Linearized collisionless drift kinetic equation

BY YOUJUN HU

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

Abstract

The notes review the Lagrangian mechanics of guiding center motion and some aspects of the linear gyrokinetic theory. These notes were initially written when I read Porcelli's paper[1] and were later revised to include more contents.

1 Guiding-center motion

The phase-space Lagrangian for guiding-center motion was first given in Littlejohn's paper[1], which takes the following form

$$\mathcal{L}(\mathbf{X}, v_{\parallel}, y, \alpha, \dot{\mathbf{X}}, \dot{v}_{\parallel}, \dot{y}, \dot{\alpha}, t) = \left(\frac{Ze}{c}\mathbf{A} + m v_{\parallel}\mathbf{b}\right) \cdot \dot{\mathbf{X}} + \frac{1}{\Omega} y \dot{\alpha} - \frac{1}{2} m v_{\parallel}^2 - y - Ze\phi, \tag{1}$$

where $\mathbf{X}$ is the location of the guiding-center, $v_{\parallel}$ is the parallel (to magnetic field) velocity of the particle (will be proved later that $v_{\parallel}$ is also the parallel velocity of the guiding center) $y \equiv m v_{\perp}^2 / 2$ with $v_{\perp}$ the perpendicular (to magnetic field) velocity of the particle, α is the gyrophase, $\Omega = ZeB/(mc)$ with Z being the charge number and e being the elementary charge, $\mathbf{b} = \mathbf{B}/B$. Note that here the (phase-space) Lagrangian of guiding center is considered to be a function of variables $\mathbf{X}$, $v_{\parallel}$, y, α, $\dot{\mathbf{X}}$, $\dot{v}_{\parallel}$, $\dot{y}$, $\dot{\alpha}$, and t. Also note that three variables, α, $\dot{v}_{\parallel}$ and $\dot{y}$ happens not to appear in Eq. (1). Further note that the explicit dependence of $\mathcal{L}$ on $\mathbf{X}$ and t is through the electromagnetic field $\mathbf{A}$, ϕ, and the cyclotron frequency Ω. The Euler-Lagrange equation corresponding to variable y is written as

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}}\right) = \frac{\partial \mathcal{L}}{\partial y}, \tag{2}$$

which, after evaluating the partial derivatives, is reduced to

$$\dot{\alpha} = \Omega, \tag{3}$$

which indicates, as expected, that the time change rate of the gyrophase α equals the cylcotron frequency Ω. The Euler-Lagrange equation corresponding to the variable α is written as

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\alpha}}\right) = \frac{\partial \mathcal{L}}{\partial \alpha}, \tag{4}$$

which can be written as

$$\frac{d}{dt}\left(\frac{y}{\Omega}\right) = 0, \tag{5}$$

which indicats that the magnetic moment $\mu \equiv y / B$ is a constant of the motion. The Euler-Lagrange equation for the variable $v_{\parallel}$ is

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{v}_{\parallel}}\right) = \frac{\partial \mathcal{L}}{\partial v_{\parallel}}, \tag{6}$$

which can be simplified to

$$v_{\parallel} = \mathbf{b} \cdot \dot{\mathbf{X}}, \tag{7}$$

which indicates that $v_{\parallel}$ is also the parallel velocity of the guiding center. Next, consider the Euler-Lagrange equation corresponding to the coordinate $\mathbf{X}$, which is given by

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{X}}}\right) = \frac{\partial \mathcal{L}}{\partial \mathbf{X}}, \tag{8}$$

which should be understood as a shorthand of the three Euler-Lagrange equations coorresponding to three coordinates. Is the above equation still valid in arbitrary coordinates system if we consider $\partial/\partial\dot{\mathbf{X}}$ and $\partial/\partial\mathbf{X}$ as gradient operators? The answer is yes. However it is not trivial for me to find the proof (the proof is provided in Sec. (4)). Using Eq. (1) and vector identies, we obtain

$$\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}}=\frac{Ze}{c}\mathbf{A}+mv_{\|}\mathbf{b}, \tag{9}$$

and

$$\frac{\partial\mathcal{L}}{\partial\mathbf{X}}=\dot{\mathbf{X}}\times\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)+\dot{\mathbf{X}}\cdot\nabla\left(\frac{Ze}{c}\mathbf{A}+mv_{\|}\mathbf{b}\right)-\frac{1}{\Omega^2}y\dot{\alpha}\nabla\Omega-Ze\nabla\phi, \tag{10}$$

where $\nabla\equiv\partial/\partial\mathbf{X}$. Using Eqs. (9) and (10), Eq. (8) is written as

$$\frac{Ze}{c}\left(\frac{\partial\mathbf{A}}{\partial t}+\dot{\mathbf{X}}\cdot\nabla\mathbf{A}\right)+mv_{\|}\left(\frac{\partial\mathbf{b}}{\partial t}+\dot{\mathbf{X}}\cdot\nabla\mathbf{b}\right)+m\dot{v}_{\|}\mathbf{b}=\dot{\mathbf{X}}\times\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)+\dot{\mathbf{X}}\cdot\nabla\left(\frac{Ze}{c}\mathbf{A}+mv_{\|}\mathbf{b}\right)-\frac{1}{\Omega^2}y\dot{\alpha}\nabla\Omega-Ze\nabla\phi \tag{11}$$

Using $\Omega=BZe/(mc)$ and $\dot{\alpha}=\Omega$, the second last term can be reduced to $-\mu\nabla B$. Then Eq. (11) is written

$$\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t}+mv_{\|}\frac{\partial\mathbf{b}}{\partial t}+\dot{\mathbf{X}}\cdot\left(\frac{Ze}{c}\nabla\mathbf{A}+mv_{\|}\nabla\mathbf{b}\right)+m\dot{v}_{\|}\mathbf{b}=\dot{\mathbf{X}}\times\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)+\dot{\mathbf{X}}\cdot\nabla\left(\frac{Ze}{c}\mathbf{A}+mv_{\|}\mathbf{b}\right)-\mu\nabla B-Ze\nabla\phi \tag{12}$$

Noting that $mv_{\|}\nabla\mathbf{b}=\nabla(mv_{\|}\mathbf{b})$ (this is because $\nabla\equiv(\partial/\partial\mathbf{X})_{v_{\|}}$, i.e., holding $v_{\|}$ constant), so that the second term on the right-hand side of the above equation is canceled by terms on the right-hand, yielding

$$\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t}+mv_{\|}\frac{\partial\mathbf{b}}{\partial t}+m\dot{v}_{\|}\mathbf{b}=\dot{\mathbf{X}}\times\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)-\mu\nabla B-Ze\nabla\phi. \tag{13}$$

Equation (13) can be further written in compact form by defining new magnetic-like and electric-like quantities. Define

$$\mathbf{A}^{\star}=\mathbf{A}+\frac{mc}{Ze}v_{\|}\mathbf{b}, \tag{14}$$

and

$$\mathbf{B}^{\star}=\nabla\times\mathbf{A}^{\star}, \tag{15}$$

then

$$\frac{\partial\mathbf{A}^{\star}}{\partial t}=\frac{\partial\mathbf{A}}{\partial t}+\frac{mc}{Ze}v_{\|}\frac{\partial\mathbf{b}}{\partial t} \tag{16}$$

$$\mathbf{B}^{\star}=\mathbf{B}+\frac{mc}{Ze}v_{\|}\nabla\times\mathbf{b}, \tag{17}$$

(Note that the time partial differential does not operate on $v_{\|}$ because it is an independent variables.) Using these, Eq. (13) is written as

$$\frac{Ze}{c}\frac{\partial\mathbf{A}^{\star}}{\partial t}+m\dot{v}_{\|}\mathbf{b}=\frac{Ze}{c}\dot{\mathbf{X}}\times\mathbf{B}^{\star}-\mu\nabla B-Ze\nabla\phi \tag{18}$$

$$\Longrightarrow\frac{Ze}{c}\dot{\mathbf{X}}\times\mathbf{B}^{\star}-Ze\nabla\phi-\frac{Ze}{c}\frac{\partial\mathbf{A}^{\star}}{\partial t}=\mu\nabla B+m\dot{v}_{\|}\mathbf{b} \tag{19}$$

Define

$$\mathbf{E}^{\star}=-\nabla\phi-\frac{1}{c}\frac{\partial\mathbf{A}^{\star}}{\partial t} \tag{20}$$

then Eq. (19) is written as

$$Ze\left(\mathbf{E}^{\star}+\frac{1}{c}\dot{\mathbf{X}}\times\mathbf{B}^{\star}\right)=\mu\nabla B+m\dot{v}_{\|}\mathbf{b}, \tag{21}$$

which agrees with Eq. (23) of Porcelli's paper[1].

1.1 Time evolution of $v_\parallel$

The time evolution of $v_\parallel$ can be obtained by dotting Eq. (13) by $\mathbf{b}$, which gives

$$\frac{Ze}{c}\frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{b}+mv_\parallel\frac{\partial \mathbf{b}}{\partial t}\cdot\mathbf{b}+m\dot{v}_\parallel=\mathbf{b}\cdot\dot{\mathbf{X}}\times(mv_\parallel\nabla\times\mathbf{b})-\mu\mathbf{b}\cdot\nabla B-Ze\mathbf{b}\cdot\nabla\phi. \tag{22}$$

Noting that

$$\mathbf{b}\cdot\frac{\partial \mathbf{b}}{\partial t}=\frac{1}{2}\frac{\partial b^2}{\partial t}=0, \tag{23}$$

Eq. (22) is written as

$$\frac{Ze}{c}\frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{b}+m\dot{v}_\parallel=\mathbf{b}\cdot\dot{\mathbf{X}}\times(mv_\parallel\nabla\times\mathbf{b})-\mu\mathbf{b}\cdot\nabla B-Ze\mathbf{b}\cdot\nabla\phi. \tag{24}$$

Using

$$\mathbf{E}=-\nabla\phi-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{25}$$

Eq. (24) is written as

$$m\dot{v}_\parallel=-\dot{\mathbf{X}}\cdot\mathbf{b}\times(mv_\parallel\nabla\times\mathbf{b})-\mu\mathbf{b}\cdot\nabla B+Ze\mathbf{b}\cdot\mathbf{E} \tag{26}$$

Noting that the magnetic curture is given by $\boldsymbol{\kappa}=-\mathbf{b}\times\nabla\times\mathbf{b}$, the above equation is written as

$$m\dot{v}_\parallel=mv_\parallel\dot{\mathbf{X}}\cdot\boldsymbol{\kappa}-\mu\mathbf{b}\cdot\nabla B+Ze\mathbf{b}\cdot\mathbf{E}, \tag{27}$$

which governing the time evolution of $v_\parallel$.

