\delta F_g/\partial\alpha$ 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 $\partial F_{g0}/\partial\mu = 0$.

3.5.1 Balance on order $O(\lambda^1)$: adiabatic response

The balance between the leading terms (terms of $O(\lambda)$) in Eq. (91) requires that

$$\Omega\frac{\partial \delta F_a}{\partial\alpha} = \frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial F_{g0}}{\partial\varepsilon}\right), \tag{94}$$

where δF_a is a unknown distribution function to be solved from the above equation. It is ready to verify that

$$\delta F_a = \frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}, \tag{95}$$

is a solution to the above equation, accurate to $O(\lambda)$. [Proof: Substitute expression (95) into the left-hand side of Eq. (94), we obtain

$$\begin{aligned} \Omega\frac{\partial \delta F_a}{\partial\alpha} &= \Omega\frac{\partial}{\partial\alpha}\left(\frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}\right) \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\Omega\frac{\partial}{\partial\alpha}(\delta\Phi) \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial\varepsilon}\Omega(\nabla_x\delta\Phi)\cdot\frac{\partial\mathbf{x}}{\partial\alpha} \end{aligned} \tag{96}$$

Using

$$\begin{aligned}\frac{\partial \mathbf{x}}{\partial \alpha} &= \frac{\partial}{\partial \alpha}\left[-\mathbf{v}\times\frac{\boldsymbol{e}_{\parallel}(\mathbf{X})}{\Omega(\mathbf{X})}\right] \\ &= \frac{\partial}{\partial \alpha}[-\mathbf{v}]\times\frac{\boldsymbol{e}_{\parallel}(\mathbf{X})}{\Omega(\mathbf{X})} \\ &= -\frac{\mathbf{v}_{\perp}}{\Omega}\end{aligned} \tag{97}$$

Eq. (96) is written as

$$\begin{aligned}\Omega\frac{\partial \delta F_a}{\partial \alpha} &= -\frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\Omega(\nabla_x \delta\Phi)\cdot\frac{\mathbf{v}_{\perp}}{\Omega} \\ &= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left(\delta\mathbf{E}+\frac{\partial \delta\mathbf{A}}{\partial t}\right)\cdot\mathbf{v}_{\perp} \qquad (98)\\ &\approx \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}(\delta\mathbf{E})\cdot\mathbf{v}_{\perp}, \qquad (99)\end{aligned}$$

where terms of $O(\lambda^2)$ have been dropped. Similarly, dropping the parallel electric field term (which is of $O(\lambda^2)$) on the right-hand side of Eq. (94), we find it is identical to the right-hand side of Eq. (99)]

3.5.2 Separate δF_g into adiabatic and non-adiabatic part

As is discussed above, the terms of $O(\lambda)$ can be eliminated by splitting a so-called adiabatic term form δF_g. Specifically, write δF_g as

$$\delta F_g = \delta F_a + \delta G, \tag{100}$$

where δF_a is given by (95), i.e.,

$$\delta F_a = \frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial \varepsilon}, \tag{101}$$

which depends on the gyro-angle via $\delta\Phi$ and this term is often called adiabatic term. Plugging expression (100) into equation (91), we obtain

$$L_g\delta G = \underbrace{\delta R F_{g0} - L_g\delta F_a}_{\text{LinearTerms}} + \underbrace{\delta R\delta F_g}_{\text{NonlinearTerms}}. \tag{102}$$

Next, let us simplify the linear term on the right-hand side, i.e, $\delta R F_{g0} - L_{gf}\delta F_a$, (which should be of $O(\lambda^2)$ or higher because $\Omega\partial\delta F_a/\partial\alpha$ cancels all the $O(\lambda^1)$ terms in $\delta R F_{g0}$).

$L_g\delta F_a$ is written

$$\begin{aligned}L_g\delta F_a &= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}L_g\delta\Phi + \frac{q}{m}\delta\Phi L_g\frac{\partial F_{g0}}{\partial \varepsilon} \\ &\approx \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}L_g\delta\Phi,\end{aligned} \tag{103}$$

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 (39) and (40). The expression (103) involves $\delta\Phi$ operated by the Vlasov propagator L_g. 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_g expressed in particle coordinates. Transforming L_g back to the particle coordinates, expression (103) is written

$$
\begin{aligned}
L_g\delta F_a &= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} + \mathbf{v}\cdot\nabla_{\mathbf{x}}\delta\Phi + \frac{q}{m}(\mathbf{E}_0 + \mathbf{v}\times\mathbf{B}_0)\cdot\frac{\partial \Phi}{\partial \mathbf{v}}|_{\mathbf{x}}\right] \\
&= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} + \mathbf{v}\cdot\nabla_{\mathbf{x}}\delta\Phi\right] \qquad (104)\\
&= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} + \mathbf{v}\cdot\left(-\delta\mathbf{E} - \frac{\partial \delta\mathbf{A}}{\partial t}|_{\mathbf{x},\mathbf{v}}\right)\right] \\
&= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \delta\Phi}{\partial t}|_{\mathbf{x},\mathbf{v}} - \mathbf{v}\cdot\delta\mathbf{E} - \frac{\partial \mathbf{v}\cdot\delta\mathbf{A}}{\partial t}|_{\mathbf{x},\mathbf{v}}\right]. \qquad (105)\\
&= \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \delta\Phi}{\partial t} - \mathbf{v}\cdot\delta\mathbf{E} - \frac{\partial \mathbf{v}\cdot\delta\mathbf{A}}{\partial t}\right]. \qquad (106)
\end{aligned}
$$

Using this and expression (92), $\delta R F_{g0} - L_g\delta F_a$ is written as

$$
\begin{aligned}
\delta R F_{g0} - L_g\delta F_a &= -\frac{q}{m}(\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial F_{g0}}{\partial \mathbf{X}} - \frac{q}{m}\delta\mathbf{E}\cdot\left(\mathbf{v}\frac{\partial F_{g0}}{\partial \varepsilon}\right) \\
&\quad - \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \delta\Phi}{\partial t} - \mathbf{v}\cdot\delta\mathbf{E} - \frac{\partial \mathbf{v}\cdot\delta\mathbf{A}}{\partial t}\right] \\
&= -\frac{q}{m}(\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\frac{\partial F_{g0}}{\partial \mathbf{X}} - \frac{q}{m}\frac{\partial F_{g0}}{\partial \varepsilon}\left[\frac{\partial \Phi}{\partial t} - \frac{\partial \mathbf{v}\cdot\delta\mathbf{A}}{\partial t}\right], \qquad (107)
\end{aligned}
$$

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. (102).

The consequence of this is that, as we will see in Sec. 3.6.1, δG is independent of the gyro-angle, accurate to order $O(\lambda^1)$. Therefore, separating δF into adiabatic and non-adiabatic parts also corresponds to separating δF into gyro-angle dependent and gyro-angle independent parts.

3.5.3 Linear term expressed in terms of $\delta\Phi$ and $\delta\mathbf{A}$

Let us rewrite the linear term (107) in terms of $\delta\Phi$ and $\delta\mathbf{A}$. The $\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B}$ term in expression (107) is written as

$$
\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B} = -\nabla_x\delta\Phi - \frac{\partial \delta\mathbf{A}}{\partial t} + \mathbf{v}\times\nabla_x\times\delta\mathbf{A}. \qquad (108)
$$

Note that this term needs to be accurate to only $O(\lambda)$. Then

$$\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B}\approx-\nabla_x\delta\Phi+\mathbf{v}\times\nabla_x\times\delta\mathbf{A}, \tag{109}$$

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

$$\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B}=-\nabla_x\delta\Phi+\nabla_x(\delta\mathbf{A}\cdot\mathbf{v})-(\mathbf{v}\cdot\nabla_x)\delta\mathbf{A} \tag{110}$$

Note that Eq. (41) 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

$$\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B}=-\nabla_X\delta\Phi+\nabla_X(\delta\mathbf{A}\cdot\mathbf{v})-(\mathbf{v}\cdot\nabla_X)\delta\mathbf{A} \tag{111}$$

Further note that the parallel gradients in the above equation are of $O(\lambda^2)$ and thus can be dropped. Then expression (111) is written

$$\begin{aligned}&\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B}\\&=-\nabla_{X\perp}\delta\Phi+\nabla_{X\perp}(\delta\mathbf{A}\cdot\mathbf{v})-(\mathbf{v}_\perp\cdot\nabla_{X\perp})\delta\mathbf{A}.\\&=-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_{X\perp}\delta\mathbf{A},\end{aligned} \tag{112}$$

where δL is defined by

$$\delta L=\delta\Phi-\mathbf{v}\cdot\delta\mathbf{A}. \tag{113}$$

Using expression (112), equation (107) is written

$$\delta RF_0-L_g\delta F_a=-\frac{q}{m}\Big[(-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_{X\perp}\delta\mathbf{A})\times\frac{\boldsymbol{e}_\parallel}{\Omega}\Big]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\delta L}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}, \tag{114}$$

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

3.6 Equation for the non-adiabatic part δG

Plugging expression (114) into Eq. (102), we obtain

$$L_g\delta G=-\frac{q}{m}\Big[(-\nabla_{X\perp}\delta L-\mathbf{v}_\perp\cdot\nabla_{X\perp}\delta\mathbf{A})\times\frac{\boldsymbol{e}_\parallel}{\Omega}\Big]\cdot\frac{\partial F_{g0}}{\partial\mathbf{X}}-\frac{q}{m}\frac{\partial\delta L}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}+\delta R\delta F_g, \tag{115}$$

where L_g is given by Eq. (93), i.e.,

$$\begin{aligned}L_g\;=\;&\frac{\partial}{\partial t}+(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_{\mathbf{E}0})\cdot\frac{\partial}{\partial\mathbf{X}}+\mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}]-\Omega\frac{\partial}{\partial\alpha}\\&+\;\frac{q}{m}\mathbf{E}_0\cdot\left(\frac{\mathbf{v}_\perp}{B_0}\frac{\partial}{\partial\mu}+\frac{\boldsymbol{e}_\alpha}{v_\perp}\frac{\partial}{\partial\alpha}\right),\end{aligned} \tag{116}$$

3.6.1 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. (115) is of $O(\lambda^2)$, then, the balance on order $O(\lambda^1)$ requires

$$\frac{\partial \delta G_0}{\partial \alpha} = 0, \tag{117}$$

i.e., δG_0 is gyro-phase independent.

The balance on order $O(\lambda^2)$ requires (for the special case of $\mathbf{E}_0 = 0$):

$$\begin{aligned} &\frac{\partial \delta G_0}{\partial t} + v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} + \mathbf{v} \cdot [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}]\delta G_0 \\ = \; & -\frac{q}{m}\Big[(-\nabla_{X\perp}\delta L - \mathbf{v}_\perp \cdot \nabla_X \delta \mathbf{A}) \times \frac{\boldsymbol{e}_\parallel}{\Omega}\Big] \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} - \frac{q}{m}\frac{\partial \delta L}{\partial t}\frac{\partial F_{g0}}{\partial \varepsilon} + \delta R \delta F_g. \end{aligned} \tag{118}$$

3.6.2 Gyro-averaging

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

$$\langle h \rangle_\alpha = (2\pi)^{-1} \int_0^{2\pi} h \, d\alpha, \tag{119}$$

where $h = h(\mathbf{X}, \alpha, \varepsilon, \mu)$ is an arbitrary function of guiding-center variables. The gyro-averaging is an integration in the velocity space. [For a field quantity, which 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. (118), we obtain

$$\begin{aligned} &\frac{\partial \delta G_0}{\partial t} + \left\langle v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} \right\rangle + \langle \mathbf{v} \cdot [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}]\delta G_0 \rangle_\alpha \\ = & -\frac{q}{m}\Big[-\nabla_{X\perp}\langle \delta L \rangle_\alpha \times \frac{\boldsymbol{e}_\parallel}{\Omega}\Big] \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} - \frac{q}{m}\frac{\partial \langle \delta L \rangle_\alpha}{\partial t}\frac{\partial F_{g0}}{\partial \varepsilon} + \langle \delta R \delta F_g \rangle_\alpha, \end{aligned} \tag{120}$$

where use has been made of $\langle (\mathbf{v}_\perp \cdot \nabla_X)\delta \mathbf{A} \rangle_\alpha \approx 0$, 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

$$\left\langle v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} \right\rangle_\alpha = \langle v_\parallel \mathbf{e}_\parallel \rangle \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} = v_\parallel \mathbf{e}_\parallel \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} \tag{121}$$

The second gyro-averaging on the left-hand side of Eq. (120) can be written as

$$\langle \mathbf{v} \cdot [\boldsymbol{\lambda}_{B1} + \boldsymbol{\lambda}_{B2}]\delta G_0 \rangle_\alpha = \mathbf{V}_D \cdot \nabla_X \delta G_0, \tag{122}$$

where $\mathbf{V}_D$ is the magnetic curvature and gradient drift (Eq. (122) is derived in Appendix xx, to do later). Then Eq. (120) is written

$$\begin{aligned} &\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D) \cdot \nabla_X\right]\delta G_0 \\ = & -\frac{q}{m}\Big[-\nabla_{X\perp}\langle \delta L \rangle_\alpha \times \frac{\boldsymbol{e}_\parallel}{\Omega}\Big] \cdot \frac{\partial F_{g0}}{\partial \mathbf{X}} - \frac{q}{m}\frac{\partial \langle \delta L \rangle_\alpha}{\partial t}\frac{\partial F_{g0}}{\partial \varepsilon} + \langle \delta R \delta F_g \rangle_\alpha. \end{aligned} \tag{123}$$

3.6.3 Simplification of the nonlinear term

Next, we try to simplify the nonlinear term $\langle \delta R \delta F_g \rangle_\alpha$ appearing in Eq. (123), which is written as

$$\begin{aligned} \langle \delta R \delta F_g \rangle_\alpha &= \left\langle \delta R \left(\frac{q}{m} \delta\Phi \frac{\partial F_{g0}}{\partial \varepsilon} + \delta G_0 \right) \right\rangle_\alpha \\ &= \left\langle \frac{q}{m} \delta R \left(\delta\Phi \frac{\partial F_{g0}}{\partial \varepsilon} \right) \right\rangle_\alpha + \langle \delta R \delta G_0 \rangle_\alpha \end{aligned} \tag{124}$$

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

$$\begin{aligned} \delta R \left(\delta\Phi \frac{\partial F_{g0}}{\partial \varepsilon} \right) &\approx -\frac{q}{m} \frac{\partial F_{g0}}{\partial \varepsilon} \left(\delta \mathbf{E} + \frac{\mathbf{v} \times \delta \mathbf{B}}{c} \right) \times \left(\frac{\boldsymbol{e}_\parallel}{\Omega} \right) \cdot \frac{\partial \delta\Phi}{\partial \mathbf{X}} - \frac{q}{m} \frac{\partial F_{g0}}{\partial \varepsilon} \left(\frac{\mathbf{v} \times \delta \mathbf{B}}{c} \right) \cdot \\ &\quad \left(\frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial \delta\Phi}{\partial \alpha} \right) \\ &\quad - \frac{q}{m} \frac{\partial F_{g0}}{\partial \varepsilon} \delta \mathbf{E} \cdot \left(\mathbf{v} \frac{\partial \delta\Phi}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0} \frac{\partial \delta\Phi}{\partial \mu} + \frac{\boldsymbol{e}_\alpha}{v_\perp} \frac{\partial \delta\Phi}{\partial \alpha} \right) + \frac{q}{m} \delta\Phi \delta \mathbf{E} \cdot \mathbf{v} \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \\ &= -\frac{q}{m} \left(\delta \mathbf{E} + \frac{\mathbf{v} \times \delta \mathbf{B}}{c} \right) \cdot \nabla_v \delta\Phi + \frac{q}{m} \delta\Phi \delta \mathbf{E} \cdot \mathbf{v} \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \\ &= \frac{q}{m} \delta\Phi \delta \mathbf{E} \cdot \mathbf{v} \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \end{aligned} \tag{125}$$

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

$$\langle \delta R \delta F \rangle_\alpha = \frac{q}{m} \left\langle \delta\Phi \delta \mathbf{E} \cdot \mathbf{v} \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \right\rangle_\alpha + \langle \delta R \delta G_0 \rangle_\alpha \tag{126}$$

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

$$\begin{aligned} \left\langle \delta\Phi \delta \mathbf{E} \cdot \mathbf{v} \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \right\rangle_\alpha &= \left\langle \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \delta\Phi \nabla \delta\Phi \cdot \mathbf{v} \right\rangle_\alpha \\ &= \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \langle \mathbf{v} \cdot \nabla (\delta\Phi)^2 \rangle_\alpha \\ &\approx \frac{\partial^2 F_{g0}}{\partial \varepsilon^2} \langle \mathbf{v}_\perp \cdot \nabla (\delta\Phi)^2 \rangle_\alpha \\ &\approx 0, \end{aligned} \tag{127}$$

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 (126) is written as

$$\langle \delta R \delta F_g \rangle_\alpha = \langle \delta R \delta G_0 \rangle_\alpha. \tag{128}$$

Using the expression of δR given by Eq. (56), the above expression is written as

$$\begin{aligned} \langle \delta R \delta G_0 \rangle_\alpha &= -\frac{q}{m} \left\langle \left(\delta \mathbf{E} + \frac{\mathbf{v} \times \delta \mathbf{B}}{c} \right) \times \left(\frac{\boldsymbol{e}_\parallel}{\Omega} \right) \right\rangle_\alpha \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} \\ &\quad - \frac{q}{m} \frac{\partial \delta G_0}{\partial \varepsilon} \langle \delta \mathbf{E} \cdot \mathbf{v} \rangle_\alpha - \frac{q}{m} \frac{\partial \delta G_0}{\partial \mu} \left\langle \delta \mathbf{E} \cdot \frac{\mathbf{v}_\perp}{B_0} \right\rangle_\alpha \end{aligned} \tag{129}$$

where use has been made of $\partial\delta G_0/\partial\alpha=0$. Using Eq. (112), we obtain

$$-\frac{q}{m}\left\langle(\delta\mathbf{E}+\mathbf{v}\times\delta\mathbf{B})\times\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\right\rangle_\alpha=\frac{q}{m}\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}. \tag{130}$$

The other two terms in Eq. (129) can be proved to be zero. [Proof:

$$\begin{aligned}-\frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\delta\mathbf{E}\cdot\mathbf{v}\rangle_\alpha &= \frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\mathbf{v}\cdot\nabla_x\Phi\rangle_\alpha\\ &\approx \frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\mathbf{v}_\perp\cdot\nabla_x\Phi\rangle_\alpha\\ &\approx \frac{q}{m}\frac{\partial\delta G_0}{\partial\varepsilon}\langle\mathbf{v}_\perp\cdot\nabla_X\Phi\rangle_\alpha\\ &\approx 0\end{aligned} \tag{131}$$

$$\begin{aligned}-\frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\delta\mathbf{E}\cdot\frac{\mathbf{v}_\perp}{B_0}\right\rangle_\alpha &= \frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\frac{1}{B_0}\mathbf{v}_\perp\cdot\nabla_x\Phi\right\rangle_\alpha\\ &\approx \frac{q}{m}\frac{\partial\delta G_0}{\partial\mu}\left\langle\frac{1}{B_0}\mathbf{v}_\perp\cdot\nabla_X\Phi\right\rangle_\alpha\\ &\approx 0\end{aligned} \tag{132}$$

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

$$\langle\delta R\delta G_0\rangle_\alpha=\frac{q}{m}\left[\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right]\cdot\nabla_X\delta G_0. \tag{133}$$

Using this in Eq. (128), we obtain

$$\langle\delta R\delta F_g\rangle_\alpha=\frac{q}{m}\left[\nabla_{X\perp}\langle\delta L\rangle_\alpha\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right]\cdot\nabla_X\delta G_0, \tag{134}$$

which is of $O(\lambda^2)$.

3.6.4 Final equation for the non-adiabatic part of distribution function perturbation

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

$$\begin{aligned}&\frac{\partial\delta G_0}{\partial t}+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\underbrace{\frac{q}{m}\nabla\langle\delta L\rangle\times\frac{\boldsymbol{e}_\parallel}{\Omega}}_{\text{nonlinear}}\right)\cdot\nabla\delta G_0\\ &=\underbrace{\left(\frac{q}{m}\nabla\langle\delta L\rangle\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\nabla F_{g0}}_{\text{spatial-drive}}-\underbrace{\frac{q}{m}\frac{\partial\langle\delta L\rangle}{\partial t}\frac{\partial F_{g0}}{\partial\varepsilon}}_{\text{velocit-space-damp}}.\end{aligned} \tag{135}$$

where $\delta G_0=\delta G_0(\mathbf{X},\varepsilon,\mu,t)$ is gyro-angle independent and is related to the distribution function perturbation δF_g by

$$\delta F_g=\frac{q}{m}\delta\Phi\frac{\partial F_{g0}}{\partial\varepsilon}+\delta G_0, \tag{136}$$

where the first term is called "adiabatic part", which depends on the gyro-phase α via $\delta\Phi$. (Equation (135) is the special case ($\partial F_{g0}/\partial\mu|_\varepsilon = 0$) of the Frieman-Chen nonlinear gyrokinetic equation given in Ref. [2]. Note that the nonlinear terms only appear on the left-hand side of Eq. (135) and all the terms on the right-hand side are linear.) Here $\mathbf{V}_D$ is the guiding-center drift velocity in the equilibrim field; $\langle ...\rangle$ is the gyro-phase averaging operator; $\delta L = \delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}$; the term

$$-\frac{q}{m}\nabla\langle\delta L\rangle\times\frac{\boldsymbol{e}_\parallel}{\Omega} \tag{137}$$

consists of the $\delta\mathbf{E}\times\mathbf{B}_0$ drift and magnetic fluttering term (refer to expression (338) in Sec. C.3). For notaiton ease, this term is denoted by $\delta\mathbf{V}_D$ in the following:

$$\delta\mathbf{V}_D \equiv -\frac{q}{m}\nabla\langle\delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}\rangle\times\frac{\mathbf{e}_\parallel}{\Omega}. \tag{138}$$

Note that $\nabla\equiv\partial/\partial\mathbf{X}$, which is the gradient operator in the guiding-center coordinates $(\mathbf{X},\varepsilon,\mu,\alpha)$ while holding (ε,μ,α) constant. How do we numerically calculate $\partial\Phi_g/\partial\mathbf{X}$ in PIC simulations? The difficulty is that $\mathbf{X}$ grid is not available, which makes it difficult to use the finite difference method. Meanwhile $\mathbf{x}$ grid is available in PIC simulations. It turns out we can make use of $\mathbf{x}$ grid to approximatly calculate the derivatives in $\mathbf{X}$ space. Using $\mathbf{x}=\mathbf{X}+\boldsymbol{\rho}$ and the definition $\delta\Phi_g(\mathbf{X},\alpha,v_\perp)=\delta\Phi_p(\mathbf{x})$ we obtain

$$\begin{aligned}\frac{\partial\delta\Phi_g(\mathbf{X},\alpha,v_\perp)}{\partial\mathbf{X}} &= \frac{\partial\delta\Phi_p(\mathbf{x})}{\partial\mathbf{X}}\\ &= \frac{\partial\delta\Phi_p(\mathbf{x})}{\partial\mathbf{x}}\cdot\frac{\partial\mathbf{x}}{\partial\mathbf{X}}\\ &= \frac{\partial\delta\Phi_p(\mathbf{x})}{\partial\mathbf{x}}\cdot\left(1+\frac{\partial\boldsymbol{\rho}}{\partial\mathbf{X}}\right)\\ &\approx \frac{\partial\delta\Phi_p(\mathbf{x})}{\partial\mathbf{x}},\end{aligned} \tag{139}$$

which is accurate up to $O(\lambda)$ in gyrokinetic ordering. The above relation indicates that the value of $\partial\delta\Phi_g/\partial\mathbf{X}$ at a guidng-center location is approximately equal to the value of $\partial\delta\Phi_p/\partial\mathbf{x}$ at the corresponding particle position. Therefore we can use $\delta\Phi$ defined on $\mathbf{x}$ grid to compute $\partial\delta\Phi_p/\partial\mathbf{x}$ on gridpoints, and interpolate it to the particle position and this is a good approximation of $\partial\delta\Phi_g/\partial\mathbf{X}$.

Also note that $\langle\delta\Phi_g\rangle$ is not known on $\mathbf{X}$ grid (it is known at random markers). So, to calculate $\partial\langle\delta\Phi_g\rangle/\partial\mathbf{X}$, we need to exchange the operation order:

$$\frac{\delta\langle\delta\Phi_g\rangle}{\partial\mathbf{X}} = \langle\frac{\partial\delta\Phi_g}{\partial\mathbf{X}}\rangle \approx \langle\frac{\partial\delta\Phi_p}{\partial\mathbf{x}}\rangle. \tag{140}$$

For notation ease, the subscripts g and p are usually dropped. Which one is intended should be obvious from the context.

4 Gyrokinetic equation suitable for numerical simulation

The gyrokinetic equation (135) contains time derivatives of unknown $\delta\Phi$ and $\delta\mathbf{A}$ on the right-hand side, which is problematic if treated by using explicit finite difference in PIC simulations. Next, we discuss some methods that can eliminate these terms, making the gyrokinetic equation more amenable to PIC simulations.

4.1 Eliminate $\partial\langle\delta\phi\rangle/\partial t$ term on the right-hand side of Eq. (135)

The coefficient before $\partial F_0/\partial\varepsilon$ in Eq. (135) involves the time derivative of $\langle\delta\Phi\rangle$, which is problematic if treated by using explicit finite difference (I test the algorithm that treats this term by implicit scheme, the result roughly agrees with the standard method discussed in Sec. 5. In GEM's split-weight scheme, $\partial\langle\delta\Phi\rangle/\partial t$ is evaluated by using the vorticity equation (time derivative of the gyrokinetic Poisson equation.). It turns out that $\partial\langle\delta\Phi\rangle/\partial t$ can be eliminated by defining another gyro-phase independent function δf by

$$\delta f = \frac{q}{m}\langle\delta\Phi\rangle\frac{\partial F_0}{\partial\varepsilon} + \delta G_0. \tag{141}$$

Substituting this into Eq. (135), we obtain the equation for δf:

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\delta f \\
&-\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\langle\delta\Phi\rangle \\
&-\frac{q}{m}\langle\delta\Phi\rangle\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\frac{\partial F_0}{\partial\varepsilon} \\
&= -\delta\mathbf{V}_D\cdot\nabla F_0 - \frac{q}{m}\frac{\partial\langle\delta L\rangle}{\partial t}\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned} \tag{142}$$

Noting that $\partial F_0/\partial t = 0$, $\mathbf{e}_\parallel\cdot\nabla F_0 = 0$, $\partial F_0/\partial\varepsilon \approx -F_0 m/T$, we find that the third line of the above equation is of order $O(\lambda^3)$ (after both sides being divided by $F_0\Omega$). Therefore the third line can be dropped. Moving the second line to the right-hand side and noting that $\langle\delta L\rangle = \langle\delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}\rangle$, the above equation is written as

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\delta f \\
&= -\delta\mathbf{V}_D\cdot\nabla F_0 + \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\langle\delta\Phi\rangle\right],
\end{aligned} \tag{143}$$

where two $\partial\langle\delta\Phi\rangle/\partial t$ terms cancel each other. [Note that the right-hand side of Eq. (143) contains a nonlinear term $\delta\mathbf{V}_D\cdot\nabla_X\langle\delta\Phi\rangle$. This is different from the original Frieman-Chen equation, where all nonlinear terms appear on the left-hand side. For the electrostatic limit, this term disappears because $\delta\mathbf{V}_D$ is perpendicular to $\nabla\langle\delta\Phi\rangle$.]

Using

$$\delta\mathbf{V}_D = -\frac{q}{m}\nabla\langle\delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}\rangle \times \frac{\mathbf{e}_\parallel}{\Omega}$$

in the right-hand side of Eq. (143) yields

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\delta f \\
= \;& -\delta\mathbf{V}_D\cdot\nabla F_0 \\
+ \;& \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle}{\partial t} + \left(v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \frac{q}{m}\nabla\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle\times\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\cdot\nabla\langle\delta\Phi\rangle\right].
\end{aligned} \tag{144}$$

[Equation (144) corresponds to Eqs. (A8-A9) in Yang Chen's paper[1], where the first minus on the right-hand side of Eq. (A8) is wrong and should be replaced with q/m; one q is missing before $\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle/\partial t$ in Eq. (A9).]