1.2 Energy conservation

Dotting Eq. (21) with $\dot{\mathbf{X}}$, we obtain

$$Ze\mathbf{E}^\star\cdot\dot{\mathbf{X}}=\mu\dot{\mathbf{X}}\cdot\nabla B+m\dot{v}_\parallel\mathbf{b}\cdot\dot{\mathbf{X}}. \tag{28}$$

Using

$$\mathbf{E}^\star=-\nabla\phi-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}-\frac{mv_\parallel}{Ze}\frac{\partial \mathbf{b}}{\partial t}, \tag{29}$$

Eq. (28) is written

$$m\dot{v}_\parallel\mathbf{b}\cdot\dot{\mathbf{X}}=-\left(Ze\nabla\phi+\frac{Ze}{c}\frac{\partial \mathbf{A}}{\partial t}+mv_\parallel\frac{\partial \mathbf{b}}{\partial t}+\mu\nabla B\right)\cdot\dot{\mathbf{X}}. \tag{30}$$

Using Eq. (7), i.e., $\mathbf{b}\cdot\dot{\mathbf{X}}=v_\parallel$, the above equation is written

$$mv_\parallel\dot{v}_\parallel=-\left(Ze\nabla\phi+\frac{Ze}{c}\frac{\partial \mathbf{A}}{\partial t}+mv_\parallel\frac{\partial \mathbf{b}}{\partial t}+\mu\nabla B\right)\cdot\dot{\mathbf{X}}, \tag{31}$$

which gives the time change rate of the parallel velocity $v_\parallel$. By using $\dot{\mu}=0$ and $y=\mu B$, the time change rate of the perpendicular velocity is written as

$$\begin{aligned}\dot{y} &= \mu\frac{dB}{dt}\\ &= \mu\left(\frac{\partial B}{\partial t}+\dot{\mathbf{X}}\cdot\nabla B\right)\end{aligned} \tag{32}$$

Next, calculate the total time derivative of the particle energy. The particle energy ε is the sum of the kinetic and potential energy, i.e.,

$$\varepsilon=\frac{1}{2}mv_\parallel^2+y+Ze\phi, \tag{33}$$

from which we obtain

$$\dot{\varepsilon}=mv_\parallel\dot{v}_\parallel+\dot{y}+Ze\dot{\phi}. \tag{34}$$

Using Eqs. (31) and (32), the right-hand side of Eq. (34) is written as

$$\begin{aligned}mv_\parallel\dot{v}_\parallel+\dot{y}+Ze\dot{\phi} &= -\left(\frac{Ze}{c}\frac{\partial \mathbf{A}}{\partial t}+mv_\parallel\frac{\partial \mathbf{b}}{\partial t}+Ze\nabla\phi+\mu\nabla B\right)\cdot\dot{\mathbf{X}}+\mu\left(\frac{\partial B}{\partial t}+\dot{\mathbf{X}}\cdot\nabla B\right)+Ze\dot{\phi}.\\ &= -\left(\frac{Ze}{c}\frac{\partial \mathbf{A}}{\partial t}+mv_\parallel\frac{\partial \mathbf{b}}{\partial t}\right)\cdot\dot{\mathbf{X}}+\mu\frac{\partial B}{\partial t}+Ze\frac{\partial \phi}{\partial t}.\end{aligned} \tag{35}$$

For the equilibrium case, electromagnetic field is independent of time, so the result of the above expression is zero, indicating that energy is a constant of the motion.

1.3 Guiding center drift

Next, we derive the guiding center drift. Using equation (13), we obtain

$$-Ze\mathbf{E}+\mu\nabla B+mv_{\|}\frac{\partial\mathbf{b}}{\partial t}+m\dot{v}_{\|}\mathbf{b}=\dot{\mathbf{X}}\times\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right) \tag{36}$$

Taking cross product of the above equation with $\mathbf{b}/(m\Omega)$, we obtain

$$\frac{1}{m\Omega}\mathbf{b}\times\left(-Ze\mathbf{E}+\mu\nabla B+mv_{\|}\frac{\partial\mathbf{b}}{\partial t}\right)=\frac{1}{m\Omega}\mathbf{b}\times\left[\dot{\mathbf{X}}\times\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)\right] \tag{37}$$

The right-hand side of the above equation is simplified as

$$\begin{aligned}&-\frac{1}{m\Omega}(\mathbf{b}\cdot\dot{\mathbf{X}})\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)+\left[\frac{1}{m\Omega}\mathbf{b}\cdot\left(\frac{Ze}{c}\mathbf{B}+mv_{\|}\nabla\times\mathbf{b}\right)\right]\dot{\mathbf{X}}\\&=-v_{\|}\left(\mathbf{b}+\frac{v_{\|}}{\Omega}\nabla\times\mathbf{b}\right)+\left[1+\frac{1}{\Omega}v_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b}\right]\dot{\mathbf{X}}\\&=-v_{\|}\mathbf{b}-\frac{1}{\Omega}v_{\|}^2\nabla\times\mathbf{b}+\dot{\mathbf{X}}+\left(\frac{1}{\Omega}v_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b}\right)\dot{\mathbf{X}}\end{aligned} \tag{38}$$

Using this, Eq. (37) is written as

$$\begin{aligned}&\frac{1}{m\Omega}\mathbf{b}\times\left(-Ze\mathbf{E}+\mu\nabla B+mv_{\|}\frac{\partial\mathbf{b}}{\partial t}\right)=-v_{\|}\mathbf{b}-\frac{1}{\Omega}v_{\|}^2\nabla\times\mathbf{b}+\dot{\mathbf{X}}+\left(\frac{1}{\Omega}v_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b}\right)\dot{\mathbf{X}}\\&\Rightarrow\dot{\mathbf{X}}=v_{\|}\mathbf{b}+\frac{1}{m\Omega}\mathbf{b}\times\left(-Ze\mathbf{E}+\mu\nabla B+mv_{\|}\frac{\partial\mathbf{b}}{\partial t}\right)+\frac{1}{\Omega}v_{\|}^2\nabla\times\mathbf{b}-\left(\frac{1}{\Omega}v_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b}\right)\dot{\mathbf{X}}\end{aligned} \tag{39}$$

$$\Rightarrow\dot{\mathbf{X}}=v_{\|}\mathbf{b}+\frac{1}{m\Omega}\mathbf{b}\times(-Ze\mathbf{E}+\mu\nabla B)+\frac{1}{m\Omega}\left[\mathbf{b}\times mv_{\|}\frac{\partial\mathbf{b}}{\partial t}+mv_{\|}^2\nabla\times\mathbf{b}-(mv_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b})\dot{\mathbf{X}}\right] \tag{40}$$

Equation (40) contains $\mathbf{E}\times\mathbf{B}$ drift and ∇B drift. The term $v_{\|}/\Omega\mathbf{b}\times\partial\mathbf{b}/\partial t$, (which is one kind of inertial drift since it is proportional to the mass), is small compared with other terms and thus is usually ignored (this term appears in equation (181) of Boozer's paper[2], and is said to be ignorably small in his paper, I do not check this). Using this, Eq. (40) is written

$$\dot{\mathbf{X}}=v_{\|}\mathbf{b}+\frac{1}{m\Omega}\mathbf{b}\times(-Ze\mathbf{E}+\mu\nabla B)+\frac{1}{m\Omega}\left[mv_{\|}^2\nabla\times\mathbf{b}-(mv_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b})\dot{\mathbf{X}}\right] \tag{41}$$

Next we examine the terms in the square bracket of Eq. (41), which can be further written

$$\begin{aligned}&\frac{1}{m\Omega}\left[mv_{\|}^2\nabla\times\mathbf{b}-(mv_{\|}\mathbf{b}\cdot\nabla\times\mathbf{b})\dot{\mathbf{X}}\right]\\&=\frac{1}{m\Omega}\left\{mv_{\|}\left[(\mathbf{b}\cdot\dot{\mathbf{X}})\nabla\times\mathbf{b}-(\mathbf{b}\cdot\nabla\times\mathbf{b})\dot{\mathbf{X}}\right]\right\}\\&=\frac{1}{m\Omega}\left\{mv_{\|}\mathbf{b}\times\left[(\nabla\times\mathbf{b})\times\dot{\mathbf{X}}\right]\right\}\\&=\frac{1}{m\Omega}\mathbf{b}\times\left\{mv_{\|}(\nabla\times\mathbf{b})\times\dot{\mathbf{X}}\right\}\\&=\frac{1}{m\Omega}\mathbf{b}\times\left\{mv_{\|}(\nabla\times\mathbf{b})\times\left(v_{\|}\mathbf{b}+\dot{\mathbf{X}}_\perp\right)\right\}\\&=\frac{1}{m\Omega}\mathbf{b}\times\left\{mv_{\|}^2\boldsymbol{\kappa}+mv_{\|}(\nabla\times\mathbf{b})\times\dot{\mathbf{X}}_\perp\right\}.\end{aligned} \tag{42}$$

The first term of expression (42) is the curvature drift. It is not yet clear what the last term stands for. We now examine this term, i.e.,

$$\frac{1}{m\Omega}\mathbf{b}\times mv_{\|}(\nabla\times\mathbf{b})\times\dot{\mathbf{X}}_\perp, \tag{43}$$

which can be written

$$-\frac{v_{\parallel}}{\Omega}(\mathbf{b}\cdot\nabla\times\mathbf{b})\dot{\mathbf{X}}_{\perp}. \tag{44}$$

In Eq. (66) of Ref. [3], it is pointed out that $\nabla\times\mathbf{b}\approx-\mathbf{b}\times(\mathbf{b}\times\nabla\times\mathbf{b})=\mathbf{b}\times\boldsymbol{\kappa}$, which is correct in the order considered here (I do not check this). This indicates that $\mathbf{b}\cdot\nabla\times\mathbf{b}\approx 0$. Using this, we know the expression in Eq. (44) is approximately zero. Thus, Eq. (41) is written

$$\dot{\mathbf{X}}=v_{\parallel}\mathbf{b}+\frac{1}{m\Omega}\mathbf{b}\times\left(-Ze\mathbf{E}+\mu\nabla B+mv_{\parallel}^2\boldsymbol{\kappa}\right). \tag{45}$$

Note that Eq. (45) does not include the polarization drift, which is proportional to the time derivative of the electric field.