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

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

$$\delta f^{(p_\parallel)} = \delta f - \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle. \tag{145}$$

Most gyrokinetic simulations approximate the vector potential as $\delta\mathbf{A}\approx\delta A_\parallel\mathbf{e}_\parallel$. Let us simplify Eq. (143) for this case. Then $\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle$ is written as

$$\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle \approx \langle v_\parallel\delta A_\parallel\rangle. \tag{146}$$

Then expression (145) is written as

$$\delta f^{(p_\parallel)} = \delta f - \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\langle v_\parallel\delta A_\parallel\rangle. \tag{147}$$

Then Eq. (143) is written in terms of $\delta f^{(p_\parallel)}$ as

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\delta f^{(p_\parallel)} \\
&+\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[\frac{\partial}{\partial t} + (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\langle v_\parallel\delta A_\parallel\rangle \\
&+\frac{q}{m}\langle v_\parallel\delta A_\parallel\rangle[(\mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla]\left(\frac{\partial F_0}{\partial\varepsilon}\right) \\
=&-\delta\mathbf{V}_D\cdot\nabla F_0 \\
&-\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\left[-\frac{\partial\langle v_\parallel\delta A_\parallel\rangle}{\partial t} - (v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\langle\delta\Phi\rangle\right],
\end{aligned} \tag{148}$$

where use has been made of that $\partial F_0/\partial t = 0$ and $\mathbf{e}_\parallel\cdot\nabla F_0 = 0$. Further noting that $v_\parallel\delta A_\parallel\sim\delta\Phi$, $q\delta\Phi/T\sim O(\lambda^1)$, $\partial F_0/\partial\varepsilon\sim -F_0 m/T$, $\mathbf{V}_D/v_t\sim O(\lambda^1)$, $\delta\mathbf{V}_D/v_t\sim O(\lambda^1)$, and $\rho\nabla F_0\sim O(\lambda^1)F_0$, we find that the red term of the above equation (after divided by

ΩF_0) is of $O(\lambda^3)$, hence can be dropped. Move the second line to the right-hand side, giving

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla\right]\delta f^{(p_\parallel)}\\&=-\delta\mathbf{V}_D\cdot\nabla F_0\\&+\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}[(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla(\langle\delta\Phi-v_\parallel\delta A_\parallel\rangle)],\end{aligned}\tag{149}$$

where $\partial\langle v_\parallel\delta A_\parallel\rangle/\partial t$ terms cancel each other. Further note that $\delta\mathbf{V}_D$, given by Eq. (138), is perpendicular to $\nabla_X\langle\delta\Phi-v_\parallel\delta A_\parallel\rangle$. Therefore the blue term in Eq. (149) is zero, then Eq. (149) simplifies to

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla\right]\delta f^{(p_\parallel)}\\&=-\delta\mathbf{V}_D\cdot\nabla F_0+\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla\langle\delta\Phi\rangle\\&-\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}[(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla\langle v_\parallel\delta A_\parallel\rangle].\end{aligned}\tag{150}$$

Note that in terms of $(\mathbf{X},\varepsilon,\mu,\alpha,\sigma)$ coordinates, $v_\parallel$ is written as

$$v_\parallel=\sigma\sqrt{2\varepsilon-2\mu B_0},\tag{151}$$

where $B_0(\mathbf{x})=B_0(\mathbf{X}+\boldsymbol{\rho})$ with $\boldsymbol{\rho}=\boldsymbol{\rho}(\mathbf{X},\varepsilon,\mu,\alpha)$. Since the scale length of B_0 is much larger than the thermal Larmor radius, $B_0(\mathbf{x})\approx B_0(\mathbf{X})$ and hence $v_\parallel$ can be approximated as a constant when gyro-angle α changes. Then $v_\parallel$ can be taken out of the gyro-averaging in expression (146), yielding

$$\langle v_\parallel\delta A_\parallel\rangle\approx v_\parallel\langle\delta A_\parallel\rangle.\tag{152}$$

Using this, the term related to $\delta A_\parallel$ in (150) can be further written as

$$\begin{aligned}(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla\langle v_\parallel\delta A_\parallel\rangle &= (v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla(v_\parallel\langle\delta A_\parallel\rangle)\\&= \langle\delta A_\parallel\rangle(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla(v_\parallel)\\&+ v_\parallel(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla\langle\delta A_\parallel\rangle.\end{aligned}\tag{153}$$

Using expression (151), $\langle\delta A_\parallel\rangle(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla(v_\parallel)$ is written as

$$\begin{aligned}\langle\delta A_\parallel\rangle(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla(v_\parallel) &\approx \langle\delta A_\parallel\rangle(v_\parallel \mathbf{e}_\parallel)\cdot\nabla(v_\parallel)\\&= \langle\delta A_\parallel\rangle(v_\parallel \mathbf{e}_\parallel)\cdot\nabla\left(\sigma\sqrt{2\varepsilon-2\mu B_0}\right)\\&= \langle\delta A_\parallel\rangle\sigma(v_\parallel \mathbf{e}_\parallel)\cdot\nabla\left(\sqrt{2\varepsilon-2\mu B_0}\right)\\&= \langle\delta A_\parallel\rangle\sigma v_\parallel\frac{-2\mu\mathbf{e}_\parallel\cdot\nabla B_0}{2\sqrt{2\varepsilon-2\mu B_0}}\\&= \langle\delta A_\parallel\rangle v_\parallel\frac{-2\mu\mathbf{e}_\parallel\cdot\nabla B_0}{2v_\parallel}\\&= -\langle\delta A_\parallel\rangle\mu\mathbf{e}_\parallel\cdot\nabla B_0.\end{aligned}\tag{154}$$

(We can also obtain $\nabla v_\parallel = -\mu\,(\nabla B_0)/v_\parallel$ by using Eq. (375).) Using the above results, equation (150) is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla\right]\delta f^{(p_\parallel)}\\ =\;& -\delta\mathbf{V}_D\cdot\nabla F_0+\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla\langle\delta\Phi\rangle\\ -\;& \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}[v_\parallel(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D)\cdot\nabla\langle\delta A_\parallel\rangle-\langle\delta A_\parallel\rangle\mu\mathbf{e}_\parallel\cdot\nabla B_0],\end{aligned}\tag{155}$$

which agrees with the so-called $p_\parallel$ formulism given in the `GEM` manual (the first line of Eq. 28).

Besides the derivation given above, the equation can also be derived by using $p_\parallel = v_\parallel + q\langle\delta A_\parallel\rangle/m$ as an independent variable (I did not verify this) and thus the name "$p_\parallel$ formulism". There is another formulism called $v_\parallel$ formulism, which uses Eq. (144) as the gyrokinetic equation to be numerically solved. The difficulty of using $v_\parallel$ formulism is that the time derivative $\partial A_\parallel/\partial t$ (in the weight evolution equation) needs to be treated by implicit schemes, otherwise it is numerical unstable[4]. On the other hand, the difficulty of using $p_\parallel$ formulism is that there is cancellation problems in Ampere's law, as we will discuss in Sec. 4.7.

4.3 Summary of distribution function split

In the above, the total distribution function F is split as

$$F = F_0 + \delta F,\tag{156}$$

where F_0 is the equilibrium distribution function and δF is the perturbation. Then δF is further split as

$$\delta F = \delta f^{(p_\parallel)} + \frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle)\frac{\partial F_0}{\partial\varepsilon} + \frac{q}{m}\langle v_\parallel\delta A_\parallel\rangle\frac{\partial F_0}{\partial\varepsilon},\tag{157}$$

where $\delta f^{(p_\parallel)}$ satisfies the gyrokinetic equation (155). In PIC simulations, $\delta f^{(p_\parallel)}$ is evolved by using markers and its moment is evaluated via Monte-Carlo integration. The blue and red terms explicitly depends on the unknown perturbed field. After being integrated in the velocity space, these two terms give the polarization density and the skin current, respectively. The polarization density is discussed in Sec. 5. The skin current is discussed in Sec. (4.4), and the so-called "cancellation problem" is discussed in Sec. 4.7.

4.4 Skin current

Let us calculate the moments of $\frac{q}{m}\langle v_\parallel\delta A_\parallel\rangle\frac{\partial F_0}{\partial\varepsilon}$ (the blue term in Eq. (157)). Denote this term by $\delta f^{(\mathrm{skin})}$. Neglect the FLR effect, then $\delta f^{(\mathrm{skin})}$ is written as

$$\delta f^{(\mathrm{skin})} = \frac{q}{m}v_\parallel\delta A_\parallel\frac{\partial F_0}{\partial\varepsilon}.\tag{158}$$

Assume that F_0 is a Maxwellian distribution:

$$F_0 = n_0\Big(\frac{m}{2\pi T}\Big)^{3/2}\exp\Big(-\frac{mv^2}{2T}\Big), \tag{159}$$

then $\frac{\partial F_0}{\partial \varepsilon} = -\frac{m}{T}F_0$, and expression (158) is written as

$$\delta f^{(\text{skin})} = -\frac{q}{T}v_\parallel \delta A_\parallel F_0. \tag{160}$$

The number density carried by $\delta f^{(\text{skin})}$ is zero. The parallel current carried by $\delta f^{(\text{skin})}$ is given by

$$\begin{aligned}\delta j_\parallel^{(\text{skin})} &= \int q v_\parallel \delta f^{(\text{skin})} d\mathbf{v} \\ &= -\int \frac{q^2}{T} v_\parallel^2 \delta A_\parallel F_0 d\mathbf{v}.\end{aligned} \tag{161}$$

Working in the spherical coordinates, then $v_\parallel = v\cos\theta$ and $d\mathbf{v} = v^2\sin\theta dv d\theta d\phi$. Then expression (161) is written as

$$\begin{aligned}\delta j_\parallel^{(\text{skin})} &= -\frac{q^2}{T}n_0\Big(\frac{m}{2\pi T}\Big)^{3/2}\delta A_\parallel \int v^2\cos^2\theta \exp\Big(-\frac{mv^2}{2T}\Big)v^2\sin\theta dv d\theta d\phi \\ &= -\frac{q^2}{T}n_0\Big(\frac{m}{2\pi T}\Big)^{3/2}\delta A_\parallel \frac{4\pi}{3}\int v^4 \exp\Big(-\frac{mv^2}{2T}\Big)dv \\ &= -\frac{q^2}{T}n_0\Big(\frac{1}{\pi}\Big)^{3/2}\frac{2T}{m}\delta A_\parallel \frac{4\pi}{3}\int x^4\exp(-x^2)dx \\ &= -\frac{q^2}{T}n_0\Big(\frac{1}{\pi}\Big)^{3/2}\frac{2T}{m}\delta A_\parallel \frac{4\pi}{3}\frac{3\sqrt{\pi}}{8} \\ &= -\frac{q^2}{m}n_0\delta A_\parallel.\end{aligned} \tag{162}$$

Using $c = 1/\sqrt{\mu_0\varepsilon_0}$ and $\omega_p^2 = n_0q^2/(m\varepsilon_0)$, the above expression can be written as

$$\mu_0\delta j_\parallel^{(\text{skin})} = -\frac{\omega_p^2}{c^2}\delta A_\parallel, \tag{163}$$

where the c/ω_p is called "skin depth" and thus this current is often called "skin current" (some authors call it "adiabatic current"). We note the skin current is inversely proportional to the particle mass. So it is contributed mainly by electrons.

Gyrokinetic simulations indicate that the skin current $\delta j_{\parallel e}^{(\text{skin})}$ is often much larger than the actual $\delta j_{\parallel}e$. This means that $\delta j_{\parallel e}^{(\text{skin})}$ nearly cancels the current carried by $\delta f_e^{(p_\parallel)}$, giving a small net current. This raises the question of whether numerical cancellation error is significant. It turns out that this error is indeed significant, which gives rise to numerical instabilities if no special treatment is used.

In TEK units:

$$\frac{1}{\lambda^2}\frac{v_u^2}{c^2}\bar{\delta} j_\parallel^{(\text{skin})} = -\frac{\omega_p^2}{c^2}\delta\overline{A}_\parallel. \tag{164}$$

(since

$$\frac{v_u q_u}{T_u}\mu_0 n_u q_u v_u = \frac{v_u q_u}{T_u}\frac{1}{c^2\varepsilon_0} n_u q_u v_u = \frac{n_u q_u^2}{T_u\varepsilon_0}\frac{v_u^2}{c^2} = \frac{1}{\lambda^2}\frac{v_u^2}{c^2}$$

)

4.5 Mixed-variable pullback method

To mitigate the skin current cancellation problem, Mishchenko et al[3] introduced the "mixed-variable pullback" method. In this method, we define $\delta A_\parallel^{(h)}$ by

$$\delta A_\parallel^{(h)} = \delta A_\parallel - \delta A_\parallel^{(s)}, \tag{165}$$

with $\delta A_\parallel^{(s)}$ determined by an evolution equation (inspired by the ideal Ohm's law):

$$\frac{\partial \delta A_\parallel^{(s)}}{\partial t} = -\mathbf{e}_\parallel \cdot \nabla \delta\Phi. \tag{166}$$

Using $\delta A_\parallel^{(h)}$, we define a new distribution function $\delta f^{(\mathrm{mv})}$ by

$$\delta f^{(\mathrm{mv})} = \delta f - \frac{q}{m}\frac{\partial F_0}{\partial \varepsilon}\langle v_\parallel \delta A_\parallel^{(h)}\rangle. \tag{167}$$

Then, starting from Eq. (143) and following the same procedure as that of Sec. 4.2, we obtain an equation for $\delta f^{(\mathrm{mv})}$:

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla\right]\delta f^{(\mathrm{mv})} \\
=&-\delta\mathbf{V}_D\cdot\nabla F_0 + \frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}(\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla\langle\delta\Phi\rangle \\
&-\frac{q}{m}\frac{\partial F_0}{\partial\varepsilon}\big[v_\parallel(v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D+\delta\mathbf{V}_D)\cdot\nabla\big\langle\delta A_\parallel^{(h)}\big\rangle - \big\langle\delta A_\parallel^{(h)}\big\rangle\mu\mathbf{e}_\parallel\cdot\nabla B_0\big].
\end{aligned} \tag{168}$$

Meanwhile, the parallel Ampere equation

$$-\nabla_\perp^2\delta A_\parallel = \mu_0\sum_j \delta J_{\parallel j}, \tag{169}$$

is written as

$$\left(\left(\sum_j \frac{\omega_{pj}^2}{c^2}\right) - \nabla_\perp^2\right)\delta A_\parallel^{(h)} = \nabla_\perp^2\delta A_\parallel^{(s)} + \mu_0\sum_j \delta J_{\parallel j}^{(\mathrm{mv})}, \tag{170}$$

where $\delta A_\parallel^{(h)}$ is the unknown to solve for, $\delta J_{\parallel j}^{(\mathrm{mv})}$ is the parallel current carried by $\delta f_j^{(\mathrm{mv})}$, where the subscript j is species index. Note that $\delta A_\parallel^{(s)}$ has been moved to the rhs because its value is already known, by solving Eq. (166), before we solve the Ampere equation for $\delta A_\parallel^{(h)}$.

In the above, a part of $\delta A_\parallel$ is solved from an evolution equation and the remainder is solved from the Ampere's law. Can this scheme help to reduce numerical error? If $A_\parallel^{(s)}$ carries the dominant part of $\delta A_\parallel$, then $\delta A_\parallel^{(h)}$ will be small, then the skin current $\delta A_\parallel^{(h)}\omega_p^2/c^2$ will be small, implying that the cancellation error will be small.

How do we ensure $A_\parallel^{(s)}$ carry the dominant part of $\delta A_\parallel$ over the entire simulation duration? In addition to a careful choice of the evolution equation for $\delta A_\parallel^{(s)}$, we have another leverage that can help $\delta A_\parallel^{(s)}$ to remain dominant: collect the whole $\delta A_\parallel$ into $\delta A_\parallel^{(s)}$ at the end of each time step:

$$\delta A_{\parallel\text{new}}^{(s)} = \delta A_{\parallel\text{old}}^{(s)} + \delta A_{\parallel\text{old}}^{(h)}. \tag{171}$$

Then, to make $\delta A_\parallel$ untouched (so that electromagnetic field remain unchanged), we set $\delta A_\parallel^{(h)}$ to zero:

$$\delta A_{\parallel\text{new}}^{(h)} = 0. \tag{172}$$

Here "old" and "new" refers to before and after the re-spliting, respectively. (This re-spliting is made at the end of each time step and does not correspond to any time evolution.) The re-splitting keeps the value of $\delta A_\parallel$ untouched and hence does not influence the electromagnetic field. Meanwhile, we need to keep δf unchanged. The definition of Eq. (167) indicates that, for a given δf, the re-splitting will make the value of $\delta f^{(\text{mv})}$ change to

$$\delta f_{\text{new}}^{(\text{mv})} = \delta f_{\text{old}}^{(\text{mv})} + \frac{q}{m}\langle v_\parallel \delta A_{\parallel\text{old}}^{(h)}\rangle \frac{\partial F_0}{\partial \varepsilon}. \tag{173}$$

This is the new initial value for $\delta f^{(\text{mv})}$. After this, the physical state of the system remains unchanged. This scheme makes $\delta A_\parallel^{(h)}$ remain small for each time step mainly because $\delta A_\parallel^{(s)}$ are set to carry all the value of $\delta A_\parallel$ at the begining of each time step. It is reasonable to assume that the variation of $\delta A_\parallel$ in a small time interval Δt is small. Denote this variation by Δ. In the best scenario, this small varaition will be captured by $\delta A_\parallel^{(s)}$ if its evlution equation is chosen wisely. In the worst scenario, the time evolution of $\delta A_\parallel^{(s)}$ may requires larger variation (than Δ) to be imposed on $\delta A_\parallel^{(h)}$. Even in this senario, $\delta A_\parallel^{(h)}$ is still of order of Δ, which is the varation of $\delta A_\parallel$ in a small time step and hence small.

Next, consider the special case that F_{g0} is local Maxwellian:

$$F_{g0}(\mathbf{X},\varepsilon) = n_0(\mathbf{X})\left(\frac{m}{2\pi T_0(\mathbf{X})}\right)^{3/2}\exp\left[-\frac{m\varepsilon}{T_0(\mathbf{X})}\right], \tag{174}$$

which is not an exact solution to the kinetic equation because $\mathbf{X}$ is not a constant of motion. We note that the radial coordinate of the guiding-center positon is approximately a constant of motion if the drift orbit width is small. So we restrict F_{g0} to depending only on ψ, i.e.,

$$F_{g0}(\mathbf{X},\varepsilon) = n_0(\psi)\left(\frac{m}{2\pi T_0(\psi)}\right)^{3/2}\exp\left[-\frac{m\varepsilon}{T_0(\psi)}\right], \tag{175}$$

Then it is straightforward to calculate $\partial F_{g0}/\partial\varepsilon$ and ∇F_{g0}:

$$\frac{\partial F_{g0}}{\partial \varepsilon} = \left(-\frac{m}{T}\right)F_{g0}, \tag{176}$$

and

$$\begin{aligned}\nabla F_{g0} &= \frac{\partial F_{g0}}{\partial n_0}\nabla n_0+\frac{\partial F_{g0}}{\partial T_0}\nabla T_0\\ &= F_{g0}\frac{\nabla n_0}{n_0}+F_{g0}\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\frac{\nabla T_0}{T_0}\\ &= F_{g0}\frac{1}{n_0}\frac{dn_0}{d\psi}\nabla\psi+F_{g0}\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\frac{1}{T_0}\frac{dT_0}{d\psi}\nabla\psi\\ &= -F_{g0}\left[\kappa_n+\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\kappa_T\right]\nabla\psi \end{aligned} \tag{177}$$

where $\kappa_n\equiv-\frac{1}{n_0}\frac{dn_0}{d\psi}$, $\kappa_T\equiv-\frac{1}{T_0}\frac{dT_0}{d\psi}$.

Using the above results, Eq. (168) is written as

$$\begin{aligned}\frac{d\delta f^{(\mathrm{mv})}}{dt} &= \delta\mathbf{V}_D\cdot\nabla\psi\left[\kappa_n+\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\kappa_T\right]F_{g0}-\frac{q}{T}(\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\langle\nabla\delta\Phi\rangle F_{g0}\\ &+ \frac{q}{T}\left[v_\parallel(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\langle\nabla\delta A_\parallel^{(h)}\rangle-\left\langle\delta A_\parallel^{(h)}\right\rangle(\mu\mathbf{e}_\parallel\cdot\nabla B_0)\right]F_{g0}, \end{aligned} \tag{178}$$

where use has been made of $\nabla\langle\delta\Phi\rangle=\langle\nabla\delta\Phi\rangle$ and $\nabla\langle\delta A_\parallel^{(h)}\rangle=\langle\nabla\delta A_\parallel^{(h)}\rangle$.

4.5.1 In field-aligned coordinates

In field-aligned coordinates (x,y,z) (where $x=\psi$, $y=\alpha$, $z=\theta$), $\partial\delta\Phi/\partial\mathbf{x}$ can be written as

$$\frac{\partial\delta\Phi}{\partial\mathbf{x}}=\frac{\partial\delta\Phi}{\partial x}\nabla x+\frac{\partial\delta\Phi}{\partial y}\nabla y+\frac{\partial\delta\Phi}{\partial z}\nabla z. \tag{179}$$

Note that the direction of $\partial\mathbf{r}/\partial x$ at $\theta=-\pi$ is different from that at $\theta=+\pi$. Therefore the value of $\partial\delta\Phi/\partial x$ at $\theta=-\pi$ is different from that at $\theta=+\pi$. I.e., $\partial\delta\Phi/\partial x$ is a non-periodic function of θ. Similarly, ∇y is also a non-perioidic functon of θ. In other words, both $\partial\delta\Phi/\partial x$ and ∇y are discontinuous across the θ cut ($\theta=\pm\pi$).

The gyro-average of expression (179) is written as

$$\langle\frac{\partial\delta\Phi}{\partial\mathbf{x}}\rangle\approx\langle\frac{\partial\delta\Phi}{\partial x}\rangle\nabla x+\langle\frac{\partial\delta\Phi}{\partial y}\rangle\nabla y+\langle\frac{\partial\delta\Phi}{\partial z}\rangle\nabla z, \tag{180}$$

where we approximate ∇x, ∇y, and ∇z as constants when performing the gyro-averge since they are determined by the equilibrim magnetic field, which is nearly constant on the Larmor radius scale. As is mentioned above, ∇y is not continuous at the θ cut. Do we need to worry about this? No. This is because we must stick to the same branch when we perform gryo-average on it, and hence ∇y is always continous. Then, do we need to worry about the disconunity of $\partial\delta\Phi/\partial x$ across the θ cut when performing the gyro-averge on it? We do not either. The reason is the same: we must stick to a single branch. The disconunity is just irrelevant here. The discontinuity only manifest itself when we need to infer value on $\theta=+\pi$ from that on $\theta=-\pi$ (vice versa), i.e., when across branch communication is explicitly needed. In `TEK`, the field equations are not solved at $\theta=+\pi$ and hence the field values are not directly obtained. Instead, the field

values at $\theta=+\pi$ are infered from the field values at $\theta=-\pi$. At the θ cut and for the same (x,ϕ), the continuity of $\nabla\delta\Phi$ requires

$$\langle\frac{\partial\delta\Phi}{\partial x}\rangle^{+}\nabla x+\langle\frac{\partial\delta\Phi}{\partial y}\rangle\nabla y^{+}=\langle\frac{\partial\delta\Phi}{\partial x}\rangle^{-}\nabla x+\langle\frac{\partial\delta\Phi}{\partial y}\rangle\nabla y^{-}, \tag{181}$$

where the superscript "+" and "−" refer to the location $\theta=+\pi$ and $\theta=-\pi$, respectively. Dotting the above by ∇x, we obtain

$$\langle\frac{\partial\delta\Phi}{\partial x}\rangle^{+}=\langle\frac{\partial\delta\Phi}{\partial x}\rangle^{-}+\langle\frac{\partial\delta\Phi}{\partial y}\rangle\frac{\nabla y^{-}\cdot\nabla x-\nabla y^{+}\cdot\nabla x}{|\nabla x|^{2}}, \tag{182}$$

which is used in TEK to infer the value of $\langle\frac{\partial\delta\Phi}{\partial x}\rangle^{+}$ from that of $\langle\frac{\partial\delta\Phi}{\partial x}\rangle^{-}$.

Similarly,

$$\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial\mathbf{x}}\rangle\approx\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial x}\rangle\nabla x+\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial y}\rangle\nabla y+\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial z}\rangle\nabla z. \tag{183}$$

Then Eq. (178) is written as

$$\begin{aligned}\frac{d\delta f^{(mv)}}{dt} &= \delta\mathbf{V}_D\cdot\nabla x\left[\kappa_n+\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\kappa_T\right]F_0\\ &- \frac{q}{T}\left[\langle\frac{\partial\delta\Phi}{\partial x}\rangle\frac{d\mathbf{X}'}{dt}\cdot\nabla x+\langle\frac{\partial\delta\Phi}{\partial y}\rangle\frac{d\mathbf{X}'}{dt}\cdot\nabla y+\langle\frac{\partial\delta\Phi}{\partial z}\rangle\frac{d\mathbf{X}'}{dt}\cdot\nabla z\right]F_0\\ &+ \frac{q}{T}v_{\parallel}\left(\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial x}\rangle\frac{d\mathbf{X}}{dt}\cdot\nabla x+\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial y}\rangle\frac{d\mathbf{X}}{dt}\cdot\nabla y+\langle\frac{\partial\delta A_{\parallel}^{(h)}}{\partial z}\rangle\frac{d\mathbf{X}}{dt}\cdot\nabla z\right)F_0\\ &- \frac{q}{T}\langle\delta A_{\parallel}^{(h)}\rangle(\mu\mathbf{e}_{\parallel}\cdot\nabla B_0)F_0,\end{aligned} \tag{184}$$

where $\frac{d\mathbf{X}}{dt}\equiv v_{\parallel}\mathbf{b}+\mathbf{V}_D+\delta\mathbf{V}_D$ and $\frac{d\mathbf{X}'}{dt}\equiv\mathbf{V}_D+\delta\mathbf{V}_D$. Define

$$\delta\overline{\Phi}=\frac{q_u\delta\Phi}{T_u},\delta\overline{A}_{\parallel}^{(h)}=\frac{q_uv_u\delta A_{\parallel}^{(h)}}{T_u}, \tag{185}$$

where q_u,T_u,v_u are units (independent of species), then Eq. (184) is written as

$$\begin{aligned}\frac{d\delta f^{(\mathrm{mv})}}{d\bar{t}} &= \delta\overline{\mathbf{V}}_D\cdot\nabla x\left[\kappa_n+\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\kappa_T\right]F_0\\ &- \frac{q/q_u}{T/T_u}\left[\langle\frac{\partial\delta\overline{\Phi}}{\partial x}\rangle\frac{d\mathbf{X}'}{d\bar{t}}\cdot\nabla x+\langle\frac{\partial\delta\overline{\Phi}}{\partial y}\rangle\frac{d\mathbf{X}'}{d\bar{t}}\cdot\nabla y+\langle\frac{\partial\delta\overline{\Phi}}{\partial z}\rangle\frac{d\mathbf{X}'}{d\bar{t}}\cdot\nabla z\right]F_0\\ &+ \frac{q/q_u}{T/T_u}\frac{v_n}{v_u}\bar{v}_{\parallel}\left[\langle\frac{\partial\delta\overline{A}_{\parallel}^{(h)}}{\partial x}\rangle\frac{d\mathbf{X}}{d\bar{t}}\cdot\nabla x+\langle\frac{\partial\delta\overline{A}_{\parallel}^{(h)}}{\partial y}\rangle\frac{d\mathbf{X}}{d\bar{t}}\cdot\nabla y+\langle\frac{\partial\delta\overline{A}_{\parallel}^{(h)}}{\partial z}\rangle\frac{d\mathbf{X}}{d\bar{t}}\cdot\nabla z\right]F_0\\ &- \frac{q/q_u}{T/T_u}\frac{v_n}{v_u}\langle\delta\overline{A}_{\parallel}^{(h)}\rangle(\bar{\mu}\mathbf{e}_{\parallel}\cdot\overline{\nabla}\overline{B}_0)F_0.\end{aligned} \tag{186}$$

$$\begin{aligned}\frac{d\delta f^{(\mathrm{mv})}}{d\bar{t}} &= \dot{x}^{(1)}\left[\kappa_n+\left(\frac{mv^2}{2T_0}-\frac{3}{2}\right)\kappa_T\right]F_0\\ &- \frac{q/q_u}{T/T_u}\left[\langle\frac{\partial\delta\overline{\Phi}}{\partial x}\rangle\dot{x}+\langle\frac{\partial\delta\overline{\Phi}}{\partial y}\rangle\dot{y}+\langle\frac{\partial\delta\overline{\Phi}}{\partial z}\rangle(\dot{z}-\dot{z}^{(0)})\right]F_0\\ &+ \frac{q/q_u}{T/T_u}\frac{v_n}{v_u}\bar{v}_\parallel\left[\langle\frac{\partial\delta\overline{A}_\parallel^{(h)}}{\partial x}\rangle\dot{x}+\langle\frac{\partial\delta\overline{A}_\parallel^{(h)}}{\partial y}\rangle\dot{y}+\langle\frac{\partial\delta\overline{A}_\parallel^{(h)}}{\partial z}\rangle\dot{z}\right]F_0\\ &- \frac{q/q_u}{T/T_u}\frac{v_n}{v_u}\langle\delta\overline{A}_\parallel^{(h)}\rangle(\bar{\mu}\mathbf{e}_\parallel\cdot\overline{\nabla}\overline{B}_0)F_0. \end{aligned}\tag{187}$$

---Normalized Ampere's equation---

Define

$$\delta\overline{A}_\parallel^{(s)}=\frac{q_u v_u\delta A_\parallel^{(s)}}{T_u},\delta\overline{J}_{\parallel j}^{(\mathrm{mv})}=\frac{\delta J_{\parallel j}^{(\mathrm{mv})}}{n_u q_u v_u},\tag{188}$$

then Ampere equation (170) is written as

$$\left(\sum_j\frac{\omega_{pj}^2}{c^2}-\nabla_\perp^2\right)\delta\overline{A}_\parallel^{(h)}=\nabla_\perp^2\delta\overline{A}_\parallel^{(s)}+\frac{n_u q_u^2}{\varepsilon_0 T_u}\frac{v_u^2}{c^2}\sum_j\delta\overline{J}_{\parallel j}^{(\mathrm{mv})},\tag{189}$$

where use has been made of $\mu_0=1/(c^2\varepsilon_0)$.

---$A_\parallel^{(s)}$ evolution---

In terms of units used in `TEK`, Eq. (166) is written as

$$\frac{\partial\delta\overline{A}_\parallel^{(s)}}{\partial\bar{t}}\frac{T_u}{q_u v_u}\frac{1}{t_u}=-\frac{T_u}{q_u L_n}\mathbf{b}\cdot\overline{\nabla}\delta\overline{\Phi},$$

where $\bar{t}=t/t_u$ with $t_u=L_n/v_u$. The equation can be simplified as

$$\frac{\partial\delta\overline{A}_\parallel^{(s)}}{\partial\bar{t}}=-\mathbf{b}\cdot\overline{\nabla}\delta\overline{\Phi},\tag{190}$$

i.e.,

$$\frac{\partial\delta\overline{A}_\parallel^{(s)}}{\partial\bar{t}}=-(\mathbf{b}\cdot\overline{\nabla}z)\frac{\partial\delta\overline{\Phi}}{\partial z}.\tag{191}$$

---pullback---

The pullback in Eq. (173) is now written as (for the case of F_0 being Maxwellian):

$$\begin{aligned}\delta f_{\mathrm{new}}^{(\mathrm{mv})} &= \delta f_{\mathrm{old}}^{(\mathrm{mv})}+\frac{q}{m}\langle v_\parallel\delta A_{\parallel\mathrm{old}}^{(h)}\rangle\left(-\frac{m}{T}\right)F_0\\ &= \delta f_{\mathrm{old}}^{(\mathrm{mv})}-\frac{q}{T}\langle v_\parallel\delta A_{\parallel\mathrm{old}}^{(h)}\rangle F_0\\ &= \delta f_{\mathrm{old}}^{(\mathrm{mv})}-\frac{q/q_u}{T/T_u}\frac{v_n}{v_u}\bar{v}_\parallel\langle\delta\overline{A}_{\parallel\mathrm{old}}^{(h)}\rangle F_0. \end{aligned}\tag{192}$$

Perturbed drfit

The $\delta\mathbf{E}\times\mathbf{B}_0$ drift term is written as

$$
\begin{aligned}
-\frac{\langle\frac{\partial\delta\Phi}{\partial\mathbf{x}}\rangle\times\mathbf{B}_0}{B_0^2} &= -\frac{1}{B_0^2}\left(\left\langle\frac{\partial\delta\Phi}{\partial x}\right\rangle\nabla x+\left\langle\frac{\partial\delta\Phi}{\partial y}\right\rangle\nabla y+\left\langle\frac{\partial\delta\Phi}{\partial z}\right\rangle\nabla z\right)\times\mathbf{B}_0\\
&= -\left\langle\frac{\partial\delta\Phi}{\partial x}\right\rangle\frac{1}{B_0^2}\nabla x\times\mathbf{B}_0-\left\langle\frac{\partial\delta\Phi}{\partial y}\right\rangle\frac{1}{B_0^2}\nabla y\times\mathbf{B}_0-\left\langle\frac{\partial\delta\Phi}{\partial z}\right\rangle\frac{1}{B_0^2}\nabla z\times\mathbf{B}_0.
\end{aligned}
$$

The ∇x, ∇y, and ∇z components of the above expression are written as

$$
\begin{aligned}
-\frac{\langle\frac{\partial\delta\Phi}{\partial\mathbf{x}}\rangle\times\mathbf{B}_0}{B_0^2}\cdot\nabla x &= -\langle\frac{\partial\delta\Phi}{\partial y}\rangle\frac{1}{B_0^2}\nabla y\times\mathbf{B}_0\cdot\nabla x-\langle\frac{\partial\delta\Phi}{\partial z}\rangle\frac{1}{B_0^2}\nabla z\times\mathbf{B}_0\cdot\nabla x\\
&= -\langle\frac{\partial\delta\Phi}{\partial y}\rangle\frac{1}{B_0^2}\nabla x\times\nabla y\cdot\mathbf{B}_0-\langle\frac{\partial\delta\Phi}{\partial z}\rangle\frac{1}{B_0^2}\nabla x\times\nabla z\cdot\mathbf{B}_0. \qquad (193)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\langle\frac{\partial\delta\Phi}{\partial\mathbf{x}}\rangle\times\mathbf{B}_0}{B_0^2}\cdot\nabla y &= -\langle\frac{\partial\delta\Phi}{\partial x}\rangle\frac{1}{B_0^2}\nabla x\times\mathbf{B}_0\cdot\nabla y-\langle\frac{\partial\delta\Phi}{\partial z}\rangle\frac{1}{B_0^2}\nabla z\times\mathbf{B}_0\cdot\nabla y\\
&= \langle\frac{\partial\delta\Phi}{\partial x}\rangle\frac{1}{B_0^2}\nabla x\times\nabla y\cdot\mathbf{B}_0-\langle\frac{\partial\delta\Phi}{\partial z}\rangle\frac{1}{B_0^2}\nabla y\times\nabla z\cdot\mathbf{B}_0. \qquad (194)
\end{aligned}
$$

$$
\begin{aligned}
-\frac{\langle\frac{\partial\delta\Phi}{\partial\mathbf{x}}\rangle\times\mathbf{B}_0}{B_0^2}\cdot\nabla z &= -\langle\frac{\partial\delta\Phi}{\partial x}\rangle\frac{1}{B_0^2}\nabla x\times\mathbf{B}_0\cdot\nabla z-\langle\frac{\partial\delta\Phi}{\partial y}\rangle\frac{1}{B_0^2}\nabla y\times\mathbf{B}_0\cdot\nabla z\\
&= \langle\frac{\partial\delta\Phi}{\partial x}\rangle\frac{1}{B_0^2}\nabla x\times\nabla z\cdot\mathbf{B}_0+\langle\frac{\partial\delta\Phi}{\partial y}\rangle\frac{1}{B_0^2}\nabla y\times\nabla z\cdot\mathbf{B}_0. \qquad (195)
\end{aligned}
$$

Note the general form:

$$
\frac{dx}{d\,\overline{t}}=t_n\mathbf{V}_G\cdot\nabla x. \qquad (196)
$$

Then, to get the normalizing factor, consider a typical term of the $\delta\mathbf{E}\times\mathbf{B}_0$ drift:

$$
\begin{aligned}
\frac{dx}{d\,\overline{t}} &= -t_n\left\langle\frac{\partial\delta\Phi}{\partial y}\right\rangle\frac{1}{B_0^2}\nabla x\times\nabla y\cdot\mathbf{B}_0+\cdots\\
&= -\frac{T_u}{q_u}\frac{t_n}{B_nL_n^2}\left\langle\frac{\partial\delta\overline{\Phi}}{\partial y}\right\rangle\frac{1}{\overline{B}_0^2}\overline{\nabla}x\times\overline{\nabla}y\cdot\overline{\mathbf{B}}_0+\cdots\\
&= -\frac{T_u}{q_u}\frac{1}{B_nL_nv_n}\left\langle\frac{\partial\delta\overline{\Phi}}{\partial y}\right\rangle\frac{1}{\overline{B}_0^2}\overline{\nabla}x\times\overline{\nabla}y\cdot\overline{\mathbf{B}}_0+\cdots, \qquad (197)
\end{aligned}
$$

where $T_u/(q_uv_nB_nL_n)$ is the normlizing factor we need when coding (in drift.f90).

Next, consider the magnetic fluttering term, which is written as

$$
-\frac{v_\parallel}{B_0^2}\mathbf{B}_0\times\langle\nabla\delta A_\parallel\rangle \approx -\frac{v_\parallel}{B_0^2}\mathbf{B}_0\times\left(\langle\frac{\partial\delta A_\parallel}{\partial x}\rangle\nabla x+\langle\frac{\partial\delta A_\parallel}{\partial y}\rangle\nabla y+\langle\frac{\partial\delta A_\parallel}{\partial z}\rangle\nabla z\right).
$$

The ∇x, ∇y, and ∇z components of the above expression are written as

$$
\begin{aligned}
-\frac{v_\parallel}{B_0^2}\langle \mathbf{B}_0 \times \nabla(\delta A_\parallel)\rangle \cdot \nabla x &= -\frac{v_\parallel}{B_0^2}\mathbf{B}_0 \times \left(\langle \frac{\partial \delta A_\parallel}{\partial y}\rangle \nabla y + \langle \frac{\partial \delta A_\parallel}{\partial z}\rangle \nabla z \right) \cdot \nabla x \\
&= \frac{v_\parallel}{B_0^2}\left(\langle \frac{\partial \delta A_\parallel}{\partial y}\rangle \nabla x \times \nabla y + \langle \frac{\partial \delta A_\parallel}{\partial z}\rangle \nabla x \times \nabla z \right) \cdot \mathbf{B}_0, \qquad (198) \\
-\frac{v_\parallel}{B_0^2}\langle \mathbf{B}_0 \times \nabla(\delta A_\parallel)\rangle \cdot \nabla y &= -\frac{v_\parallel}{B_0^2}\mathbf{B}_0 \times \left(\langle \frac{\partial \delta A_\parallel}{\partial x}\rangle \nabla x + \langle \frac{\partial \delta A_\parallel}{\partial z}\rangle \nabla z \right) \cdot \nabla y \\
&= \frac{v_\parallel}{B_0^2}\left(\langle \frac{\partial \delta A_\parallel}{\partial x}\rangle \nabla y \times \nabla x + \langle \frac{\partial \delta A_\parallel}{\partial z}\rangle \nabla y \times \nabla z \right) \cdot \mathbf{B}_0, \qquad (199) \\
-\frac{v_\parallel}{B_0^2}\langle \mathbf{B}_0 \times \nabla(\delta A_\parallel)\rangle \cdot \nabla z &= -\frac{v_\parallel}{B_0^2}\mathbf{B}_0 \times \left(\langle \frac{\partial \delta A_\parallel}{\partial x}\rangle \nabla x + \langle \frac{\partial \delta A_\parallel}{\partial y}\rangle \nabla y \right) \cdot \nabla z \\
&= \frac{v_\parallel}{B_0^2}\left(\langle \frac{\partial \delta A_\parallel}{\partial x}\rangle \nabla z \times \nabla x + \langle \frac{\partial \delta A_\parallel}{\partial y}\rangle \nabla z \times \nabla y \right) \cdot \mathbf{B}_0. \qquad (200)
\end{aligned}
$$

Also to get the normalizing factor, consider a typical term:

$$
\begin{aligned}
\frac{dx}{d\bar{t}} &= t_n \frac{v_\parallel}{B_0^2}\langle \frac{\partial \delta A_\parallel}{\partial y}\rangle \nabla x \times \nabla y \cdot \mathbf{B}_0 + \cdots \\
&= \frac{T_u}{q_u v_u} t_n \frac{v_n}{L_n^2 B_n} \frac{\overline{v}_\parallel}{\overline{B}_0^2} \langle \frac{\partial \delta \overline{A}_\parallel}{\partial y}\rangle \overline{\nabla} x \times \overline{\nabla} y \cdot \overline{\mathbf{B}}_0 + \cdots \\
&= \frac{T_u}{q_u v_u L_n B_n} \frac{\overline{v}_\parallel}{\overline{B}_0^2} \langle \frac{\partial \delta \overline{A}_\parallel}{\partial y}\rangle \overline{\nabla} x \times \overline{\nabla} y \cdot \overline{\mathbf{B}}_0 + \cdots, \qquad (201)
\end{aligned}
$$

where $T_u/(q_u v_u B_n L_n)$ is the normlizing factor I need when coding (in drift.f90).

The equilibrium terms in the above can be written as

$$
(\nabla x \times \nabla y) \cdot \mathbf{B}_0 = \frac{B_0^2}{\Psi'} \equiv \text{bdgxcgy}, \qquad (202)
$$

$$
(\nabla x \times \nabla z) \cdot \mathbf{B}_0 = \left(\frac{\partial x}{\partial Z}\frac{\partial z}{\partial R} - \frac{\partial x}{\partial R}\frac{\partial z}{\partial Z} \right) B_\phi \equiv \text{bdgxcgz}, \qquad (203)
$$

$$
(\nabla y \times \nabla z) \cdot \mathbf{B}_0 = \overline{B}_R \left(\frac{1}{R}\frac{\partial z}{\partial Z} \right) + \overline{B}_\phi \left(\frac{\partial y}{\partial Z}\frac{\partial z}{\partial R} - \frac{\partial y}{\partial R}\frac{\partial z}{\partial Z} \right) + \overline{B}_Z \left(-\frac{1}{R}\frac{\partial z}{\partial R} \right) \equiv \text{bdgycgz} \qquad (204)
$$

where the last entries in each lines are the variable names in `TEK` code.

4.6 Discretizing Laplacian operator

In `TEK`, the Laplacian operator is approximated as

$$
\nabla_\perp^2 \, \delta A_\parallel \approx \frac{\partial^2 \delta A_\parallel}{\partial x^2} |\nabla x|^2 + \frac{\partial^2 \delta A_\parallel}{\partial y^2} |\nabla y|^2 + 2 \frac{\partial^2 \delta A_\parallel}{\partial x \, \partial y} \nabla x \cdot \nabla y, \qquad (205)
$$

which is called "high-n approximation", in which all derivatives w.r.t z are dropped. This approximation reduces the Laplacian differential operator from 3D to 2D.

We assume that $\delta A_\parallel$ satisfies the zero boundary condition in the x direction: $A_\parallel(x = x_0) = 0, A_\parallel(x = x_0 + L_x) = 0$. Then the sine expansion can be used in this direction. In the y direction, full Fourier expansion is needed. I.e., at each value of z, $A_\parallel(x, y, z)$ is approximated by the following two-dimensional expansion:

$$\delta A_\parallel(x,y,z) \approx \sum_{n=-N_y/2}^{N_y/2} \exp\left(in\frac{2\pi}{L_y}y\right) \sum_{m=1}^{N_x-1} A_{mn}(z) \sin\left(\frac{m\pi}{L_x}(x-x_0)\right). \tag{206}$$

Use this expression, then $-\nabla_\perp^2 \delta A_\parallel$ of Eq. (205) is written as

$$\begin{aligned} -\nabla_\perp^2 \delta A_\parallel \;=\; & \sum_{n=-N_y/2}^{N_y/2} \exp\left(in\frac{2\pi}{L_y}y\right) \sum_{m=1}^{N_x-1} \delta A_{mn}(z) \\ & \times \left\{\left[\left(\frac{m\pi}{L_x}\right)^2 |\nabla x|^2 + \left(n\frac{2\pi}{L_y}\right)^2 |\nabla y|^2\right] \sin\left(\frac{m\pi}{L_x}(x-x_0)\right) - in\frac{2\pi}{L_y}\left(\frac{m\pi}{L_x}\right) 2\nabla x \cdot \nabla y \cos\left(\frac{m\pi}{L_x}(x-x_0)\right)\right\} \end{aligned}$$

At $x = x_j$, with $x_j = x_0 + j\Delta$ and $L_x/\Delta = N_x$, the above expression is written as

$$\begin{aligned} -\nabla_\perp^2 \delta A_\parallel \;=\; & \sum_{n=-N_y/2}^{N_y/2} \exp\left(in\frac{2\pi}{L_y}y\right) \sum_{m=1}^{N_x-1} \delta A_{m,n}(z) \\ & \times \left\{\left[\left(\frac{m\pi}{L_x}\right)^2 |\nabla x|^2 + \left(n\frac{2\pi}{L_y}\right)^2 |\nabla y|^2\right] \sin\left(\frac{jm\pi}{N_x}\right) - in\frac{2\pi}{L_y}\left(\frac{m\pi}{L_x}\right) 2\nabla x \cdot \nabla y \cos\left(\frac{jm\pi}{N_x}\right)\right\}, \end{aligned} \tag{207}$$

where ∇x and ∇y are evaluated at x_j. Plugging the DST coefficients

$$\delta A_{mn}(z) = \frac{2}{N_x} \sum_{j'=1}^{N_x-1} A_n(x_{j'}, z) \sin\left(\frac{j'm\pi}{N_x}\right), \tag{208}$$

into expression (207) gives

$$\begin{aligned} -\nabla_\perp^2 \delta A_\parallel \;=\; & \sum_{n=-N_y/2}^{N_y/2} \exp\left(in\frac{2\pi}{L_y}y\right) \sum_{m=1}^{N_x-1} \frac{2}{N_x} \sum_{j'=1}^{N_x-1} A_n(x_{j'}, z) \sin\left(\frac{j'm\pi}{N_x}\right) \\ & \times \left\{\left[\left(\frac{m\pi}{L_x}\right)^2 |\nabla x|^2 + \left(n\frac{2\pi}{L_y}\right)^2 |\nabla y|^2\right] \sin\left(\frac{jm\pi}{N_x}\right) - in\frac{2\pi}{L_y}\left(\frac{m\pi}{L_x}\right) 2\nabla x \cdot \nabla y \cos\left(\frac{jm\pi}{N_x}\right)\right\} \end{aligned}$$

For a single toroidal harmonic, the above expression reduces to

$$\begin{aligned} & \sum_{m=1}^{N_x-1} \frac{2}{N_x} \sum_{j'=1}^{N_x-1} A_n(x_{j'}, z) \sin\left(\frac{j'm\pi}{N_x}\right) \\ & \times \left\{\left[\left(\frac{m\pi}{L_x}\right)^2 |\nabla x|^2 + \left(n\frac{2\pi}{L_y}\right)^2 |\nabla y|^2\right] \sin\left(\frac{jm\pi}{N_x}\right) - in\frac{2\pi}{L_y}\left(\frac{m\pi}{L_x}\right) 2\nabla x \cdot \nabla y \cos\left(\frac{jm\pi}{N_x}\right)\right\} \end{aligned} \tag{209}$$

Define

$$\begin{aligned} M_{jj',n} \;\equiv\; & \sum_{m=1}^{N_x-1} \frac{2}{N_x} \sin\left(\frac{j'm\pi}{N_x}\right) \\ & \times \left\{\left[\left(\frac{m\pi}{L_x}\right)^2 |\nabla x|^2 + \left(n\frac{2\pi}{L_y}\right)^2 |\nabla y|^2\right] \sin\left(\frac{jm\pi}{N_x}\right) - in\frac{2\pi}{L_y}\left(\frac{m\pi}{L_x}\right) 2\nabla x \cdot \nabla y \cos\left(\frac{jm\pi}{N_x}\right)\right\} \end{aligned}$$

then expression (209) is written as

$$\sum_{j'=1}^{N_x-1} M_{jj',n} A_n(x_{j'}, z). \tag{210}$$

4.7 Parallel Ampere's Law in GEM code

$$-\nabla_\perp^2 \delta A_\parallel^{(n+1)} = \mu_0 \left(\delta J_{||i}^{(n+1)} + \delta J_{||e}^{(n+1)}\right), \tag{211}$$

where the parallel currents are given by

$$\delta J_{||i}^{(n+1)} = \delta J'_{||i}\left(\delta\phi^{(n)}, \delta A_\parallel^{(n)}\right) + \int \left(v_\parallel^{(n+1)}\right)^2 \frac{q_i^2}{m_i} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{i0}}{\partial \varepsilon} d\mathbf{v}, \tag{212}$$

$$\delta J_{||e} = \delta J'_{||e}\left(\delta\phi^{(n)}, \delta A_\parallel^{(n)}\right) + \int \left(v_\parallel^{(n+1)}\right)^2 \frac{q_e^2}{m_e} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{e0}}{\partial \varepsilon} d\mathbf{v}, \tag{213}$$

where $\delta J'_{\parallel i}$ and $\delta J'_{\parallel e}$ is the parallel current carried by the distribution function $\delta f^{(p_\parallel)}$, which are updated from the value at the nth time step using an explicit scheme and therefore does not depends on the field at the $(n+1)$th step. The blue terms in Eqs. (212) and (213) are the "skin current", which explicitly depend on the unknown field at the $(n+1)$th step. If we want to solve Ampere's law (211) by direct methods, then the blude terms need to be moved to the left-hand side. In this case, equation (211) is written as

$$\begin{aligned} &-\nabla_\perp^2 \delta A_\parallel^{(n+1)} - \mu_0 \int \left(v_\parallel^{(n+1)}\right)^2 \frac{q_i^2}{m_i} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{i0}}{\partial \varepsilon} d\mathbf{v} \\ &-\mu_0 \int \left(v_\parallel^{(n+1)}\right)^2 \frac{q_e^2}{m_e} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{e0}}{\partial \varepsilon} d\mathbf{v}. \\ &= \mu_0 \left(\delta J'_{\parallel i} + \delta J'_{||e}\right) \end{aligned} \tag{214}$$

Then we need to put the blue terms into matrix form. If we put the bule terms into martrix form by using numerical grid integration (as we do for the polarization density), then there arises the cancellation propblem (i.e., the two parts of the distribution are evaluated by different methods, one is grid-based and the other is MC marker based, there is a risk that the sum of the two terms will be inaccurate when the two terms are of opposite signs and large amplitudes, and the final result amplitude is expected to be much smaller than the amplituded of the two terms). If we get the matrix form by evaluating it numerically using MC markers (which can avoid the cancellation problem), the corresponding matrix will depends on markers and thus needs to be re-constructed each time-step, which is computationally expensive.

Therefore we go back to Eq. (211) and try to solve it using iterative methods. However, it is found numerically that directly using Eq. (211) as an iterative scheme is usually divergent. To obtain a convergent iterative scheme, we need to have an approximate form for the blue terms (bigger terms), which is independent of markers and so that it is easy to construct its matrix, and then subtract this approximate form from both sides. After doing this, the iterative scheme has better chance to be

convergent (partially due to that the right-hand side becomes smaller). An approximate form is that derived by neglecting the FLR effect given in Sec. 4.4. Using this, the iterative scheme for solving Eq. (211) is written as

$$\begin{aligned}
&-\nabla_\perp^2 \delta A_\parallel^{(n+1)} - \left(-\frac{\omega_{pi}^2}{c^2}\delta A_\parallel^{(n+1)} - \frac{\omega_{pe}^2}{c^2}\delta A_\parallel^{(n+1)} \right) \\
&= \mu_0 \left(\delta J'_{\parallel i} + \delta J'_{\parallel e} \right) \\
&+ \mu_0 \int \left(v_\parallel^{(n+1)} \right)^2 \frac{q_i^2}{m_i} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{i0}}{\partial \varepsilon} d\mathbf{v} \\
&+ \mu_0 \int \left(v_\parallel^{(n+1)} \right)^2 \frac{q_e^2}{m_e} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{e0}}{\partial \varepsilon} d\mathbf{v} \\
&- \left(-\frac{\omega_{pi}^2}{c^2}\delta A_\parallel^{(n+1)} - \frac{\omega_{pe}^2}{c^2}\delta A_\parallel^{(n+1)} \right).
\end{aligned} \tag{215}$$

In the drift-kinetic limit (i.e., neglecting the FLR effect), the blue and red terms on the right-hand side of the above equation cancel each other exactly. Even in this case, it is found numerically that these terms need to be retained and the blue terms are evaluated using markers. Otherwise, numerical inaccuracy can give numerical instabilities, which is the so-called cancellation problem. The explanation for this is as follows. The blue terms are part of the current. The remained part of the current carried by δh is computed by using Monte-Carlo integration over markers. If the blue terms are evaluated analytically, rather than using Monte-Carlo integration over markers, then the cancellation between this analytical part and Monte-Carlo part can have large error (assume that there are two large contribution that have opposite signs in the two parts) because the two parts are evaluated using different methods and thus have different accuracy, which makes the cancellation less accurate.

Because the ion skin current is smaller than its electron counterpart by a factor of m_e/m_i, its accuracy is not important. The cancellation error for ions is not a problem and hence can be neglected. In this case, equation (215) is simplified as

$$\begin{aligned}
&-\nabla_\perp^2 \delta A_\parallel^{(n+1)} - \left(-\frac{\omega_{pi}^2}{c^2}\delta A_\parallel^{(n+1)} - \frac{\omega_{pe}^2}{c^2}\delta A_\parallel^{(n+1)} \right) \\
&= \mu_0 (\delta J'_{\parallel i} + \delta J'_{\parallel e}) \\
&+ \mu_0 \int \left(v_\parallel^{(n+1)} \right)^2 \frac{q_e^2}{m_e} \left\langle \delta A_\parallel^{(n+1)} \right\rangle_\alpha \frac{\partial F_{e0}}{\partial \varepsilon} d\mathbf{v} \\
&- \left(-\frac{\omega_{pe}^2}{c^2}\delta A_\parallel^{(n+1)} \right).
\end{aligned} \tag{216}$$

Note that the blue term will be evaluated using Monte-Carlo markers.

4.8 Split-weight scheme for electrons in GEM code

The perturbed distribution function is decomposed as given by Eq. (157), i.e.,

$$\delta F = \delta f^{(p_\parallel)} + \frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial \varepsilon} + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \frac{\partial F_0}{\partial \varepsilon}, \tag{217}$$

where the term in blue is the so-called adiabatic response, which depends on the gyro-angle in guiding-center coordinates. Recall that the red term $\langle\delta\Phi\rangle_\alpha$, which is independent of the gyro-angle, is introduced in order to eliminate the time derivative $\partial\langle\delta\Phi\rangle_\alpha/\partial t$ term on the right-hand side of the original Frieman-Chen gyrokinetic equation.

The so-called generalized split-weight scheme corresponds to going back to the original Frieman-Chen gyrokinetic equation by introducing another $\langle\delta\Phi\rangle_\alpha$ term with a free small parameter ϵ_g. Specifically, δh in the above is split as

$$\delta h = \delta h_s + \epsilon_g \frac{q}{m}\langle\delta\Phi\rangle_\alpha \frac{\partial F_0}{\partial \varepsilon}. \tag{218}$$

(If $\varepsilon_g = 1$, then the two $\langle\delta\Phi\rangle_\alpha$ terms in Eq. (217) and (218) cancel each other.) Substituting this expression into Eq. (), we obtain the following equation for δh_s:

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\
&+\epsilon_g \frac{q}{m}\frac{\partial F_0}{\partial \varepsilon}\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\langle\delta\Phi\rangle_\alpha \\
&+\epsilon_g \frac{q}{m}\langle\delta\Phi\rangle_\alpha\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\frac{\partial F_0}{\partial \varepsilon} \\
&= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\
&-\frac{q}{m}[(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X(\langle\mathbf{v}\cdot\delta\mathbf{A} - \delta\Phi\rangle_\alpha)]\frac{\partial F_0}{\partial \varepsilon}.
\end{aligned} \tag{219}$$

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

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\
&= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\
&-\frac{q}{m}\left\{(v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X[\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha - \langle\delta\Phi\rangle_\alpha] + \epsilon_g\left[\frac{\partial\langle\delta\Phi\rangle_\alpha}{\partial t} + \mathbf{V}_G\cdot\nabla_X\langle\delta\Phi\rangle_\alpha\right]\right\}\frac{\partial F_0}{\partial \varepsilon}.
\end{aligned} \tag{220}$$

4.8.1 special case of $\epsilon_g = 1$

For the special case of $\epsilon_g = 1$ (the default and most used case in `GEM` code, Yang Chen said $\epsilon_g < 1$ cases are sometimes not accurate, so he gave up using it since 2009), equation (220) can be simplified as:

$$\begin{aligned}
&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\
&= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\
&-\frac{q}{m}\left[\mathbf{V}_G\cdot\nabla_X\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha + \frac{\partial\langle\delta\Phi\rangle_\alpha}{\partial t}\right]\frac{\partial F_0}{\partial \varepsilon},
\end{aligned} \tag{221}$$

where two $\mathbf{V}_G \cdot \langle \delta\Phi \rangle_\alpha$ terms cancel each other. Because the $v_\parallel E_\parallel$ term is one of the factors that make kinetic electron simulations difficult, eliminating $\mathbf{V}_G \cdot \langle \delta\Phi \rangle_\alpha$ term may be beneficial for obtaining stable algorithms.

For $\epsilon_g = 1$, δF is written as

$$\begin{aligned} \delta F &= \delta h_s + \frac{q}{m}\langle \delta\Phi \rangle_\alpha \frac{\partial F_{g0}}{\partial \varepsilon} + \frac{q}{m}(\delta\Phi - \langle \delta\Phi \rangle_\alpha)\frac{\partial F_0}{\partial \varepsilon} + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon} \\ &= \delta h_s + \frac{q}{m}\delta\Phi \frac{\partial F_{g0}}{\partial \varepsilon} + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon}, \end{aligned} \tag{222}$$

where the adiabatic term will be moved to the left-hand side of the Poisson's equation. The descretization of this term is much easier than the polarization density.

Equation (221) 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 split from the perturbed distribution function. Considering this, equation (221) can also be obtained from the original Frieman-Chen equation (135) by writing δG_0 as

$$\delta G_0 = \delta h_s + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon}, \tag{223}$$

In this case, δF is written as

$$\delta F = \delta h_s + \frac{q}{m}\delta\Phi \frac{\partial F_{g0}}{\partial \varepsilon} + \frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon}, \tag{224}$$

Substituting expression (223) into equation (135), we obtain the following equation for δh_s:

$$\begin{aligned} &\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\ &+\frac{q}{m}\frac{\partial F_0}{\partial \varepsilon}\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha \\ &+\frac{q}{m}\langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\frac{\partial F_0}{\partial \varepsilon} \\ &= -\delta\mathbf{V}_D\cdot\nabla_X F_0 - \frac{q}{m}\frac{\partial \langle \delta\Phi - \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha}{\partial t}\frac{\partial F_0}{\partial \varepsilon}, \end{aligned} \tag{225}$$

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

$$\begin{aligned} &\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X\right]\delta h_s \\ &= -\delta\mathbf{V}_D\cdot\nabla_X F_0 \\ &-\frac{q}{m}\left[\frac{\partial \langle \delta\Phi \rangle_\alpha}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D)\cdot\nabla_X \langle \mathbf{v}\cdot\delta\mathbf{A} \rangle_\alpha\right]\frac{\partial F_0}{\partial \varepsilon}, \end{aligned} \tag{226}$$

which agrees with Eq. (221).