1.4 A more accurate form of the guiding center drift

Equation (41) can also be written

$$\left(1+\frac{v_{\parallel}}{\Omega}\mathbf{b}\cdot\nabla\times\mathbf{b}\right)\dot{\mathbf{X}}=v_{\parallel}\left(\mathbf{b}+\frac{v_{\parallel}}{\Omega}\nabla\times\mathbf{b}\right)+\frac{1}{m\Omega}\mathbf{b}\times(-Ze\mathbf{E}+\mu\nabla B) \tag{46}$$

Define

$$B_{\parallel}^{\star}=B\left(1+\frac{v_{\parallel}}{\Omega}\mathbf{b}\cdot\nabla\times\mathbf{b}\right), \tag{47}$$

which is related to $\mathbf{B}^{\star}$ defined in Eq. (17) through $B_{\parallel}^{\star}=\mathbf{b}\cdot\mathbf{B}^{\star}$, then Eq. (46) is written

$$\dot{\mathbf{X}}=\frac{v_{\parallel}}{B_{\parallel}^{\star}}\left(\mathbf{B}+B\frac{v_{\parallel}}{\Omega}\nabla\times\mathbf{b}\right)-\frac{Ze}{m\Omega B_{\parallel}^{\star}}\mathbf{B}\times\mathbf{E}+\frac{1}{m\Omega B_{\parallel}^{\star}}\mathbf{B}\times\mu\nabla B, \tag{48}$$

i.e.,

$$\dot{\mathbf{X}}=\frac{\mathbf{B}^{\star}}{B_{\parallel}^{\star}}v_{\parallel}-\frac{Ze}{m\Omega B_{\parallel}^{\star}}\mathbf{B}\times\mathbf{E}+\frac{1}{m\Omega B_{\parallel}^{\star}}\mathbf{B}\times\mu\nabla B \tag{49}$$

which agrees with Eqs. (8)-(14) in Todo's paper[4]. Note that, in this form of the guiding center drift, the curvature drift is included in the $\mathbf{B}^{\star}v_{\parallel}/B_{\parallel}^{\star}$ term. Compared with Eq. (45), equation (49) is more accurate because it does not use the approximation that $\mathbf{b}\cdot\nabla\times\mathbf{b}\approx 0$. The numerical results from my numerical code indicate that Eq. (49) can conserve the toroidal angular momentum more accurately than Eq. (45). It is easy to verify that Eq. (49) reduces to Eq. (45) if we use the approximation $\mathbf{b}\cdot\nabla\times\mathbf{b}\approx 0$. Note again that both Eqs. (49) and (45) do not include the polarization drift.

1.5 A more compact form of the time evolution of $v_{\parallel}$

The time evolution of $v_{\parallel}$ is given by Eq. (27), i.e.,

$$mv_{\parallel}\dot{v}_{\parallel}=mv_{\parallel}^2\dot{\mathbf{X}}\cdot\boldsymbol{\kappa}-v_{\parallel}\mu\mathbf{b}\cdot\nabla B+v_{\parallel}Ze\mathbf{b}\cdot\mathbf{E}, \tag{50}$$

which involves the term $\dot{\mathbf{X}}\cdot\boldsymbol{\kappa}$. Next we try to simplify this term. Using Eq. (48), this term is written as

$$\begin{aligned}
mv_{\parallel}^2\dot{\mathbf{X}}\cdot\boldsymbol{\kappa} &= -\frac{mv_{\parallel}^2Ze}{m\Omega B^{\star}}(\mathbf{B}\times\mathbf{E})\cdot\boldsymbol{\kappa}+\frac{mv_{\parallel}^2}{m\Omega B^{\star}}(\mathbf{B}\times\mu\nabla B)\cdot\boldsymbol{\kappa}\\
&= \frac{Zev_{\parallel}^2}{\Omega B^{\star}}(\mathbf{B}\times\mathbf{E})\cdot(\mathbf{b}\times\nabla\times\mathbf{b})-\frac{v_{\parallel}^2}{\Omega B^{\star}}(\mathbf{B}\times\mu\nabla B)\cdot(\mathbf{b}\times\nabla\times\mathbf{b})\\
&= \frac{Zev_{\parallel}^2}{\Omega B^{\star}}(B\mathbf{E}\cdot\nabla\times\mathbf{b}-\mathbf{B}\cdot\nabla\times\mathbf{b}\mathbf{E}\cdot\mathbf{b})-\frac{v_{\parallel}^2}{\Omega B^{\star}}(B\mu\nabla B\cdot\nabla\times\mathbf{b}-\mathbf{B}\cdot\nabla\times\mathbf{b}\mu\nabla B\cdot\mathbf{b})
\end{aligned}$$

Using this in Eq. (50) gives

$$mv_{\parallel}\dot{v}_{\parallel}=\frac{Zev_{\parallel}^2}{\Omega B^{\star}}(B\mathbf{E}\cdot\nabla\times\mathbf{b}-\mathbf{B}\cdot\nabla\times\mathbf{b}\mathbf{E}\cdot\mathbf{b})-\frac{v_{\parallel}^2}{\Omega B^{\star}}(B\mu\nabla B\cdot\nabla\times\mathbf{b}-\mathbf{B}\cdot\nabla\times\mathbf{b}\mu\nabla B\cdot\mathbf{b})-v_{\parallel}\mu\mathbf{b}\cdot\nabla B+v_{\parallel}Ze\mathbf{b}\cdot\mathbf{E},$$

which can be arranged as

$$m v_\parallel \dot{v}_\parallel = \left[\frac{v_\parallel^2}{\Omega B^\star}(B\nabla\times\mathbf{b} - \mathbf{B}\cdot\nabla\times\mathbf{b}\mathbf{b}) + v_\parallel \mathbf{b} \right]\cdot Ze\mathbf{E}, -\left[\frac{v_\parallel^2}{\Omega B^\star} B\nabla\times\mathbf{b} - \frac{v_\parallel^2}{\Omega B^\star}(\mathbf{B}\cdot\nabla\times\mathbf{b})\,\mathbf{b} + v_\parallel\mathbf{b} \right]\cdot$$
$\mu\nabla B$,

which, after some straightforward algebras, can be arranged into the following forms

$$m v_\parallel \dot{v}_\parallel = \frac{\mathbf{B}^\star}{B_\parallel^\star} v_\parallel \cdot (Ze\mathbf{E} - \mu\nabla B), \tag{51}$$

i.e.,

$$\dot{v}_\parallel = \frac{1}{m}\frac{\mathbf{B}^\star}{B_\parallel^\star}\cdot(Ze\mathbf{E} - \mu\nabla B), \tag{52}$$

which agrees with Eq. (15) in Todo's paper[4].

1.6 Summary of equations of guiding center motion

Equations (49), (52), (72), and (47) are repeated here:

$$\dot{\mathbf{X}} = \frac{\mathbf{B}^\star}{B_\parallel^\star} v_\parallel - \frac{Ze}{m\Omega B^\star}\mathbf{B}\times\mathbf{E} + \frac{1}{m\Omega B^\star}\mathbf{B}\times\mu\nabla B \tag{53}$$

$$\dot{v}_\parallel = \frac{1}{m}\frac{\mathbf{B}^\star}{B_\parallel^\star}\cdot(Ze\mathbf{E} - \mu\nabla B), \tag{54}$$

$$\mathbf{B}^\star = \mathbf{B} + B\frac{v_\parallel}{\Omega}\nabla\times\mathbf{b} \tag{55}$$

$$B_\parallel^\star = B\left(1 + \frac{v_\parallel}{\Omega}\mathbf{b}\cdot\nabla\times\mathbf{b}\right), \tag{56}$$

1.7 Generalized toroidal angular momentum

Next, we work in cylindrical coordinates $(R,\ \varphi,\ Z)$ and prove that the generalized momentum conjugating to the toroidal angle φ is a constant of motion for axisymmetric electromagnetic field. The generalized momentum conjugating to φ is defined by

$$P_\varphi = \left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right)_{R,Z,\varphi,v_\parallel,y,\alpha,\dot{R},\dot{Z},\dot{v}_\parallel,\dot{y},\dot{\alpha},t}. \tag{57}$$

Using the Lagrangian given in Eq. (1), Eq. (57) is written

$$\begin{aligned} P_\varphi &= \frac{\partial}{\partial\dot{\varphi}}\left[\left(\frac{Ze}{c}\mathbf{A} + m v_\parallel\mathbf{b}\right)\cdot\dot{\mathbf{X}}\right] \\ &= \left(\frac{Ze}{c}\mathbf{A} + m v_\parallel\mathbf{b}\right)\cdot\frac{\partial\dot{\mathbf{X}}}{\partial\dot{\varphi}} \end{aligned} \tag{58}$$

Noting that $\mathbf{X} = \mathbf{X}(R, Z, \varphi)$, we obtain

$$\dot{\mathbf{X}} = \frac{\partial\mathbf{X}}{\partial R}\dot{R} + \frac{\partial\mathbf{X}}{\partial Z}\dot{Z} + \frac{\partial\mathbf{X}}{\partial\varphi}\dot{\varphi}, \tag{59}$$

from which we obtain

$$\frac{\partial\dot{\mathbf{X}}}{\partial\dot{\varphi}} = \frac{\partial\mathbf{X}}{\partial\varphi}. \tag{60}$$

Further we note that

$$\mathbf{X}(R, Z, \varphi) = R\hat{\mathbf{e}}_R(\varphi) + Z\hat{\mathbf{e}}_Z \Longrightarrow \frac{\partial\mathbf{X}}{\partial\varphi} = R\hat{\mathbf{e}}_\varphi, \tag{61}$$

where $\hat{e}_\varphi$ is the toroidal unit vector. Thus we obtain that

$$\frac{\partial\dot{\mathbf{X}}}{\partial\dot{\varphi}} = R\hat{\mathbf{e}}_\varphi, \tag{62}$$