In GEM, the split weight method is used only for electrons. When using this scheme, $\partial\delta\Phi/\partial t$ appears on the right-hand-side of the weight evolution equation. GEM makes use of the vorticity equation (time derivative of the Poissson equation) to evaluate $\partial\delta\Phi/\partial t$.

5 Poisson's equation and polarization density

Poisson's equation for the potential perturbation $\delta\Phi$ is written as

$$-\varepsilon_0\nabla^2\delta\Phi = q_i\delta n_i + q_e\delta n_e, \tag{227}$$

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. (227) is written as

$$-\varepsilon_0\nabla_\perp^2\delta\Phi = q_i\delta n_i + q_e\delta n_e. \tag{228}$$

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

Using expression (157), the density perturbation δn is written as

$$\begin{aligned}\delta n &= \int \delta F d\mathbf{v} \\ &= \int \delta h\, d\mathbf{v} + \int\left[\frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon}\right]d\mathbf{v} + \int\left[\frac{q}{m}\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}\right]d\mathbf{v},\end{aligned} \tag{229}$$

where the blue term 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 red term in expression (229) is the so-called the polarization density n_p, i.e.,

$$n_p(\mathbf{x}) = \int\frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}, \tag{230}$$

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 (228) is written as

$$-\varepsilon_0\nabla_\perp^2\delta\Phi - q_i\int\frac{q_i}{m_i}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_{i0}}{\partial\varepsilon}d\mathbf{v} = q_i\delta n_i' + q_e\delta n_e, \tag{231}$$

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.

(231) probably behaves better than the one that is based on the left-hand side of Eq. (228), i.e., $-\varepsilon_0 \nabla_\perp^2 \delta\Phi$.

5.1 Discussion on cancellation scheme

We note that the polarization density δn_p is part of the density perturbation. n_p is extracted from the source term and moved to the left-hand side of the Poisson equation. This term will be evaluated without using 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 two parts of the density perturbation can possibly introduce significant errors 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, rather than directly moving a part of the distribution function to the left-hand side, we subtract an (approximate) analytic expression from both sides of Eq. (228). The analytical expressions on both sides are evaluated based on grid values of perturbed electromagnetic fields and are independent of makers. All the original parts of the distribution functions are kept on the right-hand side and are still evaluated by using markers, which hopefully avoids the possible cancellation problem. This strategy is often called a cancellation scheme. Since unknown perturbed electromagnetic fields appear on the right-hand side, iteration is needed to solve the field equation.

Note that two things appear here: What motivates us to move parts of the distribution function to the left? It is the goal of hopefully making the left-hand side matrix more well-behaved (such as good condition number, etc.) Why do we need the cancellation scheme? Because we want to avoid the numerical inaccuracy that appears when large terms cancels each other. Note that iteration is needed when the cancellation scheme is used because the right-hand side explicitly contains unknown electromagnetic fields.

It turns out that the cancellation scheme is not necessary for Eq. (231), but for the field solver for Ampere's equation (discussed later), this cancellation scheme is necessary in order to obtain stable results.

6 Polarization density expressed as local Fourier expansion

Since $\delta\Phi$ is independent of the velocity in the particle coordinates, the dependence of the adiabatic term (the first term in expression (230)) on $\delta\Phi$ is trivial:

$$\begin{aligned} \delta n_{\mathrm{ad}} &= \int \frac{q}{m}(\delta\Phi)\frac{\partial F_0}{\partial \varepsilon} d\mathbf{v} \\ &= \frac{q}{m}\delta\Phi \int \frac{\partial F_0}{\partial \varepsilon} d\mathbf{v}. \end{aligned} \tag{232}$$

If we assume F_0 is Maxwellian, then the velocity integration can be analytically performed:

$$\begin{aligned}\delta n_{\rm ad} &= \frac{q}{m}\delta\Phi\int\Big(-\frac{m}{T}f_M\Big)d\mathbf{v}.\\ &= -\frac{q\delta\Phi}{T}n_0,\end{aligned} \tag{233}$$

Next, let us consider the second term in expression (230), i.e.,

$$-\int\frac{q}{m}\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}. \tag{234}$$

6.1 Gyro-averaging of $\delta\Phi$ in guiding-center coordinates

In order to perform the gyro-averaging of $\delta\Phi$, we Fourier expand $\delta\Phi$ in space as

$$\delta\Phi(\mathbf{x}) = \int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\frac{d\mathbf{k}}{(2\pi)^3}, \tag{235}$$

and then 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

$$\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}(\mathbf{x},\mathbf{v}) \approx \mathbf{X} - \mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}. \tag{236}$$

Using expressions (235) and (236), the gyro-average of $\delta\Phi$ is written as

$$\begin{aligned}\langle\delta\Phi\rangle_\alpha &= \left\langle\int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{x})\frac{d\mathbf{k}}{(2\pi)^3}\right\rangle_\alpha\\ &= \left\langle\int\delta\Phi_k\exp\left(i\mathbf{k}\cdot\left(\mathbf{X}-\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}\right)\right)\frac{d\mathbf{k}}{(2\pi)^3}\right\rangle_\alpha\\ &= \int\delta\Phi_k\exp(i\mathbf{k}\cdot\mathbf{X})\left\langle\exp\left(-i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}\right)\right\rangle_\alpha\frac{d\mathbf{k}}{(2\pi)^3}.\end{aligned} \tag{237}$$

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 blue expression in Eq. (237) is written as

$$\begin{aligned}-i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} &= -i\mathbf{k}\cdot v_\perp(\hat{\mathbf{e}}_1\cos\alpha+\hat{\mathbf{e}}_2\sin\alpha)\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}\\ &= -i\mathbf{k}\cdot\frac{v_\perp}{\Omega(\mathbf{X})}(-\hat{\mathbf{e}}_2\cos\alpha+\hat{\mathbf{e}}_1\sin\alpha)\\ &= -i\frac{k_\perp v_\perp}{\Omega}\sin\alpha.\end{aligned} \tag{238}$$

Then the gyro-averaging in expression (237) is written as

$$\begin{aligned}\left\langle \exp\left(-i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}\right)\right\rangle_\alpha &= \left\langle \exp\left(-i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)\right\rangle_\alpha \\ &= \frac{1}{2\pi}\int_0^{2\pi}\exp\left(-i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)d\alpha \\ &= J_0\left(\frac{k_\perp v_\perp}{\Omega}\right). \end{aligned} \tag{239}$$

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 (237) is written as

$$\langle\delta\Phi\rangle_\alpha = \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{X}) J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\frac{d\mathbf{k}}{(2\pi)^3}. \tag{240}$$

6.2 Gryo-angle integration in particle coordinates

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 (240) is written as

$$\langle\delta\Phi\rangle_\alpha = \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega}\right)\frac{d\mathbf{k}}{(2\pi)^3}. \tag{241}$$

Then the velocity integration is written as

$$\begin{aligned}&\int \langle\delta\Phi\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon} d\mathbf{v} \\ &= \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})\left[\int J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega}\right)\frac{\partial F_0}{\partial\varepsilon}d\mathbf{v}\right]\frac{d\mathbf{k}}{(2\pi)^3}. \end{aligned} \tag{242}$$

Similar to Eq. (238), 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

$$i\mathbf{k}\cdot\mathbf{v}\times\frac{\mathbf{e}_\parallel}{\Omega} = i\frac{k_\perp v_\perp}{\Omega}\sin\alpha. \tag{243}$$

Since this is at $\mathbf{x}$ rather than $\mathbf{X}$, $k_\perp$, $v_\perp$, and Ω are different from those appearing in expression (238). 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 (243) into expression (242) and using $d\mathbf{v} = v_\perp dv_\perp dv_\parallel d\alpha$, we get

$$\begin{aligned}&\int \langle\delta\Phi\rangle_\alpha \frac{\partial F_0}{\partial\varepsilon} d\mathbf{v} \\ &= \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})\left[\int J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\exp\left(i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right)\frac{\partial F_0}{\partial\varepsilon}v_\perp dv_\perp dv_\parallel d\alpha\right]\frac{d\mathbf{k}}{(2\pi)^3}. \end{aligned} \tag{244}$$

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. (244) can be performed, yielding

$$\begin{aligned}
&\int \langle \delta\Phi \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon} d\mathbf{v} \\
&= \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) \iint J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\left[\int_0^{2\pi} \exp\left(i\frac{k_\perp v_\perp}{\Omega}\sin\alpha\right) d\alpha\right] \frac{\partial F_0}{\partial \varepsilon} v_\perp dv_\perp dv_\parallel \frac{d\mathbf{k}}{(2\pi)^3} \\
&= \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) \left[\iint J_0\left(\frac{k_\perp v_\perp}{\Omega}\right) 2\pi J_0\left(\frac{k_\perp v_\perp}{\Omega}\right) \frac{\partial F_0}{\partial \varepsilon} v_\perp dv_\perp dv_\parallel\right] \frac{d\mathbf{k}}{(2\pi)^3}, \qquad (245)
\end{aligned}$$

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

6.2.1 The remaining velocity integration can be performed analytically if F_0 is Maxwellian

In order to perform the remaining velocity integration in expression (245), we assume that F_0 is a Maxwellian distribution given by

$$\begin{aligned}
F_0 &= f_M = \frac{n_0(\mathbf{X})}{(2\pi T(\mathbf{X})/m)^{3/2}} \exp\left(\frac{-mv^2}{2T(\mathbf{X})}\right) \qquad (246) \\
&= \frac{n_0}{(2\pi)^{3/2} v_t^3} \exp\left(\frac{-v^2}{2v_t^2}\right), \qquad (247)
\end{aligned}$$

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

$$\frac{\partial F_0}{\partial \varepsilon} = -\frac{m}{T} f_M. \qquad (248)$$

Again we will ignore the weak dependence of $n_0(\mathbf{X})$ and $T(\mathbf{X})$ on $\mathbf{v}$ 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.)

Parallel integration

Using Eq. (248), the expression in the square brackets of Eq. (245) is written as

$$\begin{aligned}
&2\pi \iint J_0^2\left(\frac{k_\perp v_\perp}{\Omega}\right) \frac{\partial F_0}{\partial \varepsilon} v_\perp dv_\perp dv_\parallel \\
&= -\frac{m}{T}\frac{n_0}{(2\pi)^{1/2}} \iint J_0^2\left(\frac{k_\perp v_\perp}{\Omega}\right) \frac{1}{v_t^3} \exp\left(-\frac{v_\parallel^2 + v_\perp^2}{2v_t^2}\right) v_\perp dv_\perp dv_\parallel \qquad (249) \\
&= -\frac{m}{T}\frac{n_0}{(2\pi)^{1/2}} \iint J_0^2\left(\frac{k_\perp v_\perp}{\Omega}\right) \exp\left(-\frac{\bar{v}_\parallel^2 + \bar{v}_\perp^2}{2}\right) \bar{v}_\perp d\bar{v}_\perp d\bar{v}_\parallel, \qquad (250)
\end{aligned}$$

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

$$\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{2\pi}, \tag{251}$$

the integration over $\bar{v}_\parallel$ in expression (250) can be performed, yielding

$$-\frac{m}{T} n_0 \int_0^{\infty} J_0^2\left(\frac{k_\perp v_\perp}{\Omega}\right) \exp\left(-\frac{\bar{v}_\perp^2}{2}\right) \bar{v}_\perp \, d\,\bar{v}_\perp \tag{252}$$

Perpendicular integration

Using (I verified this by using Sympy)

$$\int_0^{\infty} J_0^2(a\,x) \exp\left(-\frac{x^2}{2}\right) x\, dx = \exp(-a^2) I_0(a^2), \tag{253}$$

where $I_0(a)$ is the zeroth modified Bessel function of the first kind, expression (252) is written

$$-\frac{m}{T} n_0 \exp(-b) I_0(b) \tag{254}$$

where $b = k_\perp^2 v_t^2 / \Omega^2 = k_\perp^2 \rho_t^2$. Then the corresponding density (234) is written as

$$-\frac{q}{m} \int \langle \delta\Phi \rangle_\alpha \frac{\partial F_0}{\partial \varepsilon} d\mathbf{v} = \frac{q n_0}{T} \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) \exp(-b) I_0(b) \frac{d\mathbf{k}}{(2\pi)^3}. \tag{255}$$

6.2.2 Final form of polarization density

In Fourier space, the adiabatic term in expression (233) is written as

$$\int \frac{q}{m} (\delta\Phi) \frac{\partial F_0}{\partial \varepsilon} d\mathbf{v} = -\frac{q n_0}{T} \delta\Phi = -\frac{q n_0}{T} \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) \frac{d\mathbf{k}}{(2\pi)^3}. \tag{256}$$

Plugging expression (255) and (256) into expression (230), the polarization density n_p is written as

$$n_p = -\frac{q n_0}{T} \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) [1 - \exp(-b) I_0(b)] \frac{d\mathbf{k}}{(2\pi)^3}. \tag{257}$$

Define

$$\Gamma_0 = \exp(-b) I_0(b), \tag{258}$$

then Eq. (257) is written as

$$n_p = -\frac{q n_0}{T} \int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x}) [1 - \Gamma_0] \frac{d\mathbf{k}}{(2\pi)^3}, \tag{259}$$

Expression (259) agrees with the result given in Yang Chen's notes. Note that the dependence on species mass enters the formula through the Larmor radius ρ_t in Γ_0.

6.3 Pade approximation

Γ_0 defined in Eq. (258) can be approximated by the Pade approximation as

$$\Gamma_0 \approx \frac{1}{1+b}. \tag{260}$$

The comparison between the exact value of Γ_0 and the above Pade approximation is shown in Fig. 1.

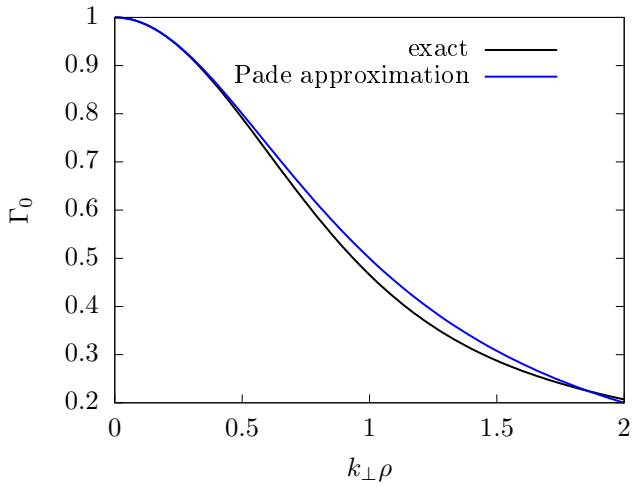

Figure 1. Comparison between the exact value of $\Gamma_0 = \exp(-(k_\perp\rho)^2)I_0((k_\perp\rho)^2)$ and the corresponding Pade approximation $1/(1+(k_\perp\rho)^2)$.

Using the Pade approximation (260), the polarization density n_p in expression (259) can be written as

$$n_p \approx -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})\frac{k_\perp^2\rho^2}{1+k_\perp^2\rho^2}\frac{d\mathbf{k}}{(2\pi)^3}. \tag{261}$$

(Pade approximate is the "best" approximation of a function by a rational function of given order – under this technique, the approximant's power series agrees with the power series of the function it is approximating.)

6.3.1 Long wavelength approximation of the polarization density

In the long wavelength limit, $k_\perp\rho \ll 1$, expression (261) can be further approximated as

$$\begin{aligned} n_p &\approx -\frac{qn_0}{T}\int \delta\Phi_k \exp(i\mathbf{k}\cdot\mathbf{x})k_\perp^2\rho^2\frac{d\mathbf{k}}{(2\pi)^3}, \\ &= \frac{qn_0}{T}\rho^2\nabla_\perp^2\delta\Phi. \end{aligned} \tag{262}$$

Then the corresponding term in the Poisson equation is written as

$$\begin{aligned} \frac{q}{\varepsilon_0}n_p &= \frac{q^2 n_0}{\varepsilon_0 T}\rho^2\nabla_\perp^2\delta\Phi \\ &= \frac{\rho^2}{\lambda_D^2}\nabla_\perp^2\delta\Phi, \end{aligned} \tag{263}$$

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. Therefore the term in expression (263) 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 in the long wavelength limit.

Equation (263) also shows that electron polarization density is smaller than the ion polarization density by a factor of $\rho_e/\rho_i \approx 1/60$. Note that this conclusion is drawn in the long wavelength limit. For short wavelength, the electron polarization and ion polarization density can be of similar magnitude (to be discussed later).

6.3.2 Polarization density expressed in terms of Laplacian operator

The polarization density expression (262) is for the long wavelength limit, which partially neglects FLR effect. Let us go back to the more general expression (261). The Poisson equation is written

$$-\varepsilon_0\nabla_\perp^2\delta\Phi = q_i\delta n_i + q_e\delta n_e. \tag{264}$$

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

$$-\varepsilon_0\nabla_\perp^2\delta\Phi - q_i n_{pi} = q_i\delta n_i' + q_e\delta n_e. \tag{265}$$

Fourier transforming in space, the above equation is written

$$-\varepsilon_0 k_\perp^2\delta\hat{\Phi} - q_i\hat{n}_{pi} = q_i\delta\hat{n}_i' + q_e\delta\hat{n}_e, \tag{266}$$

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 (261) implies that $\hat{n}_{pi}$ is given by

$$\hat{n}_{pi} = -\frac{q_i n_{i0}}{T_i}\delta\hat{\Phi}\frac{k_\perp^2\rho_i^2}{1+k_\perp^2\rho_i^2}. \tag{267}$$

Using this, equation (266) is written

$$-\varepsilon_0 k_\perp^2\delta\hat{\Phi} - q_i\left(-\frac{q_i n_{i0}}{T_i}\frac{k_\perp^2\rho_i^2}{1+k_\perp^2\rho_i^2}\delta\hat{\Phi}\right) = q_i\delta\hat{n}_i' + q_e\delta\hat{n}_e, \tag{268}$$

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

$$-(1+k_\perp^2\rho_i^2)k_\perp^2\delta\hat{\Phi} - \frac{q_i}{\varepsilon_0}\left(-\frac{q_i n_{i0}}{T_i}(k_\perp^2\rho_i^2)\delta\hat{\Phi}\right) = \frac{1}{\varepsilon_0}(1+k_\perp^2\rho_i^2)(q_i\delta\hat{n}_i' + q_e\delta\hat{n}_e). \tag{269}$$

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

$$-(1-\rho_i^2\nabla_\perp^2)\nabla_\perp^2\delta\Phi - \frac{q_i}{\varepsilon_0}\left(\frac{q_i n_{i0}}{T_i}\rho_i^2\nabla_\perp^2\delta\Phi\right) = \frac{1}{\varepsilon_0}(1-\rho_i^2\nabla_\perp^2)(q_i\delta n_i' + q_e\delta n_e). \tag{270}$$

Neglecting the Debye shielding term, the above equation is written

$$-\left(\frac{\rho_i^2}{\lambda_{Di}^2}\nabla_\perp^2\delta\Phi\right) = \frac{1}{\varepsilon_0}(1-\rho_i^2\nabla_\perp^2)(q_i\delta n_i' + q_e\delta n_e), \tag{271}$$

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

7 Polarization density obtained by numerical integration in phase space using grid

In Sec. 6, to evaluate the polarization density, the potential $\delta\Phi$ is Fourier expanded in space using local Cartesian coordinates, and then the double gyro-angle integration of each harmonic is expressed as the Bessel function. (It seems that the original motivation of using the Fourier expansion is to facilitate analytical treatment and is not designed for numerical use. GEM code does make use of the local Fourier expansion in its numerical implementation, where the local perpendicular wave number needs to be estimated numerically, which seems awkward.)

In this section, we avoid using the local Fourier expansion, and directly express the double gyro-angle integral as linear combination of values of $\delta\Phi$ at spatial grid-points. The polarization density is given by Eq. (230), i.e.,

$$n_p(\mathbf{x}) = \frac{q}{m}\int d\mathbf{v}\left((\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon}\right). \tag{272}$$

Our goal is to write n_p in the following real-space discrete from:

$$n_p(\mathbf{x}_k) = \sum_{i,j} c_{ij}\delta\Phi(\mathbf{x}_{ij}), \tag{273}$$

where c_{ij} are known coefficients, $\mathbf{x}_{ij}$ are spatial gridpoints.

7.1 Direct evaluation of the double gyrophase integration

Consider the gyro-average part of n_p:

$$A(\mathbf{x}) \equiv -\frac{q}{m}\int d\mathbf{v}\left(\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}\right). \tag{274}$$

At a space point $\mathbf{x}$, we constructe local (velocity) cylindrical coordinates $(v_\perp, v_\parallel, \alpha)$. Then the velocity element is given by $d\mathbf{v} = v_\perp dv_\perp dv_\parallel d\alpha$. Then the above integration is written as

$$A(\mathbf{x}) = -\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp \frac{\partial F_0}{\partial\varepsilon}\int_0^{2\pi} d\alpha \langle\delta\Phi\rangle_\alpha. \tag{275}$$

Note that $\partial F_0 / \partial\varepsilon$ usually dpends on α. But this dependence is usually weak and is neglected (i.e., variation of the equilibrium distribution on Larmor radius scale is assumed to be small). So it is moved outside of the α integration. Using the definition of the gyro-averaging $\langle\ldots\rangle = (2\pi)^{-1}\int_0^{2\pi}(\ldots)d\alpha'$, expression (275) is written as

$$A(\mathbf{x}) = -\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial\varepsilon} v_\perp \int_0^{2\pi} d\alpha\left(\frac{1}{2\pi}\int_0^{2\pi}\delta\Phi(\mathbf{x})d\alpha'\right). \tag{276}$$

Note that the gyro-averaging is performed in the guiding-center space, i.e., performed by varying the gyroangle α' while keeping guiding-center position $\mathbf{X}$, $v_\perp$, and $v_\parallel$ constant. So we need to transform to guiding-center varaibles. Particle location $\mathbf{x}$ is computed from the guiding-center location $\mathbf{X}$ by

$$\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}(\mathbf{x}, \mathbf{v}) \approx \mathbf{X} - \mathbf{v} \times \frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})}, \tag{277}$$

then $A(\mathbf{x})$ is written as

$$\begin{aligned}
&A(\mathbf{x}) \\
=&-\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial \varepsilon} v_\perp \int_0^{2\pi} d\alpha \left[\frac{1}{2\pi} \int_0^{2\pi} \delta\Phi\left(\mathbf{X} - \mathbf{v}_\perp(v_\perp, \alpha') \times \frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} \right) d\alpha' \right] \\
\approx&-\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial \varepsilon} v_\perp \int_0^{2\pi} d\alpha \left[\frac{1}{N_2} \sum_{j=1}^{N_2} \delta\Phi\left(\mathbf{X} - \mathbf{v}_\perp(v_\perp, \alpha_j') \times \frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} \right) \right],
\end{aligned}$$

where $\mathbf{v}_\perp = \mathbf{v}(v_\perp, \alpha_j')$ denotes perpendicular velocity corresponding to a discrete gyro-angle α_j', N_2 is the number of points chosen for α' discretization.

Next, in order to perform the remaining velocity space integration, we transform back to the particle coordinates (because the velocity integration is performed in the particle coordinates, i.e., it is performed by keeping the particle coordinate $\mathbf{x}$ constant):

$$\begin{aligned}
A(\mathbf{x}) =& -\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial \varepsilon} v_\perp \\
&\times \int_0^{2\pi} d\alpha \left[\frac{1}{N_2} \sum_{j=1}^{N_2} \delta\Phi\left(\mathbf{x} + \mathbf{v}_\perp(v_\perp, \alpha) \times \frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} - \mathbf{v}_\perp(v_\perp, \alpha_j') \times \frac{\mathbf{e}_\parallel(\mathbf{X})}{\Omega(\mathbf{X})} \right) \right] \\
\approx& -\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial \varepsilon} v_\perp \\
&\times \int_0^{2\pi} d\alpha \left[\frac{1}{N_2} \sum_{j=1}^{N_2} \delta\Phi\left(\mathbf{x} + \mathbf{v}_\perp(v_\perp, \alpha) \times \frac{\mathbf{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})} - \mathbf{v}_\perp(v_\perp, \alpha_j') \times \frac{\mathbf{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})} \right) \right] \\
\approx& -\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial \varepsilon} v_\perp \\
&\times \frac{2\pi}{N_1} \sum_{i=1}^{N_1} \frac{1}{N_2} \sum_{j=1}^{N_2} \delta\Phi\left(\mathbf{x} + \mathbf{v}_\perp(v_\perp, \alpha_i) \times \frac{\mathbf{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})} - \mathbf{v}_\perp(v_\perp, \alpha_j') \times \frac{\mathbf{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})} \right),
\end{aligned} \tag{278}$$

where N_1 is the number of points chosen for α discretization. For notation ease, define

$$\Delta\boldsymbol{\rho}_{ij} = [\mathbf{v}_\perp(v_\perp, \alpha_i) - \mathbf{v}_\perp(v_\perp, \alpha_j')] \times \frac{\mathbf{e}_\parallel(\mathbf{x})}{\Omega(\mathbf{x})}, \tag{279}$$

which is a function of $(\mathbf{x}, v_\perp, \alpha_i, \alpha_j')$. Then Eq. (278) is written as

$$A(\mathbf{x}) \;=\; -\frac{q}{m}\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty dv_\perp \frac{\partial F_0}{\partial \varepsilon} v_\perp \frac{2\pi}{N_1} \sum_{i=1}^{N_1} \frac{1}{N_2} \sum_{j=1}^{N_2} \delta\Phi(\mathbf{x} + \Delta\boldsymbol{\rho}_{ij}). \tag{280}$$

The guiding-center transform and its inverse used in the above are illustrated in Fig. 2.

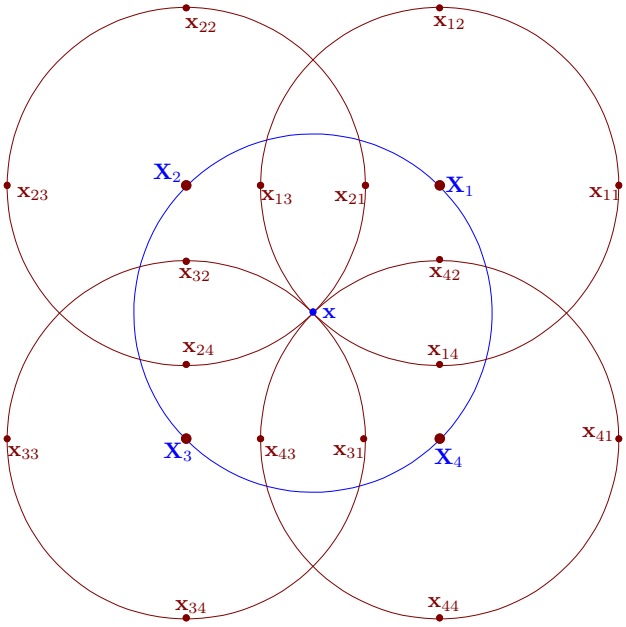

Figure 2. The guiding-center transform and its inverse used in Eq. (278). Here $\mathbf{x}$ and $v_\perp$ are given.

Let us summarize the above algorithm. Given $v_\perp$, the double gyro-angle integration at a particle location $\mathbf{x}$ is evaluated by the following steps:

- Use the guiding center transform to get guiding-center locations (four locations are shown in Fig. 2, namely $\mathbf{X}_i$ with $i=1,2,3,4$, corresponding gyro-angle α_i with $i=1,2,3,4$;
- For each guiding-center location, use the inverse guiding-center transform to calculate points on the corresponding gyro-ring (four points are shown for each guiding-center $\mathbf{X}_i$ in this case, namely $\mathbf{x}_{ij}$ with $j=1,2,3,4$, corresponding gyro angle α'_j with $j=1,2,3,4$.)
- Then the double gyro-angle integral appearing in the polarization density is approximated as

$$\int_0^{2\pi} d\alpha \left(\frac{1}{2\pi} \int_0^{2\pi} \delta\Phi(\mathbf{x}) d\alpha' \right) \approx \frac{2\pi}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \delta\Phi(\mathbf{x}_{ij}), \tag{281}$$

where $N_1=4, N_2=4$ for the case shown in Fig. 2.

The above 3 steps specify how to approximate the double gyro-angle integration using the discrete values of $\delta\Phi(\mathbf{x}_{ij})$. The spatial points $\mathbf{x}_{ij}$ appearing in Eq. (281) are not necessarily grid points. Linear interpolations are used to express $\delta\Phi(\mathbf{x}_{ij})$ as linear combination of values of $\delta\Phi$ at grid-points.

7.2 Toroidal DFT of polarization density in field-aligned coordinates

In the field-aligned coordinates (x, y, z), Fourier expand $\delta\Phi$ along y:

$$\delta\Phi(\mathbf{x}) \approx \sum_{n=-N_y/2}^{N_y/2} \exp\left(\iota n \frac{2\pi}{L_y} y\right) \delta\Phi_n(x, z), \tag{282}$$

where $\iota = \sqrt{-1}$, N_y is an even integer specifying how many harmonics are included, $\delta\Phi_n$ is the expansion coefficient. Using expression (282) in Eq. (280), we obtain

$$\begin{aligned} A(\mathbf{x}) &= -\frac{q}{m}\int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} dv_{\perp} \frac{\partial F_0}{\partial \varepsilon} v_{\perp} \frac{2\pi}{N_1} \sum_{i=1}^{N_1} \frac{1}{N_2} \sum_{j=1}^{N_2} \\ &\quad \sum_{n=-N_y/2}^{N_y/2} \exp\left(\iota n \frac{2\pi}{L_y}(y + \Delta\rho_{ijy})\right) \delta\Phi_n(x + \Delta\rho_{ijx}, z + \Delta\rho_{ijz}) \\ &= \sum_{n=-N_y/2}^{N_y/2} \exp\left(\iota n \frac{2\pi}{L_y} y\right) \left(-\frac{q}{m}\right) \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} dv_{\perp} \frac{\partial F_0}{\partial \varepsilon} v_{\perp} \frac{2\pi}{N_1} \sum_{i=1}^{N_1} \frac{1}{N_2} \sum_{j=1}^{N_2} \\ &\quad \exp\left(\iota n \frac{2\pi}{L_y} \Delta\rho_{ijy}\right) \delta\Phi_n(x + \Delta\rho_{ijx}, z + \Delta\rho_{ijz}). \end{aligned} \tag{283}$$

From Eq. (283), the Fourier expansion coefficient of $A(\mathbf{x})$ can be identified:

$$\begin{aligned} A_n(x, z) &= \left(-\frac{q}{m}\right) \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} dv_{\perp} \frac{\partial F_0}{\partial \varepsilon} v_{\perp} \frac{2\pi}{N_1} \sum_{i=1}^{N_1} \frac{1}{N_2} \sum_{j=1}^{N_2} \\ &\quad \exp\left(\iota n \frac{2\pi}{L_y} \Delta\rho_{ijy}\right) \delta\Phi_n(x + \Delta\rho_{ijx}, z + \Delta\rho_{ijz}). \end{aligned} \tag{284}$$

Assume each toroidal harmonic $\exp\left(\iota n \frac{2\pi}{L_y} y\right)\delta\Phi_n(x, z)$ is weakly varying along field linez, then we can make the following approximation:

$$\delta\Phi_n(x + \Delta\rho_{ijx}, z + \Delta\rho_{ijz}) \approx \delta\Phi_n(x + \Delta\rho_{ijx}, z), \tag{285}$$

because the variation along a field line over a distance of a Larmor radius is small. Then expression (284) is written as

$$\begin{aligned} A_n(x, z) &\approx \left(-\frac{q}{m}\right) \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} dv_{\perp} \frac{\partial F_0}{\partial \varepsilon} v_{\perp} \frac{2\pi}{N_1} \sum_{i=1}^{N_1} \frac{1}{N_2} \sum_{j=1}^{N_2} \\ &\quad \exp\left(\iota n \frac{2\pi}{L_y} \Delta\rho_{ijy}\right) \delta\Phi_n(x + \Delta\rho_{ijx}, z) \end{aligned} \tag{286}$$

The approximation (285) makes the dependence of $A_n(x,z)$ on $\delta\Phi_n(x,z)$ become local in the z direction, i.e., $A_n(x,z)$ at $z=z_0$ only involves values of $\delta\Phi_n(x,z)$ at $z=z_0$. Meanwhile $A_n(x,z)$ at $x=x_0$ involves values of $\delta\Phi_n(x,z)$ at multiple values of x.

7.3 For Maxwelliand distribution

Assume that F_0 is Maxwellian, then

$$\frac{\partial F_0}{\partial\varepsilon}=-\frac{m}{T}f_M=-\frac{m}{T}\frac{n_0}{(2\pi T/m)^{3/2}}\exp\left(\frac{-mv_\parallel^2}{2T}\right)\exp\left(\frac{-mv_\perp^2}{2T}\right). \tag{287}$$

Note that $\Delta\boldsymbol{\rho}_{ij}$ is independent of $v_\parallel$. Then the integration over $v_\parallel$ in Eq. (286) can be analytically performed:

$$\begin{aligned}
A_n(x,z) &= \left(\frac{q}{T}\right)\int_{-\infty}^{\infty}\exp\left(\frac{-mv_\parallel^2}{2T}\right)dv_\parallel\int_0^\infty dv_\perp\left(\frac{n_0}{(2\pi T/m)^{3/2}}\exp\left(\frac{-mv_\perp^2}{2T}\right)\right)v_\perp\\
&\times\ \frac{2\pi}{N_1}\sum_{i=1}^{N_1}\frac{1}{N_2}\sum_{j=1}^{N_2}\exp\left(\iota n\frac{2\pi}{L_y}\Delta\rho_{ijy}\right)\delta\Phi_n(x+\Delta\rho_{ijx},z)\\
&= \frac{q}{T}n_0\int_0^\infty d\,\bar v_\perp\bar v_\perp\exp\left(\frac{-\bar v_\perp^2}{2}\right)\\
&\times\ \frac{1}{N_1}\sum_{i=1}^{N_1}\frac{1}{N_2}\sum_{j=1}^{N_2}\left[\exp\left(\iota n\frac{2\pi}{L_y}\Delta\rho_{ijy}\right)\delta\Phi_n(x+\Delta\rho_{ijx},z)\right].
\end{aligned}$$

where $\bar v_\perp=v_\perp/\sqrt{T/m}$, and use has been made of

$$\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2}\right)dx=\sqrt{2\pi}. \tag{288}$$

7.4 Using MC integration, to be continued, not necessary

The integration we try to express is given by

$$\begin{aligned}
A_{ijk} &\approx \frac{1}{V_{ijk}}\int_{V_{ijk}}d\mathbf{x}n_p(\mathbf{x})\\
&= \frac{1}{V_{ijk}}\int_{V_{ijk}}d\mathbf{x}\int d\mathbf{v}\left(\frac{q}{m}(\delta\Phi-\langle\delta\Phi\rangle_\alpha)\frac{\partial F_0}{\partial\varepsilon}\right),
\end{aligned} \tag{289}$$

where A_{ijk} is the value of $n_p(\mathbf{x})$ at a grid-point, which is approximated by the spatial averaging of $n_p(\mathbf{x})$ over the cell whose center is the grid-point, V_{ijk} is the volume of the cell. We will use Monte-Carlo guiding-center markers to compute the above cell average.

$$\begin{aligned}
A_{ijk}^{(1)} &= -\frac{q}{m}\frac{1}{V_{ijk}}\sum_p\sum_\alpha\left(\langle\delta\Phi\rangle_\alpha\frac{\partial F_0}{\partial\varepsilon}\frac{1}{g}\right)\\
&= -\frac{q}{m}\frac{1}{V_{ijk}}\sum_p\langle\delta\Phi\rangle_\alpha\sum_\alpha\left(\frac{\partial F_0}{\partial\varepsilon}\frac{1}{g}\right)
\end{aligned}$$

Nearest neighbour interpolation

$$\begin{aligned} A^{(2)}_{ijk} &= -\frac{q}{m}\frac{1}{V_{ijk}}\sum_p \sum_\alpha \left(\delta\Phi \frac{\partial F_0}{\partial \varepsilon}\frac{1}{g}\right) \\ &= -\frac{q}{m}\frac{\delta\Phi_{ijk}}{V_{ijk}}\sum_p \sum_\alpha \left(\frac{\partial F_0}{\partial \varepsilon}\frac{1}{g}\right) \end{aligned}$$

Later, I found that evaluating the polarization density using grid works well. Therefore there is no need to evaluate it using MC markers, which is more complicated.

7.5 TEK benchmarking with GENE in DIII-D cyclone base case

The DIII-D cyclone base case is a circular and concentric magnetic configuration. The main parameters are summarized in Table 1. All parameters are the same as those in T. Gorler's paper[?]. (This configuratgion was inspired by DIII-D H-mode discharge #81499 and used in the cyclone project for benchmarking various gyrokinetic ITG turbulence simulations.)

R_0	a	r_0	q_0	$\hat{s}_0$	$B_{\rm axis}$	$\hat{\kappa}_{T_i}R_0$	$\hat{\kappa}_{n_i}R_0$	T_{i0}	$q_iT_{i0}/(eT_{e0})$
$1.67m$	$0.60m$	$0.30m$	1.41	0.84	$2.0T$	6.96	2.23	2.14keV	1

Table 1. DIII-D cyclone base case parameters[?][?] where R_0 is the major radius of the magnetic axis, a is the minor radius of the last-closed-flux-surface, q_0 is the value of safety factor at $r=r_0=0.5a$, $\hat{s}_0$ is the value of the magnetic shear $\hat{s}=rq^{-1}dq/dr$ at $r=r_0$. $B_{\rm axis}$ is the magnetic field strength at the magnetic axis. The toroidal field function $g(r)=B_\phi R$ is assumed to be a constant independent of r, i.e., $g=B_{\rm axis}R_0$. T_{i0} is the ion temperature at $r=r_0$. Assume Deuterium plasma, then $R_0/\rho_i=500$ and $a/\rho_i=180$. where $\rho_i=\sqrt{T_{i0}/m_i}/\Omega_i$ is the thermal ion gyro-radius, $\Omega_i=B_{\rm axis}q_i/m_i$ is the ion cyclotron angular frequency at the magnetic axis.

The safety factor profile is chosen as [?]

$$q(r)=0.86-0.16\,r/a+2.52\,(r/a)^2$$

This profile gives $q_0 \equiv q(0) = 1.41$ and $\hat{s} = 0.84$ at $r = r_0$. The equilibrium ion temperature and number density profile are given by

$$T_i = T_{i0}\exp\left(-\hat{\kappa}_{T_i}a\Delta_{T_i}\tanh\left(\frac{r-r_0}{\Delta_{T_i}a}\right)\right). \tag{290}$$

$$n_i = n_{i0}\exp\left(-\hat{\kappa}_{n_i}a\Delta_{n_i}\tanh\left(\frac{r-r_0}{\Delta_{n_i}a}\right)\right). \tag{291}$$

where $\Delta_{T_i} = \Delta_{n_i} = 0.3$, $n_{i0} = 4.66 \times 10^{19} m^{-3}$, $T_{i0} = 2.14$keV. Then the radial derivatives are written as

$$\begin{aligned} \frac{dT_i}{dr} &= T_i(-\hat{\kappa}_{T_i} a \Delta_{T_i}) \frac{d}{dr} \tanh\left(\frac{r - r_0}{\Delta_{T_i} a} \right) \\ &= T_i(-\hat{\kappa}_{T_i} a \Delta_{T_i}) \mathrm{sech}^2\left(\frac{r - r_0}{\Delta_{T_i} a} \right) \frac{1}{\Delta_{T_i} a} \\ &= -T_i \hat{\kappa}_{T_i} \mathrm{sech}^2\left(\frac{r - r_0}{\Delta_{T_i} a} \right) \end{aligned}$$

The corresponding gradient scale lengths of T_i and n_i are then given by

$$\kappa_{T_i} = -\frac{1}{T_i} \frac{dT_i}{dr} = \hat{\kappa}_{T_i} \mathrm{sech}^2\left(\frac{r - r_0}{\Delta_{T_i} a} \right), \tag{292}$$

and

$$\kappa_{n_i} = -\frac{1}{n_i} \frac{dn_i}{dr} = \hat{\kappa}_{n_i} \mathrm{sech}^2\left(\frac{r - r_0}{\Delta_{n_i} a} \right), \tag{293}$$

where $\mathrm{sech}^2(\cdot) \equiv 1 - \tanh^2(\cdot)$. The electron density and temperature profiles are assumed to be identical to those of ions. The electron mass is set as $m_e = 1.82 \times 10^{-30} kg$, i.e., two times the realistic value of electron mass.

The profiles of safety factor, temperature, and density are plotted in Fig. 3.

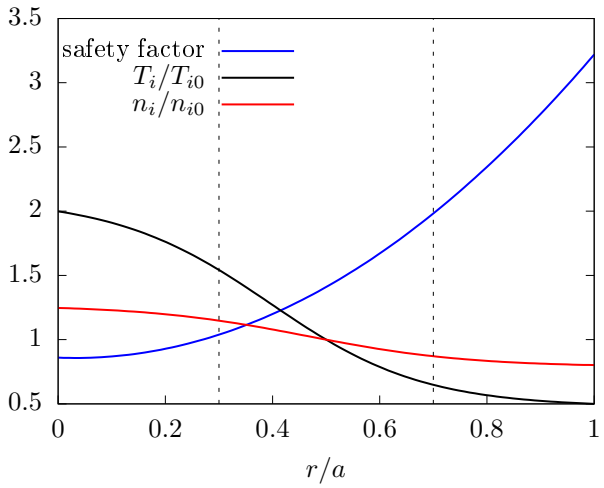

Figure 3. Radial profiles of safety factor, temperature, and density of the above DIII-D cyclone test case. The dashed lines indicate the radial simulation box $0.3 \leqslant r/a \leqslant 0.7$.

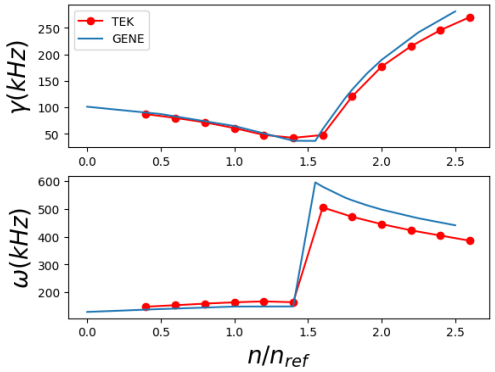

Figure 4. using gfile resolution of 257x257. left: uniform in real-space loading. right: uniform in field-algined corrdinates loading.

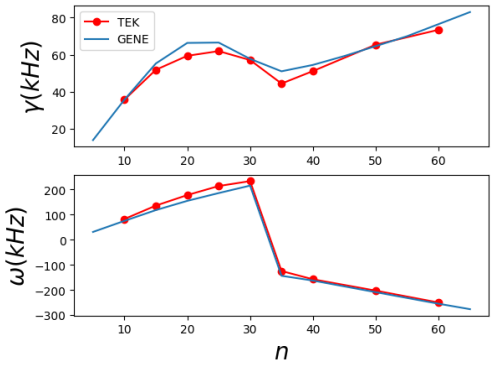

Figure 5. ITG-TEM transition benchmarking

Mode structures:

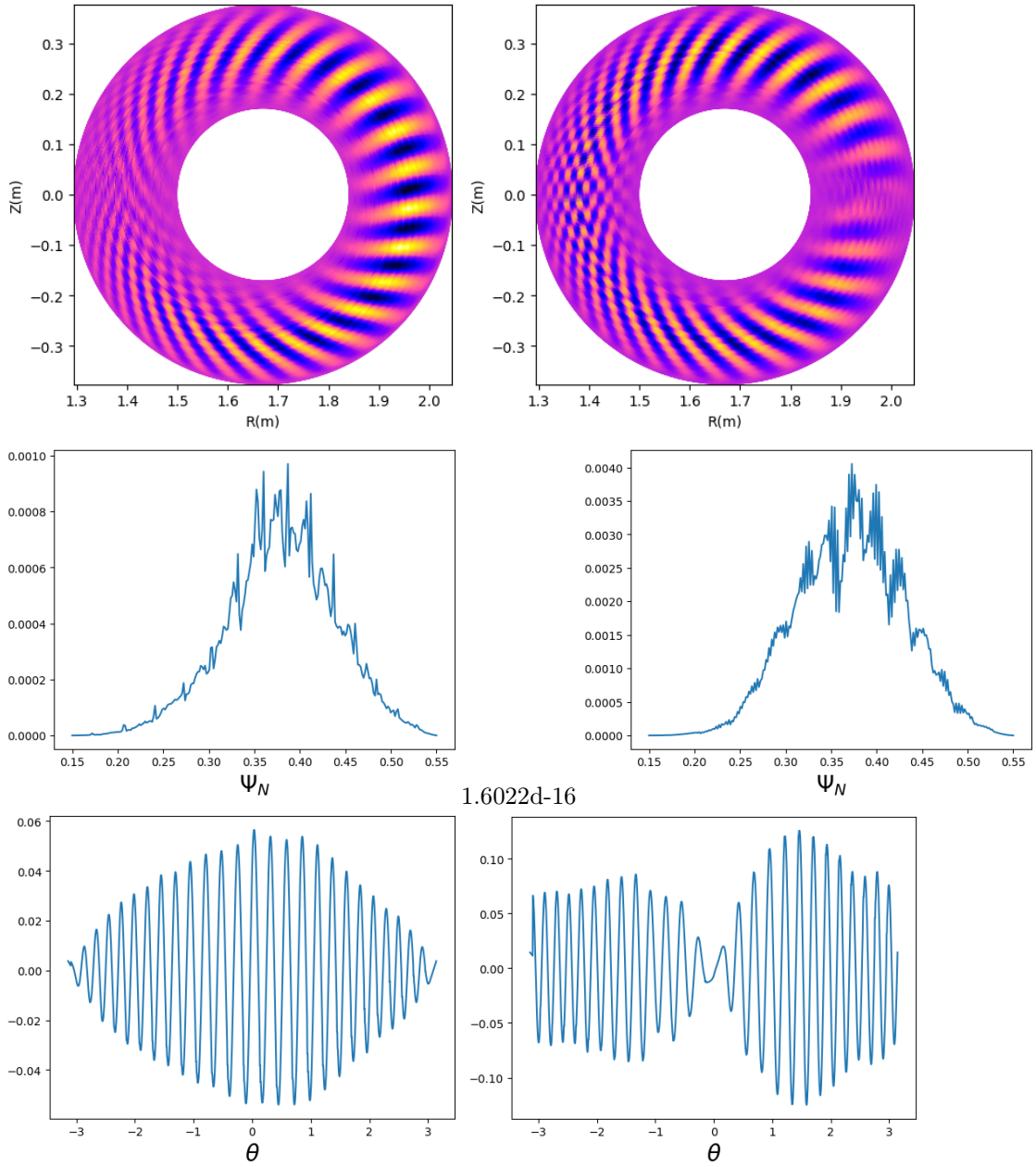

Figure 6. ITG $n = 20$.

ITG-KBM transition:

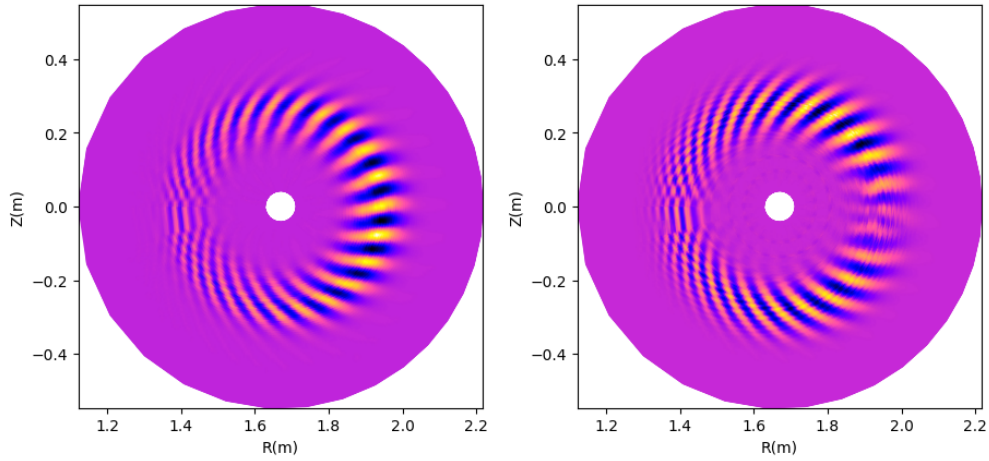

Figure 7. itg-kbm transition, seems to succeed, to be continued. left: real m_e, right: half m_e,(3) increase radial resolution
bottom left: $\delta\phi$, bottom right: $\delta A_\parallel$.

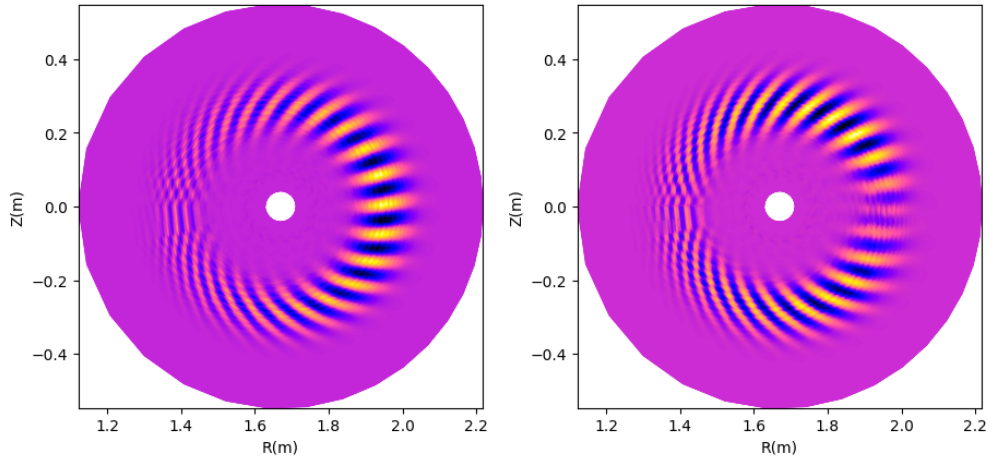

Figure 8. $\delta A_\parallel$ and $\delta\Phi$ for $n_e/n_{\mathrm{ref}} = 2.4$

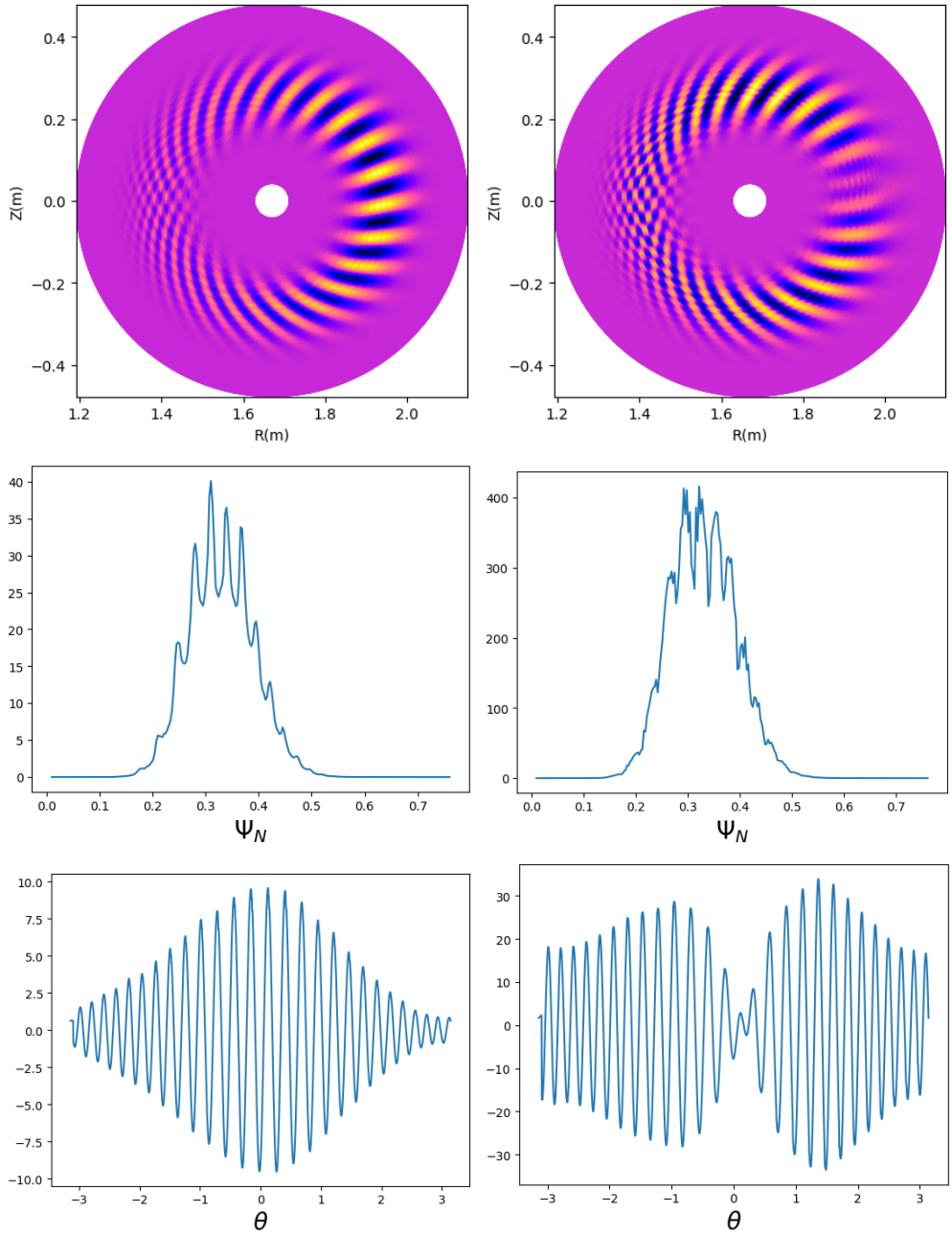

Figure 9. equal-arc-length poloidal angle.
$\delta\phi$ and $\delta A_{\|}$ at $n/n_{\mathrm{ref}} = 2.4$.

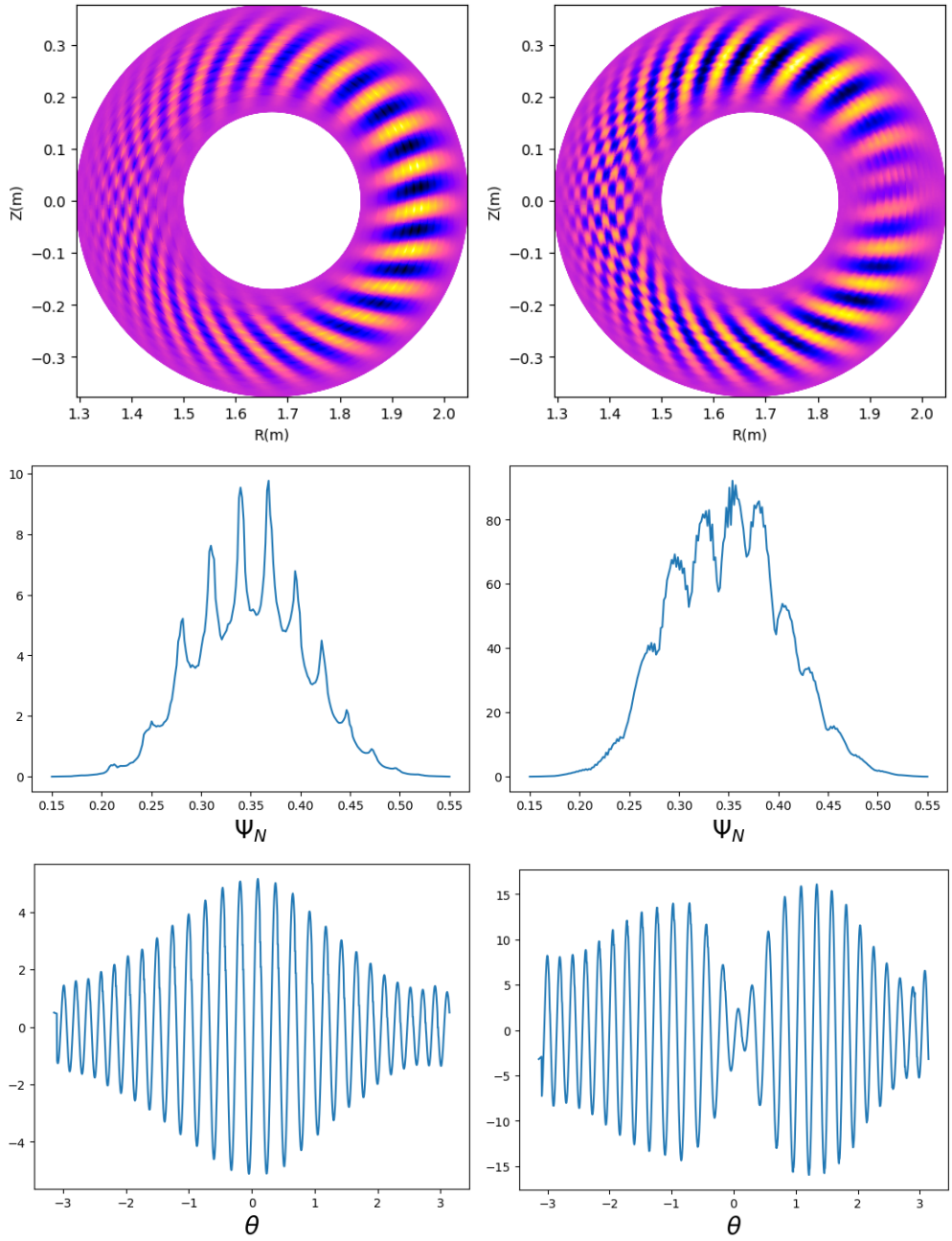

Figure 10. The same as Fig. 9 excep the radial range is narrowed.

A Adiabatic electron response

Assume that the total electron density satisfies the Boltzmann distribution on each magnetic surface, i.e.,

$$n_e = N_e \exp\left(-\frac{q_e \delta\Phi}{T_{e0}}\right) \approx N_e\left(1 - \frac{q_e \delta\Phi}{T_{e0}}\right), \tag{294}$$

where N_e is a radial function. Note that this does not imply that the equilibrium density is N_e (it just implies that the total density is N_e at the location where $\delta\Phi = 0$, which can still be different from the equilibrium density).