Using this, Eq. (58) is written as

$$\begin{aligned} P_\varphi &= \left(\frac{Ze}{c}\mathbf{A} + m v_\parallel \mathbf{b}\right) \cdot R\hat{\mathbf{e}}_\varphi \\ &= \frac{Ze}{c} A_\varphi R + m\, v_\parallel \frac{R B_\varphi}{B} \end{aligned}$$

Define $\psi = A_\varphi R$, then the above equation is written as

$$P_\varphi = \frac{Ze}{c}\psi + m\, v_\parallel \frac{R B_\varphi}{B}, \tag{63}$$

which agrees with Eq. (17) in Porcelli's paper[1]. Next we calculate the total time derivative of P_φ, which is given by the Euler-Lagrangian equation corresponding to φ,

$$\dot{P}_\varphi = \frac{\partial \mathcal{L}}{\partial \varphi}. \tag{64}$$

In order to calculate the partial derivative of $\mathcal{L}$ with respect to φ, we write $\mathbf{X}$ in terms of the cylindrical coordinate,

$$\mathbf{X} = R\hat{\mathbf{e}}_R(\varphi) + Z\hat{\mathbf{e}}_Z, \tag{65}$$

the time derivative of which is

$$\dot{\mathbf{X}} = \dot{R}\hat{\mathbf{e}}_R + R\dot{\varphi}\hat{\mathbf{e}}_\varphi + \dot{Z}\hat{\mathbf{e}}_Z. \tag{66}$$

(Note that $\partial \dot{\mathbf{X}} / \partial \varphi \neq 0$.) Then we have

$$\mathbf{A} \cdot \dot{\mathbf{X}} = A_R \dot{R} + A_\varphi R\dot{\varphi} + A_z \dot{Z}, \tag{67}$$

from which we obtain

$$\begin{aligned} \frac{\partial (\mathbf{A} \cdot \dot{\mathbf{X}})}{\partial \varphi} &= \frac{\partial A_R}{\partial \varphi}\dot{R} + \frac{\partial A_\varphi}{\partial \varphi} R\dot{\varphi} + \frac{\partial A_z}{\partial \varphi}\dot{Z} \\ &= \left(\frac{\partial A_R}{\partial \varphi}\hat{\mathbf{e}}_R + \frac{\partial A_\varphi}{\partial \varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial A_z}{\partial \varphi}\hat{\mathbf{e}}_Z\right) \cdot \dot{\mathbf{X}} \end{aligned} \tag{68}$$

Similarly, we obtain

$$\frac{\partial (\mathbf{b} \cdot \dot{\mathbf{X}})}{\partial \varphi} = \left(\frac{\partial b_R}{\partial \varphi}\hat{\mathbf{e}}_R + \frac{\partial b_\varphi}{\partial \varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial b_z}{\partial \varphi}\hat{\mathbf{e}}_Z\right) \cdot \dot{\mathbf{X}} \tag{69}$$

Using these, the partial derivative of $\mathcal{L}$ with respect to φ can be calculated as

$$\frac{\partial \mathcal{L}}{\partial \varphi} = \left[\frac{Ze}{c}\left(\frac{\partial A_R}{\partial \varphi}\hat{\mathbf{e}}_R + \frac{\partial A_\varphi}{\partial \varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial A_z}{\partial \varphi}\hat{\mathbf{e}}_Z\right) + m v_\parallel \left(\frac{\partial b_R}{\partial \varphi}\hat{\mathbf{e}}_R + \frac{\partial b_\varphi}{\partial \varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial b_z}{\partial \varphi}\hat{\mathbf{e}}_Z\right)\right] \cdot \dot{\mathbf{X}} + y\dot{\alpha}\frac{\partial}{\partial \varphi}\left(\frac{1}{\Omega}\right) - Ze\frac{\partial \phi}{\partial \varphi}. \tag{70}$$

The second last term on the right-hand side of Eq. (70) can be further calculated as

$$\begin{aligned} y\dot{\alpha}\frac{\partial}{\partial \varphi}\left(\frac{1}{\Omega}\right) &= y\dot{\alpha}\left(-\frac{1}{\Omega^2}\right)\frac{\partial \Omega}{\partial \varphi} \\ &= y\dot{\alpha}\left(-\frac{1}{B\Omega}\right)\frac{\partial B}{\partial \varphi}. \end{aligned} \tag{71}$$

Using $y = \mu B$ and $\dot{\alpha} = \Omega$ in the above equation, we obtain

$$y\dot{\alpha}\frac{\partial}{\partial \varphi}\left(\frac{1}{\Omega}\right) = -\mu\frac{\partial B}{\partial \varphi}. \tag{72}$$

Using Eq. (72) in Eq. (70) yields

$$\frac{\partial \mathcal{L}}{\partial \varphi} = \left[\frac{Ze}{c}\left(\frac{\partial A_R}{\partial \varphi}\hat{\mathbf{e}}_R + \frac{\partial A_\varphi}{\partial \varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial A_z}{\partial \varphi}\hat{\mathbf{e}}_Z\right) + m v_\parallel \left(\frac{\partial b_R}{\partial \varphi}\hat{\mathbf{e}}_R + \frac{\partial b_\varphi}{\partial \varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial b_z}{\partial \varphi}\hat{\mathbf{e}}_Z\right)\right] \cdot \dot{\mathbf{X}} - \mu\frac{\partial B}{\partial \varphi} - Ze\frac{\partial \phi}{\partial \varphi}. \tag{73}$$

For toroidal symmetrical equilibrium, the partial derivatives with respect to φ are all zeros. In this case Eq. (73) reduces to $\partial\mathcal{L}/\partial\varphi=0$. Using the Euler-Lagrange equation, we obtain $\dot{P}_\varphi=0$, i.e., P_φ is a constant of the motion in symmetrical field. [Note that Eq. (73) is different from Eq. (32) in Porcelli's paper[1], which is given by

$$\frac{\partial\mathcal{L}}{\partial\varphi}=\left[\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial\varphi}+mv_\parallel\frac{\partial\mathbf{b}}{\partial\varphi}\right]\cdot\dot{\mathbf{X}}-\mu\frac{\partial B}{\partial\varphi}-Ze\frac{\partial\phi}{\partial\varphi}. \tag{74}$$

Porcelli's equation is obviously wrong since it can not recover the correct result $\dot{P}_\phi=0$ for axisymmetrical electromagnetic field.]

1.8 A new constant of motion in coherent modes

The time change rate of ε and P_φ is given respectively by Eqs. (35) and (73), i.e.,

$$\dot{\varepsilon}=-\left(\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t}+mv_\parallel\frac{\partial\mathbf{b}}{\partial t}\right)\cdot\dot{\mathbf{X}}+\mu\frac{\partial B}{\partial t}+Ze\frac{\partial\phi}{\partial t}, \tag{75}$$

$$\dot{P}_\varphi=\left[\frac{Ze}{c}\left(\frac{\partial A_R}{\partial\varphi}\hat{\mathbf{e}}_R+\frac{\partial A_\varphi}{\partial\varphi}\hat{\mathbf{e}}_\varphi+\frac{\partial A_z}{\partial\varphi}\hat{\mathbf{e}}_Z\right)+mv_\parallel\left(\frac{\partial b_R}{\partial\varphi}\hat{\mathbf{e}}_R+\frac{\partial b_\varphi}{\partial\varphi}\hat{\mathbf{e}}_\varphi+\frac{\partial b_z}{\partial\varphi}\hat{\mathbf{e}}_Z\right)\right]\cdot\dot{\mathbf{X}}-\mu\frac{\partial B}{\partial\varphi}-Ze\frac{\partial\phi}{\partial\varphi}. \tag{76}$$

From Eqs. (75) and (76), we know that the energy ε is conserved for motion in time independent field while P_φ is conserved for motion in toroidally symmetrical field. For the motion in a toroidal symmetrical equilibrium field superposed by a coherent perturbation $h_1(R,Z)e^{i(-n\varphi-\omega t)}$ with $n\neq 0,\omega\neq 0$, neither of ε and P_φ is conserved. In this case we can construct a new conservative quantity by combining ε and P_φ. Define

$$\varepsilon'\equiv\varepsilon+\frac{\omega}{n}P_\varphi, \tag{77}$$

then it can be proved that $d\varepsilon'/dt=0$ when including only the contribution of the perturbation up to the order $O(h_1/h_0)$ (proof is needed).

1.9 Approximate expression of the drift velocity

The the combined drift due to magnetic curvature and gradient is written

$$\mathbf{v}_d=\frac{1}{B_\parallel^\star}B\frac{v_\parallel}{\Omega}\nabla\times\mathbf{b}v_\parallel+\frac{\mu}{m\Omega B_\parallel^\star}\mathbf{B}\times\nabla B. \tag{78}$$

Using the approximation $B_\parallel^\star\approx B$, $\nabla\times\mathbf{b}\approx\mathbf{b}\times\boldsymbol{\kappa}$, $\boldsymbol{\kappa}\approx-1/R\hat{\mathbf{R}}$, $\nabla B/B=-1/R\hat{\mathbf{R}}$, equation (78) is written

$$\begin{aligned}\mathbf{v}_d &= -\frac{v_\parallel^2}{\Omega R}\mathbf{b}\times\hat{\mathbf{R}}+\frac{mv_\perp^2/2}{m\Omega B}\mathbf{b}\times\nabla B\\ &= -\left[\frac{v_\parallel^2}{\Omega R}+\frac{v_\perp^2/2}{\Omega}\frac{1}{R}\right]\mathbf{b}\times\hat{\mathbf{R}}\\ &= -\frac{v_\parallel^2+v_\perp^2/2}{\Omega R}\mathbf{b}\times\hat{\mathbf{R}}\end{aligned} \tag{79}$$

Equation (79) can also be approximatedly written as

$$\mathbf{v}_d=\frac{v_\parallel^2+v_\perp^2/2}{\Omega}\frac{\mathbf{B}\times\nabla B}{B^2}, \tag{80}$$

which is the equation (2.6.9) in Wesson's book[5].