Further assume that the magnetic surface average of electron density perturbation ($\delta n_e = n_e - n_{e0}$) is zero, i.e.,

$$\langle n_e - n_{e0} \rangle_f = 0, \tag{295}$$

where n_{e0} is the equilibrium electron density, $\langle ... \rangle_f$ is the magntic surface averaging operator. Using Eq. (294) in the above condition, we get

$$N_e = n_{e0} \frac{1}{1 - \frac{q_e \langle \delta\Phi \rangle_f}{T_{e0}}} \approx n_{e0}\left(1 + \frac{q_e \langle \delta\Phi \rangle_f}{T_{e0}}\right). \tag{296}$$

Then expression (294) is written as

$$n_e = n_{e0}\left(1 + \frac{q_e \langle \delta\Phi \rangle_f}{T_{e0}}\right)\left(1 - \frac{q_e \delta\Phi}{T_{e0}}\right) \approx n_{e0}\left(1 + \frac{q_e \langle \delta\Phi \rangle_f}{T_{e0}} - \frac{q_e \delta\Phi}{T_{e0}}\right). \tag{297}$$

Then the perturbation $\delta n_e = n_e - n_{e0}$ is written as

$$\delta n_e = -n_{0e} \frac{q_e(\delta\Phi - \langle \delta\Phi \rangle_f)}{T_{e0}}. \tag{298}$$

This model for electron response is often called adiabatic electron.

A.1 Poisson's equation with adiabatic electron response

Pluging expression (298) into the Poisson equation (231), we get

$$-\varepsilon_0 \nabla_\perp^2 \delta\Phi - q_i \int \frac{q_i}{m_i} (\delta\Phi - \langle \delta\Phi \rangle_\alpha) \frac{\partial F_{i0}}{\partial \varepsilon} d\mathbf{v} + n_{0e} \frac{q_e^2 (\delta\Phi - \langle \delta\Phi \rangle_f)}{T_{e0}} = q_i \delta n_i'. \tag{299}$$

When solving the Poisson equation, the equation is Fourier expanded into toroidal harmonics. Each harmonic is independent of each other, so that they can be solved independently. For $n \neq 0$ harmonics, the $\langle \delta\Phi \rangle_f$ terms is zero and thus the the electron term is easy to handle because it's a local response. There is some difficulty in solving the $n = 0$ harmonic because the $\langle \delta\Phi \rangle_f$ term is nonzero.

I use the following method to obtain $\langle\delta\Phi\rangle_f$.

First slove the $n=0$ harmonic of the following equation:

$$-\varepsilon_0\nabla_\perp^2\delta\Phi' - q_i\int\frac{q_i}{m_i}(\delta\Phi' - \langle\delta\Phi'\rangle_\alpha)\frac{\partial F_{i0}}{\partial\varepsilon}d\mathbf{v} = q_i\delta n_i', \tag{300}$$

(i.e., Eq. (299) with the electron contribution dropped). Let $\delta\Phi'$ denote the solution to this equation. Then it can be proved that $\langle\delta\Phi'\rangle_f$ is equal to $\langle\delta\Phi\rangle_f$. [Proof:

Taking the flux surface average of Eq. (299), we obtain

$$-\varepsilon_0\langle\nabla_\perp^2\delta\Phi\rangle_f - q_i\left\langle\int\frac{q_i}{m_i}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_{i0}}{\partial\varepsilon}d\mathbf{v}\right\rangle_f = q_i\langle\delta n_i'\rangle_f, \tag{301}$$

where the adiabatic response disappears. to be continued.

]

Then solving Eq. (299) becomes easier because $\langle\delta\Phi\rangle_f$ term is known and can be moved to the right-hand side as a source term.

B Characteristic curves of Frieman-Chen gyrokinetic equation

Examining the left-hand side of Eq. (), the characteristic curves of this equation can be identified:

$$\frac{d\mathbf{X}}{dt} = v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D, \tag{302}$$

$$\frac{d\varepsilon}{dt} = 0, \tag{303}$$

$$\frac{d\mu}{dt} = 0. \tag{304}$$

(Note 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 denote the total guiding-center velocity by $\mathbf{V}_G$, i.e.,

$$\mathbf{V}_G = \frac{d\mathbf{X}}{dt} = \frac{d\mathbf{X}_0}{dt} + \frac{d\mathbf{X}_1}{dt} = v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D, \tag{305}$$

where

$$\frac{d\mathbf{X}_0}{dt} = v_\parallel\mathbf{e}_\parallel + \mathbf{V}_D. \tag{306}$$

In the above, $v_\parallel$ can be obtained by using $v_\parallel = \sigma\sqrt{2(\varepsilon - B\mu)}$. However σ is not easy to determine since it changes with time for trapped particles. In practice, we get $v_\parallel$ from an evolution equation, which can be obtained by combining the equations for ε and μ. Next, we derive this equation.

B.1 Time evolution equation for $\mathbf{v}_\parallel$

Using the definition $\mu = v_\perp^2/(2B_0)$, equation (304), i.e., $d\mu/dt = 0$, is written as

$$\frac{d}{dt}\left(\frac{v_\perp^2}{2B_0}\right) = 0, \tag{307}$$

which is written as

$$\frac{1}{B}\frac{d}{dt}(v_\perp^2) + v_\perp^2\frac{d}{dt}(\frac{1}{B_0}) = 0, \tag{308}$$

which can be further written as

$$\frac{d}{dt}(v_\perp^2) = 2\mu\frac{d}{dt}(B_0). \tag{309}$$

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

$$\frac{d}{dt}(v_\perp^2) = 2\mu\left(\frac{\partial B_0}{\partial t} + \frac{d\mathbf{X}}{dt}\cdot\nabla_X B_0 + \frac{d\varepsilon}{dt}\frac{\partial B_0}{\partial \varepsilon} + \frac{d\mu}{dt}\frac{\partial B_0}{\partial \mu}\right), \tag{310}$$

where $d\mathbf{X}/dt$, $d\varepsilon/dt$, and $d\mu/dt$ are given by Eq. (302), (303), and (304), respectively. Using Eqs. (302)-(304) and $\partial B_0/\partial t = 0$, equation (310) is reduced to

$$\frac{d}{dt}(v_\perp^2) = 2\mu\left(\frac{d\mathbf{X}}{dt}\cdot\nabla_X B_0\right). \tag{311}$$

On the other hand, equation (303), i.e., $d\varepsilon/dt = 0$, is written as

$$\frac{d}{dt}(v^2) = 0, \tag{312}$$

which can be further written as

$$\frac{d}{dt}(v_\parallel^2) = -\frac{d}{dt}(v_\perp^2). \tag{313}$$

Using Eq. (311), the above equation is written as

$$\frac{d}{dt}(v_\parallel) = -\frac{\mu}{v_\parallel}\left(\frac{d\mathbf{X}}{dt}\cdot\nabla_X B_0\right), \tag{314}$$

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. (302). Equation (314) 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. (302), 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 (314) is written as

$$\frac{d}{dt}(v_\parallel) = -\mu \mathbf{e}_\parallel \cdot \nabla_X B_0, \tag{315}$$

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. (315) and there is no parallel acceleration term $q v_\parallel \delta E_\parallel / m$ in Eq. (315). 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 deriving Frieman-Chen equation.)

C From $(\delta\Phi, \delta\mathbf{A})$ to $(\delta\mathbf{E}, \delta\mathbf{B})$

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

Note that

$$\begin{aligned} \delta\mathbf{B}_\perp &= \nabla\times\delta\mathbf{A} - (\mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A})\mathbf{e}_\parallel \\ &= \nabla\times(\delta\mathbf{A}_\perp + \delta A_\parallel \mathbf{e}_\parallel) - [\mathbf{e}_\parallel\cdot\nabla\times(\delta\mathbf{A}_\perp + \delta A_\parallel\mathbf{e}_\parallel)]\mathbf{e}_\parallel \end{aligned} \tag{316}$$

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)

$$\delta\mathbf{B}_\perp \approx \nabla\times\delta\mathbf{A}_\perp + \nabla\delta A_\parallel\times\mathbf{e}_\parallel - [\mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp + \mathbf{e}_\parallel\cdot(\nabla\delta A_\parallel\times\mathbf{e}_\parallel)]\mathbf{e}_\parallel \tag{317}$$

$$= \nabla\times\delta\mathbf{A}_\perp + \nabla\delta A_\parallel\times\mathbf{e}_\parallel - (\mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp)\mathbf{e}_\parallel \tag{318}$$

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

$$\nabla\times\delta\mathbf{A}_\perp = \left(-\frac{\partial\delta A_\phi}{\partial z}\right)\mathbf{e}_r + \left(\frac{\partial\delta A_r}{\partial z}\right)\mathbf{e}_\phi + \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial\delta A_r}{\partial\phi}\right]\mathbf{e}_\parallel. \tag{319}$$

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 (319) is written as

$$\nabla\times\delta\mathbf{A}_\perp \approx \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial\delta A_r}{\partial\phi}\right]\mathbf{e}_\parallel, \tag{320}$$

i.e., only the parallel component survive, which exactly cancels the last term in Eq. (318), i.e., equation (318) is reduced to

$$\delta\mathbf{B}_\perp = \nabla\delta A_\parallel \times \mathbf{e}_\parallel. \tag{321}$$

C.1.1 Using basis vectors in field-aligned coordinates

In terms of δB_{xL} and δB_{yL}, $\delta\mathbf{B}_\perp$ is written as

$$\delta\mathbf{B}_\perp = \delta B_{xL}\nabla x + \delta B_{yL}\nabla y = \nabla\delta A_\parallel \times \mathbf{e}_\parallel \tag{322}$$

Dotting the above equation by ∇x and ∇y, respectively, we obtain

$$\delta B_{xL}|\nabla x|^2 + \delta B_{yL}\nabla x\cdot\nabla y = \nabla x\cdot(\nabla\delta A_\parallel \times \mathbf{e}_\parallel), \tag{323}$$

$$\delta B_{xL}\nabla x\cdot\nabla y + \delta B_{yL}|\nabla y|^2 = \nabla y\cdot(\nabla\delta A_\parallel \times \mathbf{e}_\parallel). \tag{324}$$

Equations (323) and (324) can be further written as

$$\delta B_{xL}|\nabla x|^2 + \delta B_{yL}\nabla x\cdot\nabla y = -\nabla x\cdot\left(\frac{\partial\delta A_\parallel}{\partial y}\nabla y\times\mathbf{e}_\parallel + \frac{\partial\delta A_\parallel}{\partial z}\nabla z\times\mathbf{e}_\parallel\right), \tag{325}$$

and

$$\delta B_{xL}\nabla x\cdot\nabla y + \delta B_{yL}|\nabla y|^2 = -\nabla y\cdot\left(\frac{\partial\delta A_\parallel}{\partial x}\nabla x\times\mathbf{e}_\parallel + \frac{\partial\delta A_\parallel}{\partial z}\nabla z\times\mathbf{e}_\parallel\right), \tag{326}$$

The solution of this 2×2 system is expressed by Cramer's rule in the code.

Use $\mathbf{B}_0 = \Psi'\nabla x\times\nabla y$
$\mathbf{b} = \Psi'\nabla x\times\nabla y/B_0$

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

$$\begin{aligned}\delta B_\parallel &= \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}\\ &= \mathbf{e}_\parallel\cdot\nabla\times(\delta\mathbf{A}_\perp + \delta A_\parallel\mathbf{e}_\parallel)\end{aligned} \tag{327}$$

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)

$$\begin{aligned}\delta B_\parallel &\approx \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp + \mathbf{e}_\parallel\cdot(\nabla\delta A_\parallel\times\mathbf{e}_\parallel)\\ &= \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp\end{aligned} \tag{328}$$

[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

$$\nabla \times \delta\mathbf{A}_\perp = \left(-\frac{\partial \delta A_\phi}{\partial z}\right)\mathbf{e}_r + \left(\frac{\partial \delta A_r}{\partial z}\right)\mathbf{e}_r + \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial \delta A_r}{\partial \phi}\right]\mathbf{e}_\parallel \tag{329}$$

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 (319) is written as

$$\nabla \times \delta\mathbf{A}_\perp \approx \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial \delta A_r}{\partial \phi}\right]\mathbf{e}_\parallel, \tag{330}$$

Using this, equation (328) is written as

$$\delta B_\parallel = \frac{1}{r}\left[\frac{\partial}{\partial r}(r\delta A_\phi) - \frac{\partial \delta A_r}{\partial \phi}\right]. \tag{331}$$

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. (138), i.e.,

$$\delta\mathbf{V}_D = -\frac{q}{m}\nabla_X \langle \delta L \rangle_\alpha \times \frac{\mathbf{e}_\parallel}{\Omega}. \tag{332}$$

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

$$\begin{aligned}
\delta\mathbf{V}_D &= -\frac{q}{m}\nabla_X \langle \delta\Phi - \mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha \times \frac{\mathbf{e}_\parallel}{\Omega} \\
&= \frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega} \times \nabla_X \langle \delta\Phi\rangle_\alpha - \frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega} \times \nabla_X \langle v_\parallel \delta A_\parallel\rangle_\alpha \\
&\quad -\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega} \times \nabla_X \langle \mathbf{v}_\perp \cdot \delta\mathbf{A}_\perp\rangle_\alpha.
\end{aligned} \tag{333}$$

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

$$\begin{aligned}
\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega} \times \nabla_X \langle \delta\Phi\rangle_\alpha &= \frac{\mathbf{e}_\parallel}{B_0} \times \langle \nabla_X \delta\Phi\rangle_\alpha \\
&\approx \frac{\mathbf{e}_\parallel}{B_0} \times \langle \nabla_x \delta\Phi\rangle_\alpha \\
&\approx \frac{\mathbf{e}_\parallel}{B_0} \times \left\langle -\delta\mathbf{E} - \frac{\partial \delta\mathbf{A}}{\partial t}\right\rangle_\alpha \\
&\approx \frac{\mathbf{e}_\parallel}{B_0} \times \langle -\delta\mathbf{E}\rangle_\alpha \\
&\equiv \delta\mathbf{V}_E,
\end{aligned} \tag{334}$$

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. (333) is written

$$
\begin{aligned}
-\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle v_\parallel\delta A_\parallel\rangle_\alpha &\approx -\frac{q}{m}\frac{v_\parallel}{\Omega}\langle\mathbf{e}_\parallel\times\nabla_X(\delta A_\parallel)\rangle_\alpha \\
&\approx -\frac{q}{m}\frac{v_\parallel}{\Omega}\langle\mathbf{e}_\parallel\times\nabla_{\mathbf{x}}(\delta A_\parallel)\rangle_\alpha \qquad (335)\\
&= v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}, \qquad (336)
\end{aligned}
$$

which is called "magnetic fluttering" (this is actually not a real drift). In obtaining the last equality, use has been made of Eq. (321), 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 (333) 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 (328), i.e., $\delta B_\parallel = \mathbf{e}_\parallel\cdot\nabla\times\delta\mathbf{A}_\perp$, the above expression is written as

$$
\begin{aligned}
-\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle\mathbf{v}_\perp\cdot\delta\mathbf{A}_\perp\rangle_\alpha &= -\frac{1}{B_0}\langle\delta B_\parallel\mathbf{v}_\perp+\mathbf{e}_\parallel\times\mathbf{v}_\perp\cdot\nabla_x\delta\mathbf{A}_\perp\rangle_\alpha \\
&\approx -\frac{1}{B_0}\langle\delta B_\parallel\mathbf{v}_\perp+\mathbf{e}_\parallel\times\mathbf{v}_\perp\cdot\nabla_X\delta\mathbf{A}_\perp\rangle_\alpha \\
&\approx -\frac{1}{B_0}\langle\delta B_\parallel\mathbf{v}_\perp\rangle_\alpha-\frac{1}{B_0}\mathbf{e}_\parallel\times\langle\mathbf{v}_\perp\cdot\nabla_X\delta\mathbf{A}_\perp\rangle_\alpha. \\
&\approx -\frac{1}{B_0}\langle\delta B_\parallel\mathbf{v}_\perp\rangle_\alpha. \qquad (337)
\end{aligned}
$$

where use has been made of $\langle\mathbf{v}_\perp\cdot\nabla_X\delta\mathbf{A}_\perp\rangle_\alpha\approx 0$ (**seems wrong**), 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. (334), (336), and (337), expression (333) is finally written as

$$
\delta\mathbf{V}_D\equiv-\frac{q}{m}\nabla_X\langle\delta L\rangle_\alpha\times\frac{\mathbf{e}_\parallel}{\Omega}=\frac{\langle\delta\mathbf{E}\rangle_\alpha\times\mathbf{e}_\parallel}{B_0}+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}. \qquad (338)
$$

Using this, the first equation of the characteristics, equation (302), is written as

$$
\begin{aligned}
\frac{d\mathbf{X}}{dt} &= v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D \qquad (339)\\
&= v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+\frac{\langle\delta\mathbf{E}\rangle_\alpha\times\mathbf{e}_\parallel}{B_0}+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0} \\
&\equiv \mathbf{V}_G \qquad (340)
\end{aligned}
$$

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

[Note that

$$\frac{\partial\delta\mathbf{A}_\perp}{\partial t}=-(\delta\mathbf{E}_\perp+\nabla_\perp\delta\Phi), \tag{341}$$

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

$$\begin{aligned}\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha}{\partial t} &= v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t}+\mathbf{v}_\perp\cdot\frac{\partial\langle\delta\mathbf{A}\rangle_\alpha}{\partial t}\\ &= v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t}+\langle\mathbf{v}_\perp\cdot(-\delta\mathbf{E}-\nabla\delta\Phi)\rangle_\alpha\\ &\approx v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t}-\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha\end{aligned} \tag{342}$$

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. (342), the coefficient before $\partial F_0/\partial\varepsilon$ in Eq. (143) can be further written as

$$\begin{aligned}&-\frac{q}{m}\left[-\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha}{\partial t}-\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\nabla_X\langle\delta\Phi\rangle_\alpha\right]\\ &=-\frac{q}{m}\left[-v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t}+\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha-\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\left\langle-\delta\mathbf{E}-\frac{\partial\delta\mathbf{A}}{\partial t}\right\rangle_\alpha\right]\\ &\approx-\frac{q}{m}\left[-v_\parallel\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t}+\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha-\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\langle-\delta\mathbf{E}\rangle_\alpha+v_\parallel\left\langle\frac{\partial A_\parallel}{\partial t}\right\rangle_\alpha\right]\\ &=-\frac{q}{m}\left[\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D-\frac{q}{m}\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla_X\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha\right)\cdot\langle\delta\mathbf{E}\rangle_\alpha\right]\\ &\approx-\frac{q}{m}\left[\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle}{B_0}\right)\cdot\langle\delta\mathbf{E}\rangle_\alpha\right].\end{aligned} \tag{343}$$

Using Eq. (343) and (), gyrokinetic equation (143) is finally written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+\frac{\langle\delta\mathbf{E}\rangle_\alpha\times\mathbf{e}_\parallel}{B_0}+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\nabla_X\right]\delta f\\ &=-\left(\frac{\langle\delta\mathbf{E}\rangle_\alpha\times\mathbf{e}_\parallel}{B_0}+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\nabla_X F_0\\ &\quad-\frac{q}{m}\left[\langle\mathbf{v}_\perp\cdot\delta\mathbf{E}\rangle_\alpha+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\langle\delta\mathbf{E}\rangle_\alpha\right]\frac{\partial F_0}{\partial\varepsilon}.\end{aligned} \tag{344}$$

D Coordinate system and grid in **TEK** code

In (ψ, θ, ϕ) coordinates:

- ψ is the normalized poloidal magnetic flux, $\psi = (\Psi - \Psi_{\rm axis})/(\Psi_{\rm lcfs} - \Psi_{\rm axis})$.
- $\theta \in [-\pi, \pi)$ is increasing along the anti-clockwise direction when viewed along $\nabla\phi$ direction with $\theta = -\pi$ at the high-field-side midplane.
- ϕ is the toroidal angle of the right-handed cylindrical coordinates (R, ϕ, Z).

In this convention, the Jacobian of the (ψ, θ, ϕ) coordinate system, $\mathcal{J} = (\nabla\psi \cdot \nabla\theta \times \nabla\phi)$ is negative, i.e., (ψ, θ, ϕ) is a left-handed system. The field-line-following coordinate system (ψ, θ, α) is also a left-handed system. The coordinate system $(x = \psi, y = \alpha, z = \theta)$ is a right-handed system.

D.1 Poloidal grid

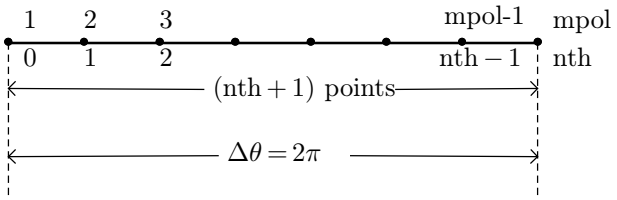

Figure 11. Poloidal grid for equilibrium quantities used in `TEK` and `GEM` code. The array starts at $\theta = -\pi$ and ends at $\theta = +\pi$ ($\theta = \pm\pi$ is chosen to be at the high-field side midplane in both the codes). `TEK` array index starts from 1 whereas GEM array index starts from 0. Hence `mpol=nth+1`. `nth` is denoted by `ntheta` in GEM. In `TEK`, `mpol` must be an odd number, so that the array has a midpoint indexed by `(mpol+1)/2`, which corresponds to $\theta = 0$.

I do not need to make connection with GEM's equilibrium poloidal array because there is no coupling of equilibrium quantities between the code written by me and the original code in GEM. The coupling happens for the perturbed quanties, whose poloidal grids need to be consistent.

- The gridpoint number along θ for perturbations is denoted by `mpol2`. Its value is not chosen by users. Instead, it is equal to `numproc/ntube`, i.e., `mpol2` is the total number of cells along θ for MPI parallelation. (`numproc` and `ntube` must be chosen in a way that makes `numproc/ntube` an integer.)

– Poloidal grid-points for perturbations are indexed as `0:mpol2-1`, with `0` corresponding to $\theta = -\pi$ and `mpol2` corresponding to $\theta = \pi$. Field equations are solved at `0: mpol2 − 1`. (The field at `mpol2`, i.e., $\theta = +\pi$, is not identical to that at $\theta = -\pi$ due to toroidal shift of a magnetic field line when it finish one poloidal loop, and is obtained by toroidally inerpolating the fields at $\theta = -\pi$.) No array for this is actually used in the code. The indexes just correspond to `gclr` of MPI processors.

– The equilibrium poloidal gridpoint number, `mpol`, must be chosen in a way that makes `(mpol-1)/mpol2` an integer, which can be larger than 1, i.e., one simulation cell can span several equilibrium cells.

D.2 Toroidal grid

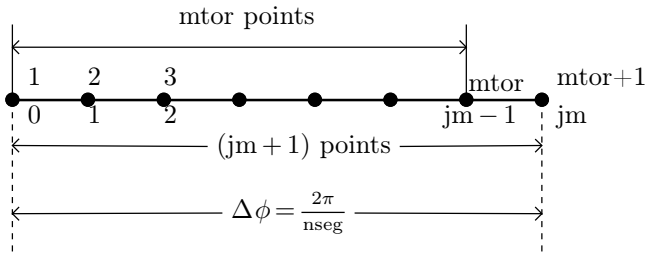

Figure 12. Toroidal grid used in `TEK` and GEM. In TEK, the toroidal array is `tor_1d_array(1:mtor+1)`. Source terms, δn and $\delta j_{\parallel}$, and EM fields are defined at `1:mtor`. GEM array index starts from 0 and ends at `jm`. It follows that `mtor=jm`.

D.3 Radial grid

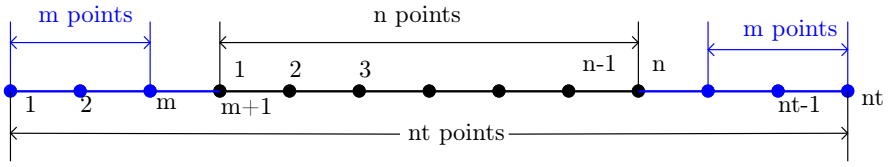

Figure 13. Radial grid used in `TEK`. Two radial arrays are used in the code, `radcor_1d_array(1:nt)` and `radcor_1d_array2(1:n)`, equilibrium is defined on the former grid, and perturbations are defined on the latter grid. Here $\mathrm{nt} = n + 2m$, m is the number of grid points in one of the two buffer regions (regions in blue color). In the code n is denoted by `nflux2` and nt is denoted by `nflux`, m is denoted by `points_in_buffer`. In `GEM` the radial array does no include the buffer regions and the index starts at `0` and ends at `imx`. Hence `n=imx+1`.

GEM array index system is better than TEK's because my system is not consistent: sometimes I use 0-based index and sometimes I use 1-based index, sometimes the index ends at n and sometimes ends at n+1. It is important to know accurately the transformation between the two systems.

E Implementation of gyrokinetics in particle-in-cell (PIC) codes

E.1 Inverse guiding-center transform in field-following coordinates (x,y,z)

Define two unit vectors:

$$\mathbf{a}_1 = \frac{\nabla x}{|\nabla x|}, \tag{345}$$

$$\mathbf{a}_2 = \mathbf{b} \times \frac{\nabla x}{|\nabla x|}, \tag{346}$$

both of which are perpendicular to the equilibrium magnetic field. Then a random gyro-radius vector $\boldsymbol{\rho}$ can be constructed by

$$\boldsymbol{\rho} = \rho[\cos(\alpha_0 + \alpha)\mathbf{a}_1 + \sin(\alpha_0 + \alpha)\mathbf{a}_2], \tag{347}$$

where α_0 is a random angle in $[0\colon 2\pi]$ among different markers, α is an angle variable in $[0\colon 2\pi]$ that will be discretized with uniform intervals for each marker, and $\rho = m\sqrt{2B_0\mu}\,/\,(B_0|q|)$ is the gyro-radius calculated using the magnetic field at the guiding-center location (x_0, y_0, z_0). Then the particle location (x_p, y_p, z_p) corresponding to this Larmor vector can be approximated by the Talor expansion at (x_0, y_0, z_0):

$$\begin{cases} x_p \approx x_0 + \boldsymbol{\rho} \cdot \nabla x \\ y_p \approx y_0 + \boldsymbol{\rho} \cdot \nabla y \\ z_p \approx z_0 + \boldsymbol{\rho} \cdot \nabla z \end{cases} \tag{348}$$

Using expressions (345), (346), (347), equation (348) is written as

$$\begin{cases} x_p = x_0 + \rho\cos(\alpha_0 + \alpha)|\nabla x| \\ y_p = y_0 + \rho\cos(\alpha_0 + \alpha)\dfrac{\nabla x}{|\nabla x|} \cdot \nabla y + \rho\sin(\alpha_0 + \alpha)\left(\mathbf{b} \times \dfrac{\nabla x}{|\nabla x|}\right) \cdot \nabla y \\ z_p = z_0 + \rho\cos(\alpha_0 + \alpha)\dfrac{\nabla x}{|\nabla x|} \cdot \nabla z + \rho\sin(\alpha_0 + \alpha)\left(\mathbf{b} \times \dfrac{\nabla x}{|\nabla x|}\right) \cdot \nabla z \end{cases}, \tag{349}$$

where ∇x, ∇y, and ∇z are all evaluated at the guiding-center position (x_0, y_0, z_0).

E.2 Monte-Carlo evaluation of distribution function moment at grid-points

Suppose that the 6D guiding-center phase-space ($\mathbf{X}$, $\mathbf{v}$) is described by (ψ, θ, ϕ, $v_\parallel$, $\alpha, v_\perp$) coordinates, where α is the gyro-angle. Denote the Jacobian of the coordinate system by $\mathcal{J}_{rv}$.

Suppose that we sample the 6D phase-space by using a probability function $P(\psi$, θ, ϕ, $v_\perp$, α, $v_\parallel)$. (Then the effecitive probability function used in rejection method is $P\mathcal{J}_r\mathcal{J}_v$.) Note that we will sample a distribution function $\delta f_g(\mathbf{X}, v_\perp, v_\parallel, \alpha)$ that happens to be independent of the gyro-angle α. Since δf_g is uniform in α, it is natural to smaple α using a uniform probability function, i.e., $P(\mathbf{X}, v_\perp, \alpha, v_\parallel)$ can be chosen to be independent of α. Then the weight is also independent of α. In terms of sampling δf_g, there is no need to specify α because both f_g and marker distribution $g = N_p P$ are indepedent of α. However, sampling α is still necessary in simulations. From the perspective of programming, it is ready to understand why α needs to be specified: when computing density at grid-points, we need to compute particle locations from the guiding-center locations which requires us to specify the gyro-angle α for each marker. Specifically, markers in guiding-center space $(\mathbf{X}, v_\perp, \alpha, v_\parallel)$ with fixed $(\mathbf{X}, v_\perp$, $v_\parallel)$ but different values of α correspond to markers in particle space with different locations (and different α'). In the PIC method of computing δn or $\delta j_\parallel$ at fixed $\mathbf{x}$, the contribution of each particle marker to the density is affected by the distance of the particle marker to the fixed $\mathbf{x}$. Therefore different values of α contributes differently to δn or $\delta j_\parallel$, and thus the resolution of α matters.

For a marker with coordinators $(\mathbf{X}, v_\perp, v_\parallel, \alpha)$, the corresponding particle position can be calculated by using the inverse guiding-center transformation (29). Then we can deposit the marker weight to grid-points in the same way that we do in conventional PIC simulations. Looping over all markers, we build the particle density at grids.

Compared with conventional PIC methods, where particle positions are directly sampled, what is the benefit of sampling guiding-center positions and then transform them back to the particle positions? More computations are involved since we need to numerically perform the inverse guiding-center transformation. The answer lies in the important fact that the distribution function $f_g(\mathbf{X}, v_\perp, v_\parallel, \alpha)$ that needs to be numerically evaluated in gyrokinetic simulation is actually independent of the gyro-angle α. Furthermore, we use a probability density function P that is independent of α to sample the 6D phase space $(\mathbf{X}, v_\perp, v_\parallel, \alpha)$. Then the marker weight $w \equiv f_g/(N_p P)$ is independent of α, where N_p is number of markers loaded. This fact enable resolution in particle coordinates can be increased in a trivial way, as described below.

Suppose we have a concrete sampling of f_g in the 6D phase-space ($\mathbf{X}$, $v_\perp$, $v_\parallel$, α), i.e., $(\mathbf{X}_j, v_{\perp j}, v_{\parallel j}, \alpha_j, w_j)$ with $j = 1, ..., N_p$, then we can do the inverse guiding-center transform and then deposit particles to the grids to obtain a moment (e.g. density).

Since $P\mathcal{J}_{rv}$ is independent of α, the sampli Maybe the real reason is that I lack the interests in learning them.ng α_j with $j = 0, 1, ..., N_p$ are uniform distributed random numbers. Therefore we can generate another sampling of α (denoted by α'_j with $j = 1, ..., N_p$) using random number generators and combine α'_j with the old sampling $(\mathbf{X}_j, v_{\perp j}, v_{\parallel j})$ to obtain a new sampling $(\mathbf{X}_j, v_{\perp j}, v_{\parallel j}, \alpha'_j, w'_j)$. The values of w at the new sampling points are equal to the original values, i.e., $w'_j = w_j$ since the particle weight $w = f_g/(N_p P)$ is independent of α. Using the new sampling and following the same procedures given above, we can estimate values of the moment at grid-points again, which will differ from the estimation obtained using the old sampling because the resulting sampling of $\mathbf{x}$ and α' for f_p is different. Taking the average of the two estimations will give a more accurate estimation because the resolution in $\boldsymbol{x}$ and α' for f_p is increased.

[Note that even if $P\mathcal{J}_{rv}$ is dependent on α due to the possible dependence of $\mathcal{J}_{rv}$ on α, we still can easily generate a new set of sampling which differs from the old sampling only in α. Specifically, in the rejection method, we use the old sampling $(\mathbf{X}_j, v_{\perp j}, v_{\parallel j})$ for each reject step and only adjust α to satisfy the acceptance criteria.]

We can also construct a new sampling by replacing α_j with $\alpha_j + \Delta$, where Δ is a constant. Then the new set of sampling $(\mathbf{X}_j, v_{\perp j}, v_{\parallel j}, \alpha_j + \Delta, w_j)$ with $j = 1, ..., N_p$ is still a consistent set of sampling.

In doing the deposition, each marker has a single gyro-angle. Due to the independence of the weight ($w = f_g/(N_p\, P)$) of the gyro-angle, the particle phase space resolution can be increased in a way that there can be several gyro-angles for a single $(\mathbf{X}_j, v_{\perp j}, v_{\parallel j})$. This gives the wrong impression that the gyro-angle of a guiding-center marker is arbitrary. The correct understanding is that given above, i.e., we do several (say 4) separate sampling of the phase space and then average the results.

In the code, the gyro-angle is defined relative to the direction $\nabla\psi/|\nabla\psi|$ at the guiding-center position. Specifically, the gyro-angle is defined as the included angle between $\nabla\psi/|\nabla\psi|$ and $-\mathbf{v}\times\mathbf{e}_\parallel(\mathbf{x})/\Omega(\mathbf{x})$, where $\mathbf{e}_\parallel(\mathbf{x})/\Omega(\mathbf{x})$ is approximated by the value at the guiding-center location.

Then ($\mathbf{X}_j$, $v_{\perp j}$, $v_{\parallel j}$) can be evolved by using the guiding-center motion equation. (It is obvious how to evalute the gyro-averaging of the electromagnetic fields needed in pusing markers.)

E.3 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 $\mathcal{J}_r = \mathcal{J}(\psi, \theta)$ is the Jacobian of the coordinates (ψ, θ, ϕ). Suppose that we sample the 6D phase-space by using the following probability density function:

$$P(\psi, \theta, \phi, v_\parallel, v_\perp, \alpha) = \frac{1}{V_r}\Big(\frac{m}{2\pi T}\Big)^{3/2}\exp\left[-\frac{m\,(v_\parallel^2 + v_\perp^2)}{2T}\right], \tag{350}$$

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

$$\int_{V_r}\int Pd\mathbf{v}d\mathbf{X} = \int_{V_r}\int_{-\infty}^{+\infty}\int_0^{\infty}\int_0^{2\pi} Pv_\perp d\alpha dv_\perp dv_\parallel \mathcal{J}_r d\psi d\theta d\phi = 1. \tag{351}$$

I 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

$$P' = |\mathcal{J}_r\mathcal{J}_v|P = |\mathcal{J}_r(\psi,\theta)|v_\perp P \tag{352}$$

Note that P' does not depend on the gyro-angle α.]

Then the weight of a marker sampling δf_g is written as

$$w = \frac{\delta f_g(\psi,\theta,\phi,v_\parallel,v_\perp)}{N_p P}. \tag{353}$$

Since both f_g and P are independent of the gyro-angle α, w is also independent of α.

The numerical representation of δf_g is written

$$\delta\tilde{f}_g = \frac{1}{\mathcal{J}_r v_\perp}\sum_j^{N_p} w_j\delta(\psi-\psi_j)\delta(\theta-\theta_j)\delta(\phi-\phi_j)\delta(v_\parallel-v_{\parallel j})\delta(v_\perp-v_{\perp j})\delta(\alpha-\alpha_j). \tag{354}$$

Although the distribution function δf_g to be sampled is independent of the gyro-angle α, we still need to specify 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 (the quantity of intest to us, e.g., density, is the integration of the particle distribution function at fixed spatial location, and the distribution at the fixed location is not uniform in the gyro-angle. So the resolution over α matters), we need to load more 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.

$$n(\mathbf{x}) = \frac{n_1(\mathbf{x})+n_2(\mathbf{x})+n_3(\mathbf{x})+n_4(\mathbf{x})}{4}. \tag{355}$$

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 implementation, we choose N sampling points that are evenly distributed on the gyro-ring (N is usually 4 as a compromise between efficiency and accuracy). And the weight is evenly split by the N sub-markers on the gyro-ring. Therefore each sub-marker have a Monte-Carlo weight w_j / N, where w_j is the weight of the jth marker. Then calculating the integration (361) 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. 14. However, interpreting in this way is confusing to me because: why can you split a single sampling into N different samplings? 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 a long time for me to finally find this way of understanding.

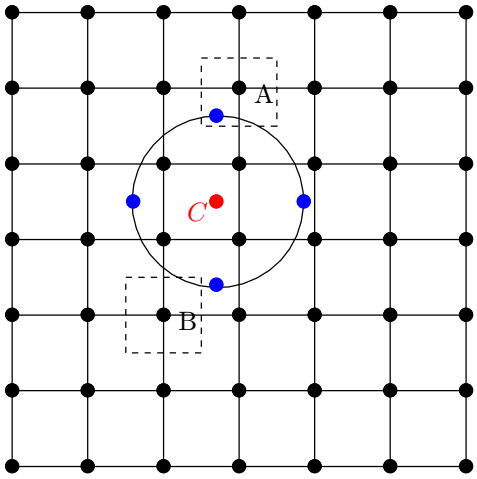

Figure 14. The spatial grid in the plane perpendicular to the equilibrium magnetic field, a guiding-center marker C, and its gyro-ring with 4 sampling points (sub-markers) on it. The 4 sub-markers are calculated by using the transformation (29) (inverse guiding-center transform). Assume that Monte-Carlo weight of the guiding-center marker C is w_j. Then The Monte-Carlo weight of each sub-marker is $w_j / 4$. The value of integration (361) 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 the 2D projection, the cell is 3D and 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 nothing to grid point 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.

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 an 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 find 4 arbitrary points on the gyro-ring that are easy to calculate. Some codes (e.g. ORB5) introduces a random variable to rotate these 4 points for different markers so that the gyro-angle can be sampled less biased.

From the view of particle simulations, the gyrokinetic model can be considered as a noise reduction method, where the averaging over the gyro-angle is equivalent to a time averaging over a gyro-period, which reduces the fluctuation level (in both time and space) associated with evaluating the Monte-Carlo phase integration. Here the averaging in gyrokinetic particle simulation refers to taking several points on a gyro-ring when depositing markers to spatial grids to obtain the density and current on the grids. (Another gyro-averaging appears in evaluating the guiding-center drift.) In gyrokinetic particle simulation, even a step size smaller than a gyro-period is taken, the quantities used in the model is still the ones averaged over one gyro-period. In this sense, a gyrokinetic simulation is only meaningful when the time step size is larger than one gyro-period. [**Some authors may disagree with that the gyro-averaging is a time-averaging. They may consider the gyro-averaging as the phase-space integration over the gyro-angle coordinate. This view seems to be right in Euler simulations but seems to be wrong in particle simulations. The reason is as follows. For each marker, choose a random gyro-phase and then do the inverse transformation to obtain particle position, and sum over all markers (this corresponds to phase-space Monte-Carlo integration, which include the gyro-angle integration, so no further gyro-angle integration is needed); choose another random gyro-phase and repeat the above procedure (this can be interpreted as do the phase-space Monte-Carlo integration at another time), choose further random gyro-phase for each marker and repeat. Finally averaging all the above values to obtain the final estimation of the phase-space integration. This amounts to a time-averaging over a gyro-motion. In summary, sampling several times with different gyro-phases for each marker and taking the average amounts to the time averaging over gyro-motion**]

When doing the time-average over the ion cyclotron motion, the time variation of the low-frequency mode is negligible and only the spatial variation of the modes is important. For the gyro-motion, only the gyro-angle is changing and all the other variables, $(\mathbf{X}, v_\perp, v_\parallel)$, are approximately constant. As a result, this time averaging finally reduces to a gyro-averaging.

I prefer to reason in terms of particle position and velocity, and consider the guiding-center location as an image of the particle position. When working in the guiding-center coordinates, I prefer to reason by transforming back to the particle position. This reasoning is clear and help me avoid some confusions I used to have.

In coding, the initial sampling of the phase space is performed for particles ($(\mathbf{v}, \mathbf{x})$ of particles) and then is transformed to guiding-center coordinates.

E.4 Distribution function transform**check

In the above, we assume that $\mathbf{X}$ and $\mathbf{x}$ are related to each other by the guiding-center transformation (27) or (29) , i.e., $\mathbf{x}$ and $\mathbf{X}$ are not independent. For some cases, it may be convenient 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,

$$f_p(\mathbf{x},\mathbf{v}) = \int f_g(\mathbf{X},\mathbf{v})\delta^3(\mathbf{X}-\mathbf{x}+\boldsymbol{\rho})d\mathbf{X}, \tag{356}$$

where $\mathbf{x}$ and $\mathbf{X}$ are considered as independent variables, $\boldsymbol{\rho}$ is the gyroradius evaluated at $\mathbf{x}$, $\delta^3(\mathbf{x}-\mathbf{X}-\boldsymbol{\rho})$ 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:

$$\delta^3(\mathbf{x}) = \frac{1}{|\mathcal{J}|}\delta(x_1)\delta(x_2)\delta(x_3), \tag{357}$$

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

$$f_g(\mathbf{X},\mathbf{v}) = \int f_p(\mathbf{x},\mathbf{v})\delta^3(\mathbf{x}-\mathbf{X}-\boldsymbol{\rho})d\mathbf{x}, \tag{358}$$

is a transformation that transforms an arbitrary function from the the particle coordinates to the guiding-center coordinates.

E.5 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

$$n(\mathbf{x}) = \int f_p(\mathbf{x},\mathbf{v})d\mathbf{v}. \tag{359}$$

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 $\mathbf{v}$, which is required by the integration in Eq. (359), means the guiding-center variable $\mathbf{X}$ is changing according to Eq. (27). Using Eq. (34), expression (359) is written as

$$n(\mathbf{x}) = \int f_g(\mathbf{X}(\mathbf{x},\mathbf{v}),\mathbf{v})d\mathbf{v}, \tag{360}$$

As is mentioned above, the $d\mathbf{v}$ integration in Eq. (360) 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)

curve integral along the the curve $\mathbf{X}(\mathbf{v}) = \mathbf{x} - \boldsymbol{\rho}(\mathbf{x}, \mathbf{v})$ with $\mathbf{x}$ being constant. Treating $\mathbf{X}$ and $\mathbf{x}$ as independent variables and using the Dirac delta function δ, this curve integral can be written as the following double integration over the independent variables $\mathbf{X}$ and $\mathbf{v}$:

$$n(\mathbf{x}) = \iint f_g(\mathbf{X}, \mathbf{v}) \delta^3(\mathbf{X} - \mathbf{x} + \boldsymbol{\rho}) d\mathbf{v} d\mathbf{X}. \tag{361}$$

Another perspective of interpreting Eq. (361) is that we are first using the transformation (356) to transform f_g to f_p and then integrating f_p in the velocity space.

F Diamagnetic flow **check**

The perturbed distribution function δF given in Eq. () contains two terms. The first term is gyro-phase dependent while the second term is gyro-phase independent. The perpendicular velocity moment of the second term will give rise to the so-called diamagnetic flow. 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 () is written as

$$\delta F_g = \frac{q}{m}(\delta\Phi - \langle\delta\Phi\rangle_\alpha)\frac{\partial F_{0g}}{\partial \varepsilon} + \delta f_g. \tag{362}$$

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,

$$n\mathbf{U}_\perp = \int \mathbf{v}_\perp \delta f_p(\mathbf{x}, \mathbf{v}) d\mathbf{v}, \tag{363}$$

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, equation (363) is written as

$$n\mathbf{U}_\perp = \int \mathbf{v}_\perp \delta f_g(\mathbf{x} - \boldsymbol{\rho}, \mathbf{v}) d\mathbf{v}. \tag{364}$$

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

$$\delta f_g(\mathbf{x} - \boldsymbol{\rho}, \mathbf{v}) \approx \delta f_g(\mathbf{x}, \mathbf{v}) - \boldsymbol{\rho} \cdot \nabla \delta f_g(\mathbf{x}, \mathbf{v}). \tag{365}$$

Plugging this expression into Eq. (364), we obtain

$$n\mathbf{U}_\perp \approx \int \mathbf{v}_\perp \delta f_g(\mathbf{x}, \mathbf{v}) d\mathbf{v} - \int \mathbf{v}_\perp \boldsymbol{\rho} \cdot \nabla \delta f_g(\mathbf{x}, \mathbf{v}) d\mathbf{v} \tag{366}$$

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. (366) is reduced to

$$n\mathbf{U}_\perp = -\int \mathbf{v}_\perp \boldsymbol{\rho} \cdot \nabla \delta f_g(\mathbf{x}, \mathbf{v}) d\mathbf{v} \tag{367}$$

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

$$\begin{aligned} n\mathbf{U}_\perp &= \int \mathbf{v}_\perp \frac{\mathbf{v}\times\mathbf{e}_\parallel}{\Omega}\cdot\nabla\delta f_g(\mathbf{x},\mathbf{v})d\mathbf{v} \\ &= \int \mathbf{v}_\perp\Big(\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla\delta f_g(\mathbf{x},\mathbf{v})\Big)\cdot\mathbf{v}_\perp d\mathbf{v}. \\ &= \int \mathbf{v}_\perp\mathbf{H}\cdot\mathbf{v}_\perp d\mathbf{v}, \end{aligned} \tag{368}$$

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. (368). In terms of velocity space cylindrical coordinates $(v_\parallel, v_\perp, \alpha)$, $\mathbf{v}_\perp$ is written as

$$\mathbf{v}_\perp = v_\perp(\hat{\mathbf{x}}\cos\alpha + \hat{\mathbf{y}}\sin\alpha), \tag{369}$$

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

$$\mathbf{H} = H_x\hat{\mathbf{x}} + H_y\hat{\mathbf{y}}, \tag{370}$$

where H_x and H_y are independent of α. Using these in Eq. (368), we obtain

$$\begin{aligned} n\mathbf{U}_\perp &= \int v_\perp(\hat{\mathbf{x}}\cos\alpha + \hat{\mathbf{y}}\sin\alpha)\, v_\perp(H_x\cos\alpha + H_y\sin\alpha)d\mathbf{v} \\ &= \int v_\perp^2[\hat{\mathbf{x}}(H_x\cos^2\alpha + H_y\sin\alpha\cos\alpha) + \hat{\mathbf{y}}(H_x\cos\alpha\sin\alpha + H_y\sin^2\alpha)]\, d\mathbf{v}. \end{aligned} \tag{371}$$

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

$$\begin{aligned} n\mathbf{U}_\perp &= \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp \int_0^{2\pi} v_\perp^2[\hat{\mathbf{x}}(H_x\cos^2\alpha + H_y\sin\alpha\cos\alpha) + \hat{\mathbf{y}}(H_x\cos\alpha\sin\alpha + \\ &\quad H_y\sin^2\alpha)]d\alpha \\ &= \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp \int_0^{2\pi} v_\perp^2(\hat{\mathbf{x}}H_x\cos^2\alpha + \hat{\mathbf{y}}H_y\sin^2\alpha)d\alpha \\ &= \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp [v_\perp^2(\hat{\mathbf{x}}H_x\pi + \hat{\mathbf{y}}H_y\pi)] \\ &= \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp [v_\perp^2\mathbf{H}\pi] \\ &= \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp [v_\perp^2\frac{\mathbf{e}_\parallel}{\Omega}\times\nabla\delta f_g(\mathbf{x},\mathbf{v})\pi] \\ &= \frac{\mathbf{e}_\parallel}{\Omega}\times\nabla\int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp \delta f_g(\mathbf{x},\mathbf{v})\frac{v_\perp^2}{2}2\pi \\ &= \frac{\mathbf{e}_\parallel}{m\Omega}\times\nabla\delta p_\perp, \end{aligned} \tag{372}$$

where

$$\begin{aligned}\delta p_\perp &\equiv \int_{-\infty}^{\infty} dv_\parallel \int_0^\infty v_\perp dv_\perp \delta f_g(\mathbf{x},\mathbf{v}) \frac{m v_\perp^2}{2} 2\pi \\ &= \int \delta f_g(\mathbf{x},\mathbf{v}) \frac{m v_\perp^2}{2} d\mathbf{v}, \end{aligned} \tag{373}$$

is the perpendicular pressure carried by $\delta f_g(\mathbf{x},\mathbf{v})$. The flow given by Eq. (372) is called the diamagnetic flow.

G Transform gyrokinetic equation from $(\mathbf{X}, \mu, \varepsilon, \alpha)$ to $(\mathbf{X}, \mu, v_\parallel, \alpha)$ coordinates

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

$$\begin{cases} \mathbf{X}' = \mathbf{X} \\ \mu' = \mu \\ \alpha' = \alpha \\ v_\parallel = \sigma\sqrt{2(\varepsilon - \mu B_0(\mathbf{X}))} \end{cases} \tag{374}$$

Use this definition and the chain rule, we obtain

$$\begin{aligned}\left.\frac{\partial}{\partial \mathbf{X}}\right|_{\mu,\varepsilon,\alpha} &= \frac{\partial \mathbf{X}'}{\partial \mathbf{X}} \cdot \frac{\partial}{\partial \mathbf{X}'} + \frac{\partial \mu'}{\partial \mathbf{X}} \frac{\partial}{\partial \mu'} + \frac{\partial v_\parallel}{\partial \mathbf{X}} \frac{\partial}{\partial v_\parallel} + \frac{\partial \alpha'}{\partial \mathbf{X}} \frac{\partial}{\partial \alpha'} \\ &= \frac{\partial}{\partial \mathbf{X}'} - \frac{\mu}{v_\parallel} \frac{\partial B_0}{\partial \mathbf{X}} \frac{\partial}{\partial v_\parallel}, \end{aligned} \tag{375}$$

and

$$\begin{aligned}\left.\frac{\partial}{\partial \varepsilon}\right|_{\mathbf{X},\mu,\alpha} &= \frac{\partial \mathbf{X}'}{\partial \varepsilon} \frac{\partial}{\partial \mathbf{X}'} + \frac{\partial \mu'}{\partial \varepsilon} \frac{\partial}{\partial \mu'} + \frac{\partial v_\parallel}{\partial \varepsilon} \frac{\partial}{\partial v_\parallel} + \frac{\partial \alpha'}{\partial \varepsilon} \frac{\partial}{\partial \alpha'} \\ &= \frac{1}{v_\parallel} \frac{\partial}{\partial v_\parallel}. \end{aligned} \tag{376}$$

Then, in terms of $(\mathbf{X}', \mu', v_\parallel, \alpha')$, equation (135) is written

$$\begin{aligned}&\left[\frac{\partial}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \frac{\partial}{\partial \mathbf{X}'}\right] \delta G_0 - (v_\parallel \mathbf{e}_\parallel + \mathbf{V}_D + \delta\mathbf{V}_D) \cdot \frac{\mu}{v_\parallel} \frac{\partial B_0}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial v_\parallel} \\ =& -\delta\mathbf{V}_D \cdot \left(\frac{\partial F_0}{\partial \mathbf{X}'} - \frac{\mu}{v_\parallel} \frac{\partial B_0}{\partial \mathbf{X}} \frac{\partial F_0}{\partial v_\parallel}\right) - \frac{q}{m} \frac{\partial \langle \delta L \rangle_\alpha}{\partial t} \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel}, \end{aligned} \tag{377}$$

Dropping terms of order higher than $O(\lambda^2)$, equation (377) is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\frac{\partial}{\partial \mathbf{X}'}\right]\delta G_0-\mathbf{e}_\parallel\cdot\mu\frac{\partial B_0}{\partial \mathbf{X}}\frac{\partial \delta G_0}{\partial v_\parallel}\\&=-\delta\mathbf{V}_D\cdot\left(\frac{\partial F_0}{\partial \mathbf{X}'}\right)+\left(\delta\mathbf{V}_D\cdot\mu\frac{\partial B_0}{\partial \mathbf{X}}-\frac{q}{m}\frac{\partial\langle\delta L\rangle_\alpha}{\partial t}\right)\frac{\partial F_0}{\partial v_\parallel}\frac{1}{v_\parallel},\end{aligned}\tag{378}$$

The above equation 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., $-\mathbf{e}_\parallel\cdot\mu\nabla B_0$, which is independent of any perturbations.

Note that the gyro-averaging operator in ($\mathbf{X}'$, μ', $v_\parallel$, α') coordinates is identical 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

$$\delta\mathbf{V}_D=\frac{\mathbf{e}_\parallel\times\nabla_X\langle\delta\phi\rangle_\alpha}{B_0}+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0},\tag{379}$$

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. (375) indicates that $\partial\delta\phi/\partial\mathbf{X}=\partial\delta\phi/\partial\mathbf{X}'$.

Following the same procedures, equation (143) in terms of $(\mathbf{X}', \mu', v_\parallel, \alpha')$ is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\frac{\partial}{\partial \mathbf{X}'}\right]\delta f-\mathbf{e}_\parallel\cdot\mu\nabla B_0\frac{\partial\delta f}{\partial v_\parallel}\\&=-\delta\mathbf{V}_D\cdot\left(\frac{\partial F_0}{\partial \mathbf{X}'}\right)+\delta\mathbf{V}_D\cdot\mu\nabla B_0\frac{\partial F_0}{\partial v_\parallel}\frac{1}{v_\parallel}\\&-\frac{q}{m}\left[-\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha}{\partial t}-\left(v_\parallel\mathbf{e}_\parallel+\mathbf{V}_D+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\nabla_X\langle\delta\phi\rangle_\alpha\right]\frac{\partial F_0}{\partial v_\parallel}\frac{1}{v_\parallel}.\end{aligned}\tag{380}$$

next, try to recover the equation in Mishchenk's paper:

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

$\delta L=\delta\Phi-\mathbf{v}\cdot\delta\mathbf{A}$,

$$\frac{\delta\mathbf{V}_D}{v_\parallel}=-\frac{q}{m v_\parallel}\nabla_X\langle\delta\Phi-v_\parallel A_\parallel\rangle_\alpha\times\frac{\mathbf{e}_\parallel}{\Omega}.\tag{381}$$

$$-\frac{q}{m}\left[-\frac{\partial\langle\delta A_{\parallel}\rangle_{\alpha}}{\partial t}-\left(\mathbf{e}_{\parallel}+\frac{\mathbf{V}_D}{v_{\parallel}}+\frac{q}{m}\nabla_X\langle A_{\parallel}\rangle_{\alpha}\times\frac{\mathbf{e}_{\parallel}}{\Omega}\right)\cdot\nabla_X\langle\delta\phi\rangle_{\alpha}\right]\frac{\partial F_0}{\partial v_{\parallel}}.$$

$$-\frac{q}{m}\left[-\frac{\partial\langle\delta A_{\parallel}{}^{(h)}\rangle_{\alpha}}{\partial t}-\left(\frac{\mathbf{V}_D}{v_{\parallel}}+\frac{q}{m}\nabla_X\langle A_{\parallel}\rangle_{\alpha}\times\frac{\mathbf{e}_{\parallel}}{\Omega}\right)\cdot\nabla_X\langle\delta\phi\rangle_{\alpha}\right]\frac{\partial F_0}{\partial v_{\parallel}}.$$

G.1 Recover equation in W. Deng's 2011 NF paper

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

$$v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D=\frac{\mathbf{B}_0^{\star}}{B_{\parallel 0}^{\star}}v_{\parallel}+\frac{\mu}{\Omega B_{\parallel 0}^{\star}}\mathbf{B}_0\times\nabla B_0+\frac{1}{B_0\,B_{\parallel 0}^{\star}}\mathbf{E}_0\times\mathbf{B}_0 \tag{382}$$

where

$$\mathbf{B}_0^{\star}=\mathbf{B}_0+B_0\frac{v_{\parallel}}{\Omega}\nabla\times\mathbf{b}, \tag{383}$$

$$B_{\parallel}^{\star}\equiv\mathbf{b}\cdot\mathbf{B}^{\star}=B\left(1+\frac{v_{\parallel}}{\Omega}\mathbf{b}\cdot\nabla\times\mathbf{b}\right), \tag{384}$$

Using $B_{\parallel 0}^{\star}\approx B_0$, then expression (382) is written as

$$v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D=v_{\parallel}\mathbf{b}+\underbrace{\frac{v_{\parallel}^2}{\Omega}\nabla\times\mathbf{b}}_{\text{curvature drift}}+\underbrace{\frac{\mu}{\Omega B_0}\mathbf{B}_0\times\nabla B_0}_{\nabla B\text{ drift}}+\underbrace{\frac{1}{B_0^2}\mathbf{E}_0\times\mathbf{B}_0}_{E\times B\text{ drift}}, \tag{385}$$

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 (refer to Sec. C.3)

$$\delta\mathbf{V}_D=\frac{\mathbf{e}_{\parallel}\times\nabla_X\langle\delta\phi\rangle_{\alpha}}{B_0}+v_{\parallel}\frac{\langle\delta\mathbf{B}_{\perp}\rangle_{\alpha}}{B_0}. \tag{386}$$

Using the above results, equation (380) is written as

$$\begin{aligned}&\left[\frac{\partial}{\partial t}+(v_{\parallel}\mathbf{e}_{\parallel}+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\frac{\partial}{\partial\mathbf{X}'}\right]\delta f-\mathbf{e}_{\parallel}\cdot\mu\nabla B_0\frac{\partial\delta f}{\partial v_{\parallel}}\\&=-\delta\mathbf{V}_D\cdot\left(\frac{\partial F_0}{\partial\mathbf{X}'}\right)+\left(\frac{\mathbf{e}_{\parallel}\times\nabla_X\langle\delta\phi\rangle_{\alpha}}{B_0}+v_{\parallel}\frac{\langle\delta\mathbf{B}_{\perp}\rangle_{\alpha}}{B_0}\right)\cdot\left(\frac{\mu}{v_{\parallel}}\nabla B_0\frac{\partial F_0}{\partial v_{\parallel}}\right)\\&-\frac{q}{m}\left[-\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_{\alpha}}{\partial t}-\left(v_{\parallel}\mathbf{e}_{\parallel}+\frac{v_{\parallel}^2}{\Omega}\nabla\times\mathbf{b}+\frac{\mu}{\Omega B_0}\mathbf{B}_0\times\nabla B_0+\frac{1}{B_0^2}\mathbf{E}_0\times\mathbf{B}_0+v_{\parallel}\frac{\langle\delta\mathbf{B}_{\perp}\rangle_{\alpha}}{B_0}\right)\cdot\right.\\&\left.\nabla_X\langle\delta\phi\rangle_{\alpha}\right]\frac{\partial F_0}{\partial v_{\parallel}}\frac{1}{v_{\parallel}},\end{aligned} \tag{387}$$

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

$$\begin{aligned}
&\left[\frac{\partial}{\partial t}+(v_\parallel \mathbf{e}_\parallel+\mathbf{V}_D+\delta\mathbf{V}_D)\cdot\frac{\partial}{\partial \mathbf{X}'}\right]\delta f-\mathbf{e}_\parallel\cdot\mu\nabla B_0\frac{\partial \delta f}{\partial v_\parallel}\\
=&-\delta\mathbf{V}_D\cdot\left(\frac{\partial F_0}{\partial \mathbf{X}}\right)\\
&+\frac{q}{m}\left[\frac{m}{q}v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\cdot(\mu\nabla B_0)+\frac{\partial\langle\mathbf{v}\cdot\delta\mathbf{A}\rangle_\alpha}{\partial t}+\left(v_\parallel\mathbf{b}+\frac{v_\parallel^2}{\Omega}\nabla\times\mathbf{b}+\frac{1}{B_0^2}\mathbf{E}_0\times\mathbf{B}_0+\right.\right.\\
&\left.\left.v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0}\right)\cdot\nabla_X\langle\delta\phi\rangle_\alpha\right]\frac{\partial F_0}{\partial v_\parallel}\frac{1}{v_\parallel},
\end{aligned}\tag{388}$$

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.