For a strongly trapped particle, the toroidal drift velocity of the banana orbit is written

2 Drift kinetic equation

The guiding center distribution function is constant along the trajectory of the guiding center in phase space, i.e.,

$$\frac{df}{dt} = 0. \tag{81}$$

Consider the case that the distribution function is independent of the gyro-phase angle α, i.e., $f = f(\mathbf{X}, v_\parallel, y, t)$, then Eq. (81) is written as

$$\frac{\partial f}{\partial t} + \dot{\mathbf{X}} \cdot \nabla f + \dot{v_\parallel}\frac{\partial f}{\partial v_\parallel} + \dot{y}\frac{\partial f}{\partial y} = 0, \tag{82}$$

(which is equation (15) in Porcelli's paper) where the guiding center orbits, $\dot{\mathbf{X}}$, $\dot{v_\parallel}$, and $\dot{y}$, are given by Eqs. (45), (27), and (32), i.e.,

$$\dot{\mathbf{X}} = v_\parallel \mathbf{b} + \frac{1}{m\Omega}\mathbf{b} \times \left(\frac{y}{B}\nabla B + m v_\parallel^2 \boldsymbol{\kappa} - Ze\mathbf{E}\right), \tag{83}$$

$$\dot{v_\parallel} = \frac{1}{m}\left(-\frac{y}{B}\mathbf{b} \cdot \nabla B + Ze\mathbf{b} \cdot \mathbf{E} + m v_\parallel \boldsymbol{\kappa} \cdot \dot{\mathbf{X}}\right), \tag{84}$$

and

$$\dot{y} = \left(\frac{\partial B}{\partial t} + \dot{\mathbf{X}} \cdot \nabla B\right)\frac{y}{B}. \tag{85}$$

We note that, besides the independent variables ($\mathbf{X}$, $v_\parallel$, y), the right-hand side of the Eqs. (83), (84), and (85) depends on the electromagnetic field. Note that, in the perturbation theory, only the electromagnetic field can be perturbed, the independent variables (variables used as the phase space coordinates) are kept fixed.

2.1 Linearized drift kinetic equation

Next, we derive the linearized version of Eq. (82). The perturbation in electromagnetic field causes perturbation in both distribution function and particle orbits $\dot{\mathbf{X}}$, $\dot{v_\parallel}$, and $\dot{y}$. Thus we write

$$f = f_0 + f_1 \tag{86}$$

$$\dot{\mathbf{X}} = \dot{\mathbf{X}}^{(0)} + \dot{\mathbf{X}}^{(1)}, \tag{87}$$

$$\dot{v_\parallel} = \dot{v_\parallel}^{(0)} + \dot{v_\parallel}^{(1)}, \tag{88}$$

$$\dot{y} = \dot{y}^{(0)} + \dot{y}^{(1)}. \tag{89}$$

and substitute this into Eq. (82), we obtain

$$\frac{\partial f_0}{\partial t} + \frac{\partial f_1}{\partial t} + \left(\dot{\mathbf{X}}^{(0)} + \dot{\mathbf{X}}^{(1)}\right) \cdot \nabla f_0 + \left(\dot{\mathbf{X}}^{(0)} + \dot{\mathbf{X}}^{(1)}\right) \cdot \nabla f_1 + (\dot{v_\parallel}^{(0)} + \dot{v_\parallel}^{(1)})\frac{\partial f_0}{\partial v_\parallel} + (\dot{v_\parallel}^{(0)} + \dot{v_\parallel}^{(1)})\frac{\partial f_1}{\partial v_\parallel} + \left(\dot{y}^{(0)} + \dot{y}^{(1)}\right)\frac{\partial f_0}{\partial y} + \left(\dot{y}^{(0)} + \dot{y}^{(1)}\right)\frac{\partial f_1}{\partial y} = 0. \tag{90}$$

The zero order equation is

$$\frac{\partial f_0}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \nabla f_0 + \dot{v_\parallel}^{(0)}\frac{\partial f_0}{\partial v_\parallel} + \dot{y}^{(0)}\frac{\partial f_0}{\partial y} = 0. \tag{91}$$

The first order equation is

$$\frac{\partial f_1}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \nabla f_1 + \dot{v_\parallel}^{(0)}\frac{\partial f_1}{\partial v_\parallel} + \dot{y}^{(0)}\frac{\partial f_1}{\partial y} = -\left[\dot{\mathbf{X}}^{(1)} \cdot \nabla f_0 + \dot{v_\parallel}^{(1)}\frac{\partial f_0}{\partial v_\parallel} + \dot{y}^{(1)}\frac{\partial f_0}{\partial y}\right] \tag{92}$$

Define

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \nabla + \dot{v_\parallel}^{(0)}\frac{\partial}{\partial v_\parallel} + \dot{y}^{(0)}\frac{\partial}{\partial y}, \tag{93}$$

which is the unperturbed orbit propagator, then Eq. (92) is written as

$$\frac{Df_1}{Dt} = -\left[\dot{\mathbf{X}}^{(1)} \cdot \nabla f_0 + \dot{v_\parallel}^{(1)} \frac{\partial f_0}{\partial v_\parallel} + \dot{y}^{(1)} \frac{\partial f_0}{\partial y}\right], \tag{94}$$

which agrees with Eq. (21) in Porcelli's paper[1].

At this point, I would like to discuss the equilibrium distribution function. We know that functions of constants of the motion are solutions to the kinetic equation. Noting that $P_{\varphi 0}$, ε_0, and μ_0 are constants of the motion in equilibrium field. Then $f_0 = F(P_{\varphi 0}, \varepsilon_0, \mu_0)$ is a solution to the kinetic equation Eq. (82). Noting that the right-hand side of Eq. (94) contains partial derivative with respect to variables $(\mathbf{X}, v_\parallel, y)$, we would like to transform the partial derivatives with respect to $(\mathbf{X}, v_\parallel, y)$ to one with respect to $(P_{\varphi 0}, \varepsilon_0, \mu_0)$ because f_0 is usually specified in terms of the vairables $(P_{\varphi 0}, \varepsilon_0, \mu_0)$. The right-hand side of Eq. (94) can be written term by term as

$$\begin{aligned}\dot{\mathbf{X}}^{(1)} \cdot \nabla f_0 &= \dot{\mathbf{X}}^{(1)} \cdot \left(\frac{\partial F}{\partial P_{\varphi 0}} \nabla P_{\varphi 0} + \frac{\partial F}{\partial \varepsilon_0} \nabla \varepsilon_0 + \frac{\partial F}{\partial \mu_0} \nabla \mu_0\right) \\ &= \dot{\mathbf{X}}^{(1)} \cdot \left(\frac{\partial F}{\partial P_{\varphi 0}} \nabla P_{\varphi 0} + \frac{\partial F}{\partial \varepsilon_0} Ze \nabla \phi_0 - \frac{\partial F}{\partial \mu_0} \frac{y}{B_0^2} \nabla B_0\right)\end{aligned} \tag{95}$$

$$\begin{aligned}\dot{v_\parallel}^{(1)} \frac{\partial f_0}{\partial v_\parallel} &= \dot{v_\parallel}^{(1)} \left(\frac{\partial F}{\partial P_{\varphi 0}} \frac{\partial P_{\varphi 0}}{\partial v_\parallel} + \frac{\partial F}{\partial \varepsilon_0} \frac{\partial \varepsilon_0}{\partial v_\parallel} + \frac{\partial F}{\partial \mu_0} \frac{\partial \mu_0}{\partial v_\parallel}\right) \\ &= \dot{v_\parallel}^{(1)} \left(\frac{\partial F}{\partial P_{\varphi 0}} \frac{\partial P_{\varphi 0}}{\partial v_\parallel} + \frac{\partial F}{\partial \varepsilon_0} m v_\parallel\right)\end{aligned} \tag{96}$$

$$\begin{aligned}\dot{y}^{(1)} \frac{\partial f_0}{\partial y} &= \dot{y}^{(1)} \left(\frac{\partial F}{\partial P_{\varphi 0}} \frac{\partial P_{\varphi 0}}{\partial y} + \frac{\partial F}{\partial \varepsilon_0} \frac{\partial \varepsilon_0}{\partial y} + \frac{\partial F}{\partial \mu_0} \frac{\partial \mu_0}{\partial y}\right) \\ &= \dot{y}^{(1)} \left(\frac{\partial F}{\partial \varepsilon_0} + \frac{\partial F}{\partial \mu_0} \frac{1}{B_0}\right)\end{aligned} \tag{97}$$

Using these results, Eq. (94) is written as

$$\frac{Df_1}{Dt} = -\left[\dot{\mathbf{X}}^{(1)} \cdot \left(\frac{\partial F}{\partial P_{\varphi 0}} \nabla P_{\varphi 0} + \frac{\partial F}{\partial \varepsilon_0} Ze \nabla \phi_0 - \frac{\partial F}{\partial \mu_0} \frac{y}{B_0^2} \nabla B_0\right) + \dot{v_\parallel}^{(1)} \left(\frac{\partial F}{\partial P_{\varphi 0}} \frac{\partial P_{\varphi 0}}{\partial v_\parallel} + \frac{\partial F}{\partial \varepsilon_0} m v_\parallel\right) + \dot{y}^{(1)} \left(\frac{\partial F}{\partial \varepsilon_0} + \frac{\partial F}{\partial \mu_0} \frac{1}{B_0}\right)\right],$$

which can be arranged in the form

$$\frac{Df_1}{Dt} = -\left[\left(\dot{\mathbf{X}}^{(1)} \cdot \nabla P_{\varphi 0} + \dot{v_\parallel}^{(1)} \frac{\partial P_{\varphi 0}}{\partial v_\parallel}\right) \frac{\partial F}{\partial P_{\varphi 0}} + \left(Ze \dot{\mathbf{X}}^{(1)} \cdot \nabla \phi_0 + m v_\parallel \dot{v_\parallel}^{(1)} + \dot{y}^{(1)}\right) \frac{\partial F}{\partial \varepsilon_0} + \left(\frac{\dot{y}^{(1)}}{B_0} - \frac{y}{B_0^2} \dot{\mathbf{X}}^{(1)} \cdot \nabla B_0\right) \frac{\partial F}{\partial \mu_0}\right], \tag{98}$$

which agrees with Eq. (22) in Porcelli's paper[1]. Next, we need to express $\dot{\mathbf{X}}^{(1)}$, $\dot{v}_\parallel^{(1)}$, and $\dot{y}^{(1)}$ in terms of the perturbed electromagnetic field. Let us first consider the coefficient befor the $\partial F / \partial \varepsilon_0$ term in Eq. (98). We note that Eq. (35) takes the following form:

$$m v_\parallel \dot{v_\parallel} + \dot{y} + Ze \dot{\phi} = -\left(\frac{Ze}{c} \frac{\partial \mathbf{A}}{\partial t} + m v_\parallel \frac{\partial \mathbf{b}}{\partial t}\right) \cdot \dot{\mathbf{X}} + \frac{y}{B} \frac{\partial B}{\partial t} + Ze \frac{\partial \phi}{\partial t}, \tag{99}$$

whose linearized version is (noting that $\mathbf{A}^{(0)}$, $\mathbf{b}^{(0)}$, and $B^{(0)}$ is time independent, thus $\partial \mathbf{A}^{(0)} / \partial t$, $\partial \mathbf{b}^{(0)} / \partial t$, and $\partial B^{(0)} / \partial t$ are all zeros)

$$m v_\parallel \dot{v_\parallel}^{(1)} + \dot{y}^{(1)} + Ze \dot{\phi}^{(1)} = -\left(\frac{Ze}{c} \frac{\partial \mathbf{A}^{(1)}}{\partial t} + m v_\parallel \frac{\partial \mathbf{b}^{(1)}}{\partial t}\right) \cdot \dot{\mathbf{X}}^{(0)} + \mu_0 \frac{\partial B^{(1)}}{\partial t} + Ze \frac{\partial \phi^{(1)}}{\partial t}, \tag{100}$$

where $\mu_0 = y/B_0$. Using

$$\begin{aligned}\dot{\phi}^{(1)} &\equiv \left(\frac{d\phi}{dt}\right)^{(1)} \\ &= \left(\frac{\partial\phi}{\partial t} + \dot{\mathbf{X}}\cdot\nabla\phi + 0 + 0 + 0\right)^{(1)} \\ &= \frac{\partial\phi^{(1)}}{\partial t} + \dot{\mathbf{X}}^{(0)}\cdot\nabla\phi^{(1)} + \dot{\mathbf{X}}^{(1)}\cdot\nabla\phi^{(0)} \\ &= \frac{D\phi^{(1)}}{Dt} + \dot{\mathbf{X}}^{(1)}\cdot\nabla\phi^{(0)} \end{aligned} \tag{101}$$

(Note that $\frac{D}{Dt}$ here denotes total time derivative along the unperturbed orbit, instead of the perturbed orbit) in Eq. (100) gives

$$m v_{\parallel}\dot{v_{\parallel}}^{(1)} + \dot{y}^{(1)} + Ze\dot{\mathbf{X}}^{(1)}\cdot\nabla\phi^{(0)} = -L_t^{(1)} - Ze\frac{D\phi^{(1)}}{Dt}, \tag{102}$$

where, for notation ease, we have defined

$$-\mathcal{L}_t^{(1)} = -\left(\frac{Ze}{c}\frac{\partial\mathbf{A}^{(1)}}{\partial t} + m v_{\parallel}\frac{\partial\mathbf{b}^{(1)}}{\partial t}\right)\cdot\dot{\mathbf{X}}^{(0)} + \mu_0\frac{\partial B^{(1)}}{\partial t} + Ze\frac{\partial\phi^{(1)}}{\partial t}. \tag{103}$$

Equation (102) agrees with Eq. (30) in Porcelli's paper[1]. The right-hand side of Eq. (102) provides the desired expression for the coefficent before the $\partial F/\partial\varepsilon_0$ term of Eq. (98).

The first order equation of Eq. (73) is written as (Noting that we are considering toroidal symmetrical equilibrium, thus terms such as $\partial A_R^{(0)}/\partial\varphi$, $\partial b_R^{(0)}/\partial\varphi$, and $\partial B^{(0)}/\partial\varphi$ are all zeros.)

$$\left(\frac{\partial\mathcal{L}}{\partial\varphi}\right)^{(1)} = \left[\frac{Ze}{c}\left(\frac{\partial A_R^{(1)}}{\partial\varphi}\hat{\mathbf{e}}_R + \frac{\partial A_\varphi^{(1)}}{\partial\varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial A_z^{(1)}}{\partial\varphi}\hat{\mathbf{e}}_Z\right) + m v_{\parallel}\left(\frac{\partial b_R^{(1)}}{\partial\varphi}\hat{\mathbf{e}}_R + \frac{\partial b_\varphi^{(1)}}{\partial\varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial b_z^{(1)}}{\partial\varphi}\hat{\mathbf{e}}_Z\right)\right]\cdot\dot{\mathbf{X}}^{(0)} - \mu_0\frac{\partial B^{(1)}}{\partial\varphi} - Ze\frac{\partial\phi^{(1)}}{\partial\varphi}. \tag{104}$$

Using this in the Euler equation (64), we obtain

$$(\dot{P}_\varphi)^{(1)} = \left(\frac{\partial\mathcal{L}}{\partial\varphi}\right)^{(1)} \tag{105}$$

Note that

$$P_\varphi = \frac{Ze}{c}A_\varphi R + m R v_{\parallel}\frac{B_\varphi}{B}, \tag{106}$$

then

$$\dot{P}_\varphi = \frac{\partial P_\varphi}{\partial t} + \dot{\mathbf{X}}\cdot\nabla P_\varphi + \dot{v}_{\parallel}\frac{\partial P_\varphi}{\partial v_{\parallel}}$$

Using this in Eq. (105), we obtain

$$\left(\frac{\partial P_\varphi}{\partial t} + \dot{\mathbf{X}}\cdot\nabla P_\varphi + \dot{v}_{\parallel}\frac{\partial P_\varphi}{\partial v_{\parallel}}\right)^{(1)} = \left(\frac{\partial\mathcal{L}}{\partial\varphi}\right)^{(1)}, \tag{107}$$

which can be further written as

$$\dot{\mathbf{X}}^{(1)}\cdot\nabla P_\varphi^{(0)} + \dot{v}_{\parallel}^{(1)}\frac{\partial P_\varphi^{(0)}}{\partial v_{\parallel}} + \frac{\partial P_\varphi^{(1)}}{\partial t} + \dot{\mathbf{X}}^{(0)}\cdot\nabla P_\varphi^{(1)} + \dot{v}_{\parallel}^{(0)}\frac{\partial P_\varphi^{(1)}}{\partial v_{\parallel}} = \left(\frac{\partial\mathcal{L}}{\partial\varphi}\right)^{(1)}. \tag{108}$$

$$\Longrightarrow \dot{\mathbf{X}}^{(1)}\cdot\nabla P_\varphi^{(0)} + \dot{v}_{\parallel}^{(1)}\frac{\partial P_\varphi^{(0)}}{\partial v_{\parallel}} = \left(\frac{\partial\mathcal{L}}{\partial\varphi}\right)^{(1)} - \frac{DP_\varphi^{(1)}}{Dt} \tag{109}$$

The right-hand side of Eq. (109) gives desired expression for the coefficient before the term $\partial F/\partial P_{\varphi 0}$ of Eq. (98). The linearized version of Eq. (32)

$$\dot{y} = \mu\left(\frac{\partial B}{\partial t} + \dot{\mathbf{X}}\cdot\nabla B\right), \tag{110}$$

is written as

$$\dot{y}^{(1)} = \mu_0 \left(\frac{\partial B_1}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \nabla B_1 + \dot{\mathbf{X}}^{(1)} \cdot \nabla B_0 \right) + \mu^{(1)} \dot{\mathbf{X}}^{(0)} \cdot \nabla B_0. \tag{111}$$

Noting that $\mu_0 = y / B_0$, $\mu^{(1)} = -(y / B_0^2) B_1$, the above equation is written as

$$\frac{\dot{y}^{(1)}}{B_0} = \frac{y}{B_0} \left(\frac{\partial B_1 / B_0}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \frac{1}{B_0} \nabla B_1 + \frac{1}{B_0} \dot{\mathbf{X}}^{(1)} \cdot \nabla B_0 \right) - \frac{y}{B_0^2} B_1 \dot{\mathbf{X}}^{(0)} \cdot \frac{1}{B_0} \nabla B_0 \tag{112}$$

$$\frac{\dot{y}^{(1)}}{B_0} - y \frac{1}{B_0^2} \dot{\mathbf{X}}^{(1)} \cdot \nabla B^{(0)} = \mu_0 \left(\frac{\partial B^{(1)} / B_0}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \frac{1}{B_0} \nabla B^{(1)} - B^{(1)} \dot{\mathbf{X}}^{(0)} \cdot \frac{1}{B_0^2} \nabla B^{(0)} \right) \tag{113}$$

$$\frac{\dot{y}^{(1)}}{B_0} - y \frac{1}{B_0^2} \dot{\mathbf{X}}^{(1)} \cdot \nabla B^{(0)} = \mu_0 \left[\frac{\partial B^{(1)} / B_0}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \left(\frac{1}{B_0} \nabla B^{(1)} - B^{(1)} \cdot \frac{1}{B_0^2} \nabla B^{(0)} \right) \right] \tag{114}$$

$$\frac{\dot{y}^{(1)}}{B_0} - y \frac{1}{B_0^2} \dot{\mathbf{X}}^{(1)} \cdot \nabla B^{(0)} = \mu_0 \left(\frac{\partial B^{(1)} / B_0}{\partial t} + \dot{\mathbf{X}}^{(0)} \cdot \nabla \frac{B^{(1)}}{B^{(0)}} \right) \tag{115}$$

$$\frac{\dot{y}^{(1)}}{B_0} - y \frac{1}{B_0^2} \dot{\mathbf{X}}^{(1)} \cdot \nabla B_0 = \mu_0 \frac{D}{Dt} \left(\frac{B_1}{B_0} \right). \tag{116}$$

Eq. (116) agrees with Eq. (31) in Porceli's paper. The right-hand side of Eq. (116) provide the desired expression for the coefficient before $\partial F / \partial \mu_0$ term in Eq. (98). Using Eqs. (102), (109), and (116), Eq. (98) is finally written as

$$\frac{D f_1}{Dt} = - \left\{ \left[\left(\frac{\partial \mathcal{L}}{\partial \varphi} \right)^{(1)} - \frac{D P_\varphi^{(1)}}{Dt} \right] \frac{\partial F}{\partial P_{\varphi 0}} + \left(- L_t^{(1)} - Ze \frac{D \phi^{(1)}}{Dt} \right) \frac{\partial F}{\partial \varepsilon_0} + \mu_0 \frac{D}{Dt} \left(\frac{B_1}{B_0} \right) \frac{\partial F}{\partial \mu_0} \right\}, \tag{117}$$

2.2 Separation of perturbed distribution into adiabatic and non-adiabatic parts

Write f_1 as

$$f_1 = P_\varphi^{(1)} \frac{\partial F}{\partial P_{\varphi 0}} + Ze \phi^{(1)} \frac{\partial F}{\partial \varepsilon_0} - \mu_0 \frac{B_1}{B_0} \frac{\partial F}{\partial \mu_0} + h^{(1)}, \tag{118}$$

and substitute this into Eq. (117), giving an equation of $h^{(1)}$,

$$\frac{D h^{(1)}}{Dt} = - \left(\frac{\partial \mathcal{L}}{\partial \varphi} \right)^{(1)} \frac{\partial F}{\partial P_{\varphi 0}} + L_t^{(1)} \frac{\partial F}{\partial \varepsilon_0}. \tag{119}$$

Eq. (119) agrees with Eq. (35) in Porcelli's paper[1]. For notation convenience, we define $\mathcal{L}^{(1)}$ as

$$\mathcal{L}^{(1)} = \left(\frac{Ze}{c} \mathbf{A}^{(1)} + m v_\parallel \mathbf{b}^{(1)} \right) \cdot \dot{\mathbf{X}}^{(0)} - \mu_0 B^{(1)} - Ze \phi^{(1)}, \tag{120}$$

which can be called "perturbed Lagrangian" (I do not care the name of $\mathcal{L}^{(1)}$, which is only a notation without any physical meaning since I do not need this meaning to derive anything). Since we are considering toroidal symmetrical case, different toroidal harmonics of perturbation are independent. Thus we can consider a single toroidal harmonic, i.e, the φ dependence of components of $\mathbf{A}^{(1)}$ and $\mathbf{b}^{(1)}$ is $\exp(-in\varphi)$. Further due to that the equilibrium is time-independent, we can consider a single time harmonic, i.e., the time dependence of the perturbation is $\exp(-i\omega t)$. Then, using Eq. (120), Eq. (104) is written as

$$\left(\frac{\partial \mathcal{L}}{\partial \varphi} \right)^{(1)} = -in \mathcal{L}^{(1)}, \tag{121}$$

and $\mathcal{L}_t^{(1)}$ in Eq. (103) is written as

$$\mathcal{L}_t^{(1)} = -i\omega\mathcal{L}^{(1)}. \tag{122}$$

Using Eqs. (121) and (122), Eq. (119) is written as

$$\frac{Dh^{(1)}}{Dt} = in\mathcal{L}^{(1)}\frac{\partial F}{\partial P_{\varphi 0}} - i\omega\mathcal{L}^{(1)}\frac{\partial F}{\partial \varepsilon_0}. \tag{123}$$

Define

$$\omega_\star \equiv \frac{\partial F}{\partial P_{\varphi 0}} \Big/ \frac{\partial F}{\partial \varepsilon_0}, \tag{124}$$

then Eq. (123) is written as

$$\frac{Dh^{(1)}}{Dt} = -i\,(\omega - n\omega_\star)\frac{\partial F}{\partial \varepsilon_0}\mathcal{L}^{(1)}. \tag{125}$$

Eq. (125) agrees with Eq. (40) in Porcelli's paper. Define $g^{(1)}$ as

$$h^{(1)} = -i\,(\omega - n\omega_\star)\frac{\partial F}{\partial \varepsilon_0}g^{(1)}, \tag{126}$$

then $g^{(1)}$ satisfies (note that $\omega_\star$ and $\partial F/\partial\varepsilon_0$ are both functions of constants of motion, thus can be taken out of the orbit integration)

$$\frac{Dg^{(1)}}{Dt} = \mathcal{L}^{(1)}. \tag{127}$$

2.3 Further simplification of the perturbed Lagrangian $\mathcal{L}^{(1)}$

The expression $\mathcal{L}^{(1)}$ in Eq. (120) can be further simplified, by noticing that the term $mv_\parallel\mathbf{b}^{(1)}\cdot\dot{\mathbf{X}}^{(0)}$ is of the order $O(\delta^2)$ thus can be neglected, giving

$$\mathcal{L}^{(1)} = \frac{Ze}{c}\mathbf{A}^{(1)}\cdot\dot{\mathbf{X}}^{(0)} - \mu_0 B^{(1)} - Ze\phi^{(1)}. \tag{128}$$

Next, we provide the proof that the term $mv_\parallel\mathbf{b}^{(1)}\cdot\dot{\mathbf{X}}^{(0)}$ is on the order $O(\delta^2)$. The unperturbed velocity of guiding center is given by

$$\begin{aligned}\dot{\mathbf{X}}^{(0)} &= v_\parallel\mathbf{b} + \frac{1}{m\Omega}\mathbf{b}^{(0)}\times\left(\frac{y}{B_0}\nabla B_0 + mv_\parallel^2\boldsymbol{\kappa}^{(0)}\right)\\ &= v_\parallel\mathbf{b} + \mathbf{v}_d\end{aligned} \tag{129}$$

where $v_\parallel\sim O(\delta^0)$, $v_d\sim O(\delta)$. Next, we derive the expression of $\mathbf{b}^{(1)}$. Using

$$\begin{aligned}\frac{1}{B} &= \frac{1}{\sqrt{B_0^2 + \left(B^{(1)}\right)^2 + 2\mathbf{B}_0\cdot\mathbf{B}^{(1)}}}\\ &\approx \frac{1}{B_0} - \frac{1}{2B_0^3}\left[\left(B^{(1)}\right)^2 + 2\mathbf{B}_0\cdot\mathbf{B}^{(1)}\right]\end{aligned} \tag{130}$$

we obtain

$$\begin{aligned}\mathbf{b} &= \frac{\mathbf{B}_0 + \mathbf{B}^{(1)}}{B}\\ &= \left(\mathbf{B}_0 + \mathbf{B}^{(1)}\right)\left\{\frac{1}{B_0} - \frac{1}{2B_0^3}\left[\left(B^{(1)}\right)^2 + 2\mathbf{B}_0\cdot\mathbf{B}^{(1)}\right]\right\}\\ &\approx \mathbf{b}^{(0)} - \frac{\mathbf{B}_0}{2B_0^3}\left(2\mathbf{B}_0\cdot\mathbf{B}^{(1)}\right) + \frac{\mathbf{B}^{(1)}}{B_0}\\ &= \mathbf{b}^{(0)} - \mathbf{b}^{(0)}\left(\mathbf{b}^{(0)}\cdot\frac{\mathbf{B}^{(1)}}{B_0}\right) + \frac{\mathbf{B}^{(1)}}{B_0}.\end{aligned}$$

Therefore

$$\mathbf{b}^{(1)} = -\mathbf{b}^{(0)}\left(\mathbf{b}^{(0)} \cdot \frac{\mathbf{B}^{(1)}}{B_0}\right) + \frac{\mathbf{B}^{(1)}}{B_0} \tag{131}$$

Using this, the term $m v_\parallel \mathbf{b}^{(1)} \cdot \dot{\mathbf{X}}^{(0)}$ is written as

$$\begin{aligned} m v_\parallel \mathbf{b}^{(1)} \cdot \dot{\mathbf{X}}^{(0)} &= -m v_\parallel^2 \mathbf{b}^{(0)} \cdot \frac{\mathbf{B}^{(1)}}{B_0} + m v_\parallel^2 \mathbf{b}^{(0)} \cdot \frac{\mathbf{B}^{(1)}}{B_0} + m v_\parallel \mathbf{b}^{(1)} \cdot \mathbf{v}_d & (132)\\ &= 0 + m v_\parallel \mathbf{b}^{(1)} \cdot \mathbf{v}_d, & (133) \end{aligned}$$

where the first two terms on the right-hand side of Eq. (132), which are on the order $O(\delta)$, happen to cancel each other. Since $\mathbf{b}^{(1)} \sim O(\delta)$ and $\mathbf{v}_d \sim O(\delta)$, the product of these two terms are on the order $O(\delta^2)$. Therefore Eq. (133) indicates that the term $m v_\parallel \mathbf{b}^{(1)} \cdot \dot{\mathbf{X}}^{(0)}$ is on the order $O(\delta^2)$.

Next we show that, in the linear approximation, the perturbation in the strength of the magnetic field, $B^{(1)}$, is equal to $B_\parallel^{(1)}$, where $B_\parallel^{(1)} \equiv \mathbf{b}^{(0)} \cdot \mathbf{B}^{(1)}$. The total magnetic field can be written as

$$\begin{aligned} B &= \sqrt{\mathbf{B} \cdot \mathbf{B}} \\ &= \sqrt{(\mathbf{B}_0 + \mathbf{B}^{(1)}) \cdot (\mathbf{B}_0 + \mathbf{B}^{(1)})} \\ &= \sqrt{B_0^2 + (B^{(1)})^2 + 2\mathbf{B}_0 \cdot \mathbf{B}^{(1)}} & (134) \end{aligned}$$

Expanding the right-hand side of the above equation at B_0, we obtain

$$B \approx \sqrt{B_0^2} + \frac{1}{2\sqrt{B_0^2}}\left[(B^{(1)})^2 + 2\mathbf{B}_0 \cdot \mathbf{B}^{(1)}\right] + \ldots \tag{135}$$

Neglecting the second order term, the above equation is written as

$$\begin{aligned} B &\approx B_0 + \frac{\mathbf{B}_0 \cdot \mathbf{B}^{(1)}}{B_0} \\ &= B_0 + \mathbf{b}^{(0)} \cdot \mathbf{B}^{(1)} \\ &= B_0 + B_\parallel^{(1)} & (136) \end{aligned}$$

Thus we obtain

$$B^{(1)} = B_\parallel^{(1)}. \tag{137}$$

Eq. (137) seems strange at first glance (actually I think it is wrong at first glance, Dr. Fu let me know it is right and how to prove it (as given in the above)). Using Eq. (137), the Lagrangian in Eq. (128) is written as

$$\mathcal{L}^{(1)} = \frac{Ze}{c}\mathbf{A}^{(1)} \cdot \dot{\mathbf{X}}^{(0)} - \mu_0 B_\parallel^{(1)} - Ze\phi^{(1)}. \tag{138}$$