H 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 (344) is written as

$$\begin{aligned}
&\left[\frac{\partial}{\partial t}+\left(v_\parallel\mathbf{e}_\parallel+\mathbf{v}_D+\mathbf{v}_E+v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X\right]\delta f\\
=&-\left(\mathbf{v}_E+v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0-\frac{q}{m}\left[\left(v_\parallel\mathbf{e}_\parallel+\mathbf{v}_D+v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\delta\mathbf{E}\right]\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned}\tag{389}$$

H.1 Linear case

Neglecting the nonlinear terms, drift-kinetic equation (389) is written

$$\begin{aligned}
&\left[\frac{\partial}{\partial t}+(v_\parallel\mathbf{e}_\parallel+\mathbf{v}_D)\cdot\nabla_X\right]\delta f\\
=&-\left(\mathbf{v}_E+v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0-\frac{q}{m}[(v_\parallel\mathbf{e}_\parallel+\mathbf{v}_D)\cdot\delta\mathbf{E}]\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned}\tag{390}$$

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 (390) by $qv_\parallel$ and then integrating over velocity space, we obtain

$$\begin{aligned}
\frac{\partial\delta j_\parallel}{\partial t} &= -q\int d\mathbf{v}v_\parallel(v_\parallel\mathbf{e}_\parallel+\mathbf{v}_D)\cdot\nabla_X\delta f\\
&- q\int d\mathbf{v}v_\parallel\left(\mathbf{v}_E+v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0-\frac{q}{m}q\int d\mathbf{v}v_\parallel[(v_\parallel\mathbf{e}_\parallel+\mathbf{v}_D)\cdot\delta\mathbf{E}]\frac{\partial F_0}{\partial\varepsilon}.
\end{aligned}\tag{391}$$

Equation (391) 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 (390) into variable $(\mathbf{X}, \mu, v_\parallel)$.

H.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

$$\mathbf{X}'(\mathbf{X}, \mu, \varepsilon) = \mathbf{X}, \tag{392}$$

$$\mu'(\mathbf{X}, \mu, \varepsilon) = \mu, \tag{393}$$

and

$$v_\parallel(\mathbf{X}, \mu, \varepsilon) = \sqrt{2(\varepsilon - \mu B_0(\mathbf{X}))}. \tag{394}$$

Use this, we have

$$\begin{aligned} \frac{\partial \delta G_0}{\partial \mathbf{X}}|_{\mu,\varepsilon} &= \frac{\partial \mathbf{X}'}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial \mathbf{X}'} + \frac{\partial \mu'}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial \mu'} + \frac{\partial v_\parallel}{\partial \mathbf{X}} \frac{\partial \delta G_0}{\partial v_\parallel} \\ &= \frac{\partial \delta G_0}{\partial \mathbf{X}'}|_{\mu, v_\parallel} + 0 \frac{\partial \delta G_0}{\partial \mu'} - \frac{\mu}{v_\parallel} \frac{\partial B_0}{d\mathbf{X}} \frac{\partial \delta G_0}{\partial v_\parallel}, \end{aligned} \tag{395}$$

and

$$\begin{aligned} \frac{\partial F_0}{\partial \varepsilon} &= \frac{\partial F_0}{\partial \mu'} \frac{\partial \mu'}{\partial \varepsilon} + \frac{\partial F_0}{\partial v_\parallel} \frac{\partial v_\parallel}{\partial \varepsilon} \\ &= 0 \frac{\partial F_0}{\partial \mu'} + \frac{\partial F_0}{\partial v_\parallel} \frac{\partial v_\parallel}{\partial \varepsilon} \\ &= \frac{\partial F_0}{\partial v_\parallel} \frac{1}{v_\parallel} \end{aligned} \tag{396}$$

Then, in terms of variable $(\mathbf{X}', \mu, v_\parallel)$, equation (390) is written

$$\begin{aligned} &\frac{\partial \delta f}{\partial t} + (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D) \cdot \nabla \delta f - \mathbf{e}_\parallel \cdot \mu \nabla B \frac{\partial \delta f}{\partial v_\parallel} \\ =& -\left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0} \right) \cdot \nabla F_0 + \left(\frac{\mathbf{v}_E}{v_\parallel} + \frac{\delta \mathbf{B}_\perp}{B_0} \right) \cdot \mu \nabla B \frac{\partial F_0}{\partial v_\parallel} - \frac{q}{m} \left[\left(\mathbf{e}_\parallel + \frac{\mathbf{v}_D}{v_\parallel} \right) \cdot \delta \mathbf{E} \right] \frac{\partial F_0}{\partial v_\parallel}, \end{aligned} \tag{397}$$

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

H.3 Parallel momentum equation

Multiplying the linear drift kinetic equation (397) by $qv_\parallel$ and then integrating over velocity space, we obtain

$$\begin{aligned}
&\frac{\partial \delta j_\parallel}{\partial t} + q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D)\cdot\nabla_X \delta f - q\int d\mathbf{v} v_\parallel \mathbf{e}_\parallel \cdot \mu\nabla B \frac{\partial \delta f}{\partial v_\parallel} \\
=&-q\int d\mathbf{v} v_\parallel \left(\mathbf{v}_E + v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0 + q\int d\mathbf{v} v_\parallel \left(\frac{\mathbf{v}_E}{v_\parallel} + \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot \mu\nabla B \frac{\partial F_0}{\partial v_\parallel} \\
&-\frac{q}{m} q\int d\mathbf{v} v_\parallel \left[\left(\mathbf{e}_\parallel + \frac{\mathbf{v}_D}{v_\parallel}\right)\cdot \delta\mathbf{E}\right]\frac{\partial F_0}{\partial v_\parallel}.
\end{aligned} \tag{398}$$

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. (398) are all zero. Among the rest terms, only the following term

$$-\frac{q}{m} q\int d\mathbf{v} v_\parallel [(v_\parallel \mathbf{e}_\parallel)\cdot \delta\mathbf{E}]\frac{\partial F_0}{\partial v_\parallel}\frac{1}{v_\parallel} \tag{399}$$

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

$$\begin{aligned}
&-\frac{q}{m} q\int d\mathbf{v} v_\parallel [(v_\parallel \mathbf{e}_\parallel)\cdot \delta\mathbf{E}]\frac{\partial F_0}{\partial v_\parallel}\frac{1}{v_\parallel} \\
=&-\frac{q^2}{m}\int 2\pi B dv_\parallel d\mu v_\parallel \delta E_\parallel \frac{\partial F_0}{\partial v_\parallel} \\
=&-\frac{q^2}{m}\int 2\pi B d\mu \delta E_\parallel \int v_\parallel \frac{\partial F_0}{\partial v_\parallel} dv_\parallel \\
=&-\frac{q^2}{m}\int 2\pi B d\mu \delta E_\parallel \left(0 - \int F_0 dv_\parallel\right) \\
=&\frac{q^2}{m}\delta E_\parallel n_0.
\end{aligned} \tag{400}$$

Using these results, the parallel momentum equation (398) is written

$$\begin{aligned}
\frac{\partial \delta j_\parallel}{\partial t} \;=\; & \frac{q^2}{m}\delta E_\parallel n_0 - q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D)\cdot\nabla_X \delta f + q\int d\mathbf{v} v_\parallel \mathbf{e}_\parallel \cdot \mu\nabla B \frac{\partial \delta f}{\partial v_\parallel} \\
& -q\int d\mathbf{v} v_\parallel \left(v_\parallel \frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot\nabla F_0 + q\int d\mathbf{v} v_\parallel \left(\frac{\delta \mathbf{B}_\perp}{B_0}\right)\cdot \mu\nabla B \frac{\partial F_0}{\partial v_\parallel},
\end{aligned} \tag{401}$$

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 (401) 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

$$\begin{aligned}\frac{\partial \delta j_\parallel}{\partial t} &= \frac{q^2}{m}\delta E_\parallel n_0 - q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D)\cdot\nabla_X \delta f - q(\mathbf{e}_\parallel\cdot\nabla B_0)\int \mu\delta f d\mathbf{v} \\ &\quad -q\int d\mathbf{v} v_\parallel\left(v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla F_0 - q\left(\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot(\nabla B_0)\int \mu F_0\, d\mathbf{v}, \end{aligned} \tag{402}$$

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

$$\begin{aligned}\frac{\partial \delta j_\parallel}{\partial t} &= \frac{q^2}{m}\delta E_\parallel n_0 - q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel + \mathbf{v}_D)\cdot\nabla_X \delta f - q(\mathbf{e}_\parallel\cdot\nabla B_0)\frac{\delta p_\perp}{m B_0} \\ &\quad -q\int d\mathbf{v} v_\parallel\left(v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla F_0 - q\left(\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot(\nabla B_0)\frac{p_{\perp 0}}{m B_0}, \end{aligned} \tag{403}$$

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. (403) is written

$$\begin{aligned}&-q\int d\mathbf{v} v_\parallel (v_\parallel \mathbf{e}_\parallel)\cdot\nabla_X\delta f, \\ =&-q\int 2\pi B_0\, dv_\parallel d\mu v_\parallel^2 \mathbf{e}_\parallel\cdot\nabla_X\delta f \\ =&-q2\pi B_0\,\mathbf{e}_\parallel\cdot\nabla_X\int v_\parallel^2\delta f dv_\parallel d\mu \\ =&-qB_0\,\mathbf{e}_\parallel\cdot\nabla_X\left(\frac{1}{mB_0}\int m v_\parallel^2\delta f d\mathbf{v}\right) \\ =&-qB_0\,\mathbf{e}_\parallel\cdot\nabla_X\left(\frac{\delta p_\parallel}{mB_0}\right), \end{aligned} \tag{404}$$

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

$$\begin{aligned}&-q\int d\mathbf{v} v_\parallel\left(v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0 \\ =&-q\int 2\pi B_0\, dv_\parallel d\mu\left(v_\parallel^2\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot\nabla_X F_0 \\ =&-q\left(\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot B_0\nabla_X\int (v_\parallel^2 F_0 2\pi dv_\parallel d\mu) \\ =&-q\left(\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot B_0\,\nabla_X\left[\frac{1}{B}\int (m v_\parallel^2 F_0 d\mathbf{v})\right] \\ =&-q\delta\mathbf{B}_\perp\cdot\nabla_X\left(\frac{p_{\parallel 0}}{mB_0}\right) \end{aligned} \tag{405}$$

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

$$\mathbf{v}_D = \frac{B_0 \frac{v_\parallel}{\Omega} \nabla \times \mathbf{b}}{B_\parallel^\star} v_\parallel + \frac{\mu}{\Omega B_\parallel^\star} \mathbf{B}_0 \times \nabla B_0. \tag{406}$$

where $B_\parallel^\star = B_0\big(1 + \frac{v_\parallel}{\Omega}\mathbf{b}\cdot\nabla\times\mathbf{b}\big)$ (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

$$\begin{aligned}
-q\int d\mathbf{v} v_\parallel \mathbf{v}_D \cdot \nabla_X \delta f &= -q\int 2\pi B_0\, dv_\parallel d\mu v_\parallel \left(\frac{v_\parallel^2}{\Omega}\nabla\times\mathbf{b} + \frac{\mu}{\Omega}\mathbf{b}\times\nabla B_0\right)\cdot\nabla_X\delta f \\
&= -q2\pi B_0 \frac{1}{\Omega}(\nabla\times\mathbf{b})\cdot\nabla_X \int v_\parallel^3 \delta f dv_\parallel d\mu - q2\pi B_0 \frac{1}{\Omega}(\mathbf{b}\times\nabla B_0)\cdot\nabla_X \int \\
&\quad v_\parallel \mu \delta f dv_\parallel d\mu \\
&= -q B_0\ \frac{1}{\Omega}(\nabla\ \times\ \mathbf{b}\ )\ \cdot\ \nabla_X\left(\frac{1}{B_0}\int v_\parallel^3 \delta f d\mathbf{v}\ \right)\ -\ q B_0 \frac{1}{\Omega}(\mathbf{b}\ \times \\
&\quad \nabla B_0)\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel \mu \delta f d\mathbf{v}\right), \\
&= -m(\nabla\times\mathbf{b}\,)\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel^3\delta f d\mathbf{v}\ \right) - m(\mathbf{b}\times\nabla B_0)\cdot\nabla_X\left(\frac{1}{B_0}\int\right. \\
&\quad \left. v_\parallel\mu\delta f d\mathbf{v}\right),
\end{aligned} \tag{407}$$

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

$$\begin{aligned}
\frac{\partial \delta j_\parallel}{\partial t} &= \frac{e^2 n_{e0}}{m}\delta E_\parallel + e B_0\, \mathbf{b}\cdot\nabla_X\left(\frac{\delta p_\parallel}{m B_0}\right) + e(\mathbf{b}\cdot\nabla B_0)\frac{\delta p_\perp}{m B_0} \\
&\quad + e\delta\mathbf{B}_\perp\cdot\nabla_X\left(\frac{p_{\parallel 0}}{m B_0}\right) + e\left(\frac{\delta\mathbf{B}_\perp}{B_0}\right)\cdot(\nabla B_0)\frac{p_{\perp 0}}{m B_0}. \\
&\quad - m(\nabla\times\mathbf{b}\,)\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel^3\delta f d\mathbf{v}\ \right) - m(\mathbf{b}\times\nabla B_0)\cdot\nabla_X\left(\frac{1}{B_0}\int v_\parallel\mu\delta f d\mathbf{v}\right)
\end{aligned} \tag{408}$$

Define

$$\mathbf{D}_0 = \nabla\left(\frac{p_{\parallel 0}}{m B_0}\right) + \frac{\nabla B_0}{B_0}\frac{p_{\perp 0}}{m B_0}, \tag{409}$$

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

$$\mathbf{D}_0 = \frac{\nabla p_0}{m B_0}. \tag{410}$$

then Eq. (408) is written as

$$\begin{aligned}\frac{\partial \delta j_{\parallel}}{\partial t} &= \frac{e^2 n_0}{m}\delta E_{\parallel} + e\delta \mathbf{B}_{\perp}\cdot \mathbf{D}_0 \\ &+ e B_0 \mathbf{b}\cdot\nabla_X\left(\frac{\delta p_{\parallel}}{m B_0}\right) + e(\mathbf{b}\cdot\nabla B_0)\frac{\delta p_{\perp}}{m B_0} \\ &- m(\nabla\times\mathbf{b})\cdot\nabla_X\left(\frac{1}{B_0}\int v_{\parallel}^3 \delta f d\mathbf{v}\right) - m(\mathbf{b}\times\nabla B_0)\cdot\nabla_X\left(\frac{1}{B_0}\int v_{\parallel}\mu\delta f d\mathbf{v}\right). \end{aligned} \tag{411}$$

H.4 Special case in uniform magnetic field

In the case of uniform magnetic field, the parallel momentum equation (408) is written as

$$\frac{\partial \delta j_{\parallel}}{\partial t} = \frac{q}{m} q E_{\parallel} n_{e0} - q\,\mathbf{e}_{\parallel}\cdot\nabla_X(\delta p_{\parallel}) - q\frac{\delta \mathbf{B}_{\perp}}{B_0}\cdot\nabla_X p_{\parallel 0}. \tag{412}$$

H.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. F or Appendix in Yang Chen's paper[1])

$$n_{e0}\delta\mathbf{V}_{e\perp} = \underbrace{\frac{n_{e0}}{B_0}\delta\mathbf{E}\times\mathbf{b}}_{E\times B\text{flow}} - \underbrace{\frac{1}{e B_0}\mathbf{b}\times\nabla\delta p_{\perp e}}_{\text{diamagneticflow}} \tag{413}$$

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

H.5.1 Drift kinetic equation

Drift kinetic equation is written

$$\frac{\partial f}{\partial t} + \left(v_{\parallel}\tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta\mathbf{E}\times\mathbf{b}}{B_0}\right)\cdot\nabla f + \left(-\frac{e}{m}\delta E_{\parallel} - \mu\tilde{\mathbf{b}}\cdot\nabla B\right)\frac{\partial f}{\partial v_{\parallel}} = 0, \tag{414}$$

where $f = f(\mathbf{x},\, \mu,\, v_{\parallel},\, t)$, $\mu = m v_{\perp}^2 / B_0$ 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.

H.5.2 Parallel momentum equation

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

$$\int \frac{\partial f_e}{\partial t} v_\parallel d\mathbf{v} + \int v_\parallel \Big(v_\parallel \tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta \mathbf{E} \times \mathbf{b}}{B_0} \Big) \cdot \nabla f_e d\mathbf{v} + \int v_\parallel \Big(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \Big) \frac{\partial f_e}{\partial v_\parallel} d\mathbf{v} = 0, \tag{415}$$

which can be written as

$$\frac{\partial J_{\parallel e}}{\partial t} + \int v_\parallel \Big(v_\parallel \tilde{\mathbf{b}} + \mathbf{v}_D + \frac{\delta \mathbf{E} \times \mathbf{e}_\parallel}{B_0} \Big) \cdot \nabla f_e \, d\mathbf{v} + \int v_\parallel \Big(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \Big) \frac{\partial f_e}{\partial v_\parallel} d\mathbf{v} = 0, \tag{416}$$

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

$$\begin{aligned}
&\iint v_\parallel \Big(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \Big) \frac{\partial f_e}{\partial v_\parallel} d\mathbf{v} \\
&= \iint v_\parallel \Big(-\frac{e}{m} \delta E_\parallel - \mu \tilde{\mathbf{b}} \cdot \nabla B \Big) \frac{\partial f_e}{\partial v_\parallel} 2\pi \frac{B}{m} \, dv_\parallel d\mu \\
&= -\frac{e}{m} \delta E_\parallel 2\pi \frac{B}{m} \iint v_\parallel \frac{\partial f_e}{\partial v_\parallel} \, dv_\parallel d\mu - (\tilde{\mathbf{b}} \cdot \nabla B) 2\pi \frac{B}{m} \int \mu \int v_\parallel \frac{\partial f_e}{\partial v_\parallel} \, dv_\parallel d\mu \\
&= -\frac{e}{m} \delta E_\parallel 2\pi \frac{B}{m} \int \Big(v_\parallel f_e |_{-\infty}^{+\infty} - \int f_e dv_\parallel \Big) d\mu - (\tilde{\mathbf{b}} \cdot \nabla B) 2\pi \frac{B}{m} \int \mu \Big(v_\parallel f_e |_{-\infty}^{+\infty} - \int f_e dv_\parallel \Big) d\mu \\
&= \frac{e}{m} \delta E_\parallel 2\pi \frac{B}{m} \iint f_e dv_\parallel d\mu + (\tilde{\mathbf{b}} \cdot \nabla B) 2\pi \frac{B}{m} \int \mu \int f_e dv_\parallel d\mu \\
&= \frac{e}{m} \delta E_\parallel n_e + \iint \mu (\tilde{\mathbf{b}} \cdot \nabla B) f_e d\mathbf{v} \\
&\approx \frac{e}{m} \delta E_\parallel n_e
\end{aligned} \tag{417}$$

$$\begin{aligned}
&\iint v_\parallel (v_\parallel \tilde{\mathbf{b}}) \cdot \nabla f_e d\mathbf{v} \\
&= \iint \tilde{\mathbf{b}} \cdot \nabla (v_\parallel^2 f_e) d\mathbf{v} \\
&= \iint \tilde{\mathbf{b}} \cdot \nabla (v_\parallel^2 f_e) 2\pi \frac{B}{m} \, dv_\parallel d\mu \\
&= \tilde{\mathbf{b}} \cdot \nabla \Big(\iint v_\parallel^2 f_e 2\pi \frac{1}{m} \, dv_\parallel d\mu \Big) B_0 \\
&= \tilde{\mathbf{b}} \cdot \nabla \Big(\frac{p_\parallel}{B_0} \Big) B_0
\end{aligned}$$

$$
\begin{aligned}
=&\tilde{\mathbf{b}}\cdot\nabla\left(\frac{p_{\|0}+\delta p_{\|}}{B_0}\right)B_0\\
=&\tilde{\mathbf{b}}\cdot\nabla\left(\frac{p_{\|0}}{B_0}\right)B_0+\tilde{\mathbf{b}}\cdot\nabla\left(\frac{\delta p_{\|}}{B_0}\right)B_0\\
\approx&\mathbf{b}\cdot\nabla\left(\frac{p_{\|0}}{B_0}\right)B_0+\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla\left(\frac{p_{\|0}}{B_0}\right)B_0+\mathbf{b}\cdot\nabla\left(\frac{\delta p_{\|}}{B_0}\right)B_0\\
\approx&\mathbf{b}\cdot\nabla(p_{\|0})+\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla(p_{\|0})+\mathbf{b}\cdot\nabla(\delta p_{\|})\\
=&\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla(p_{\|0})+\mathbf{b}\cdot\nabla(\delta p_{\|})
\end{aligned}
\tag{418}
$$

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

$$
\frac{\partial\delta J_{e\|}}{\partial t}=-\frac{e}{m}\delta E_\|n_e-\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla(p_{\|0})-\mathbf{b}\cdot\nabla(\delta p_{\|})
\tag{419}
$$

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

$$
\mu_0e\,\frac{e}{m}\delta E_\|n_e+\mathbf{b}\cdot\nabla\times\nabla\times\delta\mathbf{E}=-\mu_0e\left[\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla(p_{e\,\|0})+\mathbf{b}\cdot\nabla(\delta p_{e\,\|})\right]
\tag{420}
$$

$$
-\mathbf{b}\cdot\nabla\times\nabla\times\delta\mathbf{E}=-\mu_0e\left(-q\frac{\delta\mathbf{B}_\perp}{B_0}\cdot\nabla_Xp_{\|0}+\frac{q}{m}qE_\|n_{e0}-q\,\mathbf{e}_\|\cdot\nabla_X(\delta p_\|)\right).
\tag{421}
$$

ddddd

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

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

I Derivation of Eq. (122), not finished

From the definition of μ, we obtain

$$
\frac{\partial \mu}{\partial \mathbf{x}} = -\frac{v_{\perp}^2}{2B_0^2} \frac{\partial B_0}{\partial \mathbf{x}} + \frac{1}{2B_0} \frac{\partial v_{\perp}^2}{\partial \mathbf{x}} = -\frac{\mu}{B_0} \frac{\partial B_0}{\partial \mathbf{x}} + \frac{\partial \mathbf{v}_{\perp}}{\partial \mathbf{x}} \cdot \frac{\mathbf{v}_{\perp}}{B_0} \tag{422}
$$

Using

$$
\frac{\partial \mathbf{v}_{\perp}}{\partial \mathbf{x}} = \frac{\partial [\mathbf{v} - v_{\parallel} \mathbf{e}_{\parallel}]}{\partial \mathbf{x}} = -v_{\parallel} \frac{\partial \mathbf{e}_{\parallel}}{\partial \mathbf{x}} - \frac{\partial v_{\parallel}}{\partial \mathbf{x}} \mathbf{e}_{\parallel}, \tag{423}
$$

expression (422) is written as

$$
\frac{\partial \mu}{\partial \mathbf{x}} = -\frac{\mu}{B_0} \frac{\partial B_0}{\partial \mathbf{x}} - v_{\parallel} \frac{\partial \mathbf{e}_{\parallel}}{\partial \mathbf{x}} \cdot \frac{\mathbf{v}_{\perp}}{B_0}, \tag{424}
$$

which agrees with Eq. (10) in Frieman-Chen's paper[2].

$$
\begin{aligned}
&\frac{1}{2\pi} \int_0^{2\pi} d\alpha \mathbf{v} \cdot \frac{\partial \mu}{\partial \mathbf{x}} \\
&= \frac{1}{2\pi} \int_0^{2\pi} d\alpha \mathbf{v} \cdot \left(-\frac{\mu}{B_0} \frac{\partial B_0}{\partial \mathbf{x}} - v_{\parallel} \frac{\partial \mathbf{e}_{\parallel}}{\partial \mathbf{x}} \cdot \frac{\mathbf{v}_{\perp}}{B_0} \right) \\
&= ?\, 0
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{2\pi} \int_0^{2\pi} d\alpha \mathbf{v} \cdot \left[\mathbf{v} \times \frac{\partial}{\partial \mathbf{x}} \Big(\frac{\boldsymbol{e}_{\parallel}}{\Omega} \Big) \right] \cdot \frac{\partial \delta G_0}{\partial \mathbf{X}} \\
&=
\end{aligned}
$$

Using the fact that δG_0 is independent of α, the left-hand side of Eq. (122) is written as

$$\begin{aligned}
& \langle \mathbf{v}\cdot[\boldsymbol{\lambda}_{B1}+\boldsymbol{\lambda}_{B2}]\delta G_0\rangle_\alpha \\
= & \frac{1}{2\pi}\int_0^{2\pi} d\alpha\, \mathbf{v}\cdot\left\{\left[\mathbf{v}\times\frac{\partial}{\partial\mathbf{x}}\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\right]\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}+\frac{\partial\mu}{\partial\mathbf{x}}\frac{\partial\delta G_0}{\partial\mu}+\frac{\partial\alpha}{\partial\mathbf{x}}\frac{\partial\delta G_0}{\partial\alpha}\right\} \\
= & \frac{1}{2\pi}\int_0^{2\pi} d\alpha\, \mathbf{v}\cdot\left\{\left[\mathbf{v}\times\frac{\partial}{\partial\mathbf{x}}\left(\frac{\boldsymbol{e}_\parallel}{\Omega}\right)\right]\cdot\frac{\partial\delta G_0}{\partial\mathbf{X}}+\frac{\partial\mu}{\partial\mathbf{x}}\frac{\partial\delta G_0}{\partial\mu}\right\}
\end{aligned}$$

J Modern gyrokinetic formulation

Frimain-Chen equation was obtained by first transforming to the guiding-center coordinates and then using the classical perturbation expansion for the distribution function (to separate gyro-phase independent part form the gyro-phase dependent one). Note that the guiding-center transform does not involve the perturbed field.

The modern form of the nonlinear gyrokinetic equation is in the total-f form. This way of deriving the gyrokinetic equation is to use coordinate transforms to eliminate gyro-phase dependence of the total distribution function (rather than splitting the distribution function itself) and thus obtain an equation for the resulting gyro-phase independent distribution function (called gyro-center distribution function). The coordinate transform involves the perturbed fields besides the equilibrium field.

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

$$\left(\frac{\partial}{\partial t}+\dot{\mathbf{X}}\cdot\nabla+\dot{v}_\parallel\frac{\partial}{\partial v_\parallel}\right)f(\mathbf{X},v_\parallel,\mu,t)=0, \tag{425}$$

where

$$\dot{\mathbf{R}}=\mathbf{V}_D+\frac{\mathbf{e}_\parallel}{B_0}\times\langle\nabla\delta\Phi\rangle_\alpha+v_\parallel\frac{\langle\delta\mathbf{B}_\perp\rangle_\alpha}{B_0} \tag{426}$$

$$\dot{v}_\parallel=-\frac{1}{m}\frac{\mathbf{B}^\star}{B_\parallel^\star}\cdot(q\nabla\langle\delta\Phi\rangle+\mu\nabla B_0)-\frac{q}{m}\frac{\partial\langle\delta A_\parallel\rangle_\alpha}{\partial t}. \tag{427}$$

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 (inverse gyro-center transform) gives rise to the polarization density term. (phase-space-Lagrangian Lie perturbation method (Littlejohn, 1982a, 1983), I need to read these two papers.).

The learning curve of Lie tansform gyrokinetic theory seems steep (the name Lie transform itself is already very scaring). I tried to learn it but failed. To derive the Frieman-Chen equation, you just need some patience. Meanwhile, patience is all that you need to learn any mathematic stuff, but some requires more patience than others, and Lie transform gyrokinetic is one that requires more for me.

8 About this document

These notes were initially written when I visited University of Colorado at Boulder (Sept.-Nov. 2016), where I worked with Dr. Yang Chen, who said that most gyrokinetic simulations essentially employ Frieman-Chen's nonlinear gyrokinetic equation. Therefore a careful re-derivation of the equation to get familiar to the gyrokinetic orderings and physics included in the model is highly helpful, which motivates me to write this note. Update: After the visit, I went there again, working with Yang Chen and Scott Parker as a postdoc researcher for 2 years, during which I tried to develop a fully kinetic code (including gyromotion of ions) for electromagnetic ITG turbulence simulation. I got an electrostatic version (with adiabatic electrons) work, but failed to deliver a working electromagnetic version.

This document was written by using TeXmacs[5], a structured wysiwyw (what you see is what you want) editor for scientists. The HTML version of this document is generated by first converting the TeXmacs file to TeX format and then using `htlatex` to convert the TeX file to HTML format.

Bibliography

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

[2] 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.

[3] A. Mishchenko, A. Bottino, A. Biancalani, R. Hatzky, T. Hayward-Schneider, N. Ohana, E. Lanti, S. Brunner, L. Villard, M. Borchardt, R. Kleiber, and A. Könies. Pullback scheme implementation in orb5. *Computer Physics Communications*, 238:194–202, 2019.

[4] Benjamin J. Sturdevant, S. Ku, L. Chacón, Y. Chen, D. Hatch, M. D. J. Cole, A. Y. Sharma, M. F. Adams, C. S. Chang, S. E. Parker, and R. Hager. Verification of a fully implicit particle-in-cell method for the v ‖-formalism of electromagnetic gyrokinetics in the xgc code. *Physics of Plasmas*, 28(7):72505, jul 2021.

[5] Joris van der Hoeven. Gnu texmacs, a structured wysiwyw (what you see is what you want) editor for scientists. `http://www.texmacs.org/`, 2007. [Online].