2.4 Unperturbed orbit integration———not finished————

For the ease of notation, in the following we drop the zero superscript on the unperturbed orbit. And to distinguish instantaneous and the initial value of orbit, we add a prime to $\mathbf{X}$ and $v_\parallel$ to denote the instantaneous value. Integrating along the unperturbed orbit, Eq. (127) is written as

$$g^{(1)}(\mathbf{X}, v_\parallel, y) = \int_{-\infty}^{0} \mathcal{L}^{(1)}(\tau) d\tau. \tag{139}$$

with the boundary condition

$$\mathbf{X}'(\tau = 0) = \mathbf{X}, \tag{140}$$

$$v_\parallel'(\tau = 0) = v_\parallel, \tag{141}$$

and the value of the conserved magnetic moment is determined by$\mu = y / B(\mathbf{X})$. Using the expression of $\mathcal{L}^{(1)}$ in Eq. (128), Eq. (139) is written as

$$\begin{aligned} g^{(1)} &= \int_{-\infty}^{0} \left[\frac{Ze}{c} \mathbf{A}^{(1)} \cdot \dot{\mathbf{X}}' - \mu_0 B_{\parallel}^{(1)} - Ze\phi^{(1)} \right] d\tau \\ &= \end{aligned}$$

$$B^{(1)}(\psi, \theta, \varphi, t) = \hat{B}^{(1)}(\psi, \theta) \exp[-i(\omega t + n\varphi)] \tag{142}$$

$$\dot{\psi} = \dot{\mathbf{X}} \cdot \nabla \psi \tag{143}$$

$$\dot{\theta} = \dot{\mathbf{X}} \cdot \nabla \theta \tag{144}$$

$$\dot{\varphi} = \dot{\mathbf{X}} \cdot \nabla \varphi \tag{145}$$

$$\varphi'(t + l t_p) = \varphi'(t) + 2l\pi \tag{146}$$

3 Perturbed Lagrangian for ideal MHD perturbation

For ideal MHD perturbation, the perturbed magnetic field is written as

$$\mathbf{B}^{(1)} = \nabla \times (\boldsymbol{\xi}_\perp \times \mathbf{B}_0) \tag{147}$$

$$= -\mathbf{B}_0 (\nabla \cdot \boldsymbol{\xi}_\perp) + \mathbf{B}_0 \cdot \nabla \boldsymbol{\xi}_\perp - \boldsymbol{\xi}_\perp \cdot \nabla \mathbf{B}_0, \tag{148}$$

Using Eq. (147), the vector potential of magnetic perturbation is written as

$$\mathbf{A}_1 = \boldsymbol{\xi}_\perp \times \mathbf{B}_0 \tag{149}$$

Using Eq. (148), the parallel component of the magnetic perturbation is written as

$$\begin{aligned} B_{\parallel}^{(1)} &\equiv \mathbf{b}^{(0)} \cdot \mathbf{B}^{(1)} \\ &= -B_0 (\nabla \cdot \boldsymbol{\xi}_\perp) + [\mathbf{B}_0 \cdot \nabla \boldsymbol{\xi}_\perp] \cdot \mathbf{b}^{(0)} - [\boldsymbol{\xi}_\perp \cdot \nabla \mathbf{B}_0] \cdot \mathbf{b}^{(0)} \end{aligned} \tag{150}$$

Here I have some important remarks about tensor identities (I had not known these identities before CaiHuiShan told me). First we note the associate law applies in this case (Important!), thus

$$B_{\parallel}^{(1)} = -B_0 (\nabla \cdot \boldsymbol{\xi}_\perp) + \mathbf{B}_0 \cdot \left(\nabla \boldsymbol{\xi}_\perp \cdot \mathbf{b}^{(0)} \right) - \boldsymbol{\xi}_\perp \cdot \left(\nabla \mathbf{B}_0 \cdot \mathbf{b}^{(0)} \right). \tag{151}$$

Second we have the tensor identity (Important!, CaiHuiShan let me know this identity),

$$\nabla (\mathbf{A} \cdot \mathbf{B}) = (\nabla \mathbf{A}) \cdot \mathbf{B} + (\nabla \mathbf{B}) \cdot \mathbf{A} \tag{152}$$

Using this, the second term of Eq. (151) is written as

$$\begin{aligned} \mathbf{B}_0 \cdot (\nabla \boldsymbol{\xi}_\perp \cdot \mathbf{b}) &= \mathbf{B}_0 \cdot [\nabla (\boldsymbol{\xi}_\perp \cdot \mathbf{b}) - \nabla \mathbf{b} \cdot \boldsymbol{\xi}_\perp] \\ &= \mathbf{B}_0 \cdot [0 - \nabla \mathbf{b} \cdot \boldsymbol{\xi}_\perp] \\ &= -B_0 \boldsymbol{\kappa} \cdot \boldsymbol{\xi}_\perp, \end{aligned} \tag{153}$$

where $\boldsymbol{\kappa} = \mathbf{b} \cdot \nabla \mathbf{b}$. The last term of Eq. (151) is written as

$$\begin{aligned} \boldsymbol{\xi}_\perp \cdot (\nabla \mathbf{B}_0 \cdot \mathbf{b}) &= \boldsymbol{\xi}_\perp \cdot [\nabla (B_0 \mathbf{b}) \cdot \mathbf{b}] \\ &= \boldsymbol{\xi}_\perp \cdot [(\mathbf{b} \nabla B_0 + B_0 \nabla \mathbf{b}) \cdot \mathbf{b}]. \end{aligned} \tag{154}$$

We note that $\nabla \mathbf{b} \cdot \mathbf{b} = 0$ (CaiHuiShan let me know this), since the tensor identity in Eq. (152) indicates

$$0 = \nabla (\mathbf{b} \cdot \mathbf{b}) = 2 \nabla \mathbf{b} \cdot \mathbf{b}. \tag{155}$$

(It is interesting to note that $\mathbf{b}\cdot\nabla\mathbf{b}\equiv\boldsymbol{\kappa}\neq 0$ while the above result proves that $\nabla\mathbf{b}\cdot\mathbf{b}=0$.) Thus Eq. (154) becomes

$$\begin{aligned}\boldsymbol{\xi}_\perp\cdot(\nabla\mathbf{B}_0\cdot\mathbf{b}) &= \boldsymbol{\xi}_\perp\cdot[(\mathbf{b}\nabla B_0)\cdot\mathbf{b}]\\ &= \boldsymbol{\xi}_\perp\cdot\nabla B_0\end{aligned} \tag{156}$$

Using the above results, the parallel component of the perturbed magnetic field Eq. (151) becomes

$$B_\parallel^{(1)}=-B_0(\nabla\cdot\boldsymbol{\xi}_\perp)-B_0\boldsymbol{\kappa}\cdot\boldsymbol{\xi}_\perp-\boldsymbol{\xi}_\perp\cdot\nabla B_0 \tag{157}$$

The perturbed Lagrangian, Eq. (138), is

$$\mathcal{L}^{(1)}=\frac{Ze}{c}\mathbf{A}^{(1)}\cdot\dot{\mathbf{X}}^{(0)}-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)} \tag{158}$$

Using Eq. (149) in Eq. (158) yields

$$\mathcal{L}^{(1)}=\frac{Ze}{c}(\boldsymbol{\xi}_\perp\times\mathbf{B}_0)\cdot\dot{\mathbf{X}}^{(0)}-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)} \tag{159}$$

Using Eq. (83) for $\dot{\mathbf{X}}^{(0)}$, Eq. (159) is written as

$$\begin{aligned}\mathcal{L}^{(1)} &= \frac{Ze}{c}\big(\boldsymbol{\xi}_\perp\times\mathbf{B}^{(0)}\big)\cdot\left[v_\parallel\mathbf{b}^{(0)}+\frac{1}{m\Omega}\mathbf{b}^{(0)}\times\left(\frac{y}{B^{(0)}}\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}-Ze\mathbf{E}^{(0)}\right)\right]-\mu_0B_\parallel^{(1)}-\\ &\quad Ze\phi^{(1)}\\
&= \frac{Ze}{c}\big(\boldsymbol{\xi}_\perp\times\mathbf{B}^{(0)}\big)\cdot\left[\frac{1}{m\Omega}\mathbf{b}^{(0)}\times\left(\frac{y}{B^{(0)}}\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}+Ze\nabla\varphi^{(0)}\right)\right]-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)}\\
&= \frac{Ze}{c}\boldsymbol{\xi}_\perp\cdot\left\{\mathbf{B}^{(0)}\times\left[\frac{1}{m\Omega}\mathbf{b}^{(0)}\times\left(\frac{y}{B^{(0)}}\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}+Ze\nabla\varphi^{(0)}\right)\right]\right\}-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)}\\
&= \frac{Ze}{c}\boldsymbol{\xi}_\perp\cdot\left\{-\frac{1}{m\Omega}B^{(0)}\left(\frac{y}{B^{(0)}}\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}+Ze\nabla\varphi^{(0)}\right)+\frac{1}{m\Omega}\mathbf{b}^{(0)}(\ldots.)\right\}-\mu_0B_\parallel^{(1)}-\\ &\quad Ze\phi^{(1)}\\
&= \frac{Ze}{c}\boldsymbol{\xi}_\perp\cdot\left\{-\frac{1}{m\Omega}B^{(0)}\left(\frac{y}{B^{(0)}}\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}+Ze\nabla\varphi^{(0)}\right)\right\}-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)}\\
&= \frac{Ze}{c}\boldsymbol{\xi}_\perp\cdot\left\{-\frac{c}{Ze}\big(\mu_0\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}+Ze\nabla\varphi^{(0)}\big)\right\}-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)}\\
&= -\boldsymbol{\xi}_\perp\cdot\big(\mu_0\nabla B^{(0)}+mv_\parallel^2\boldsymbol{\kappa}^{(0)}+Ze\nabla\varphi^{(0)}\big)-\mu_0B_\parallel^{(1)}-Ze\phi^{(1)}\\
&= -mv_\parallel^2\boldsymbol{\xi}_\perp\cdot\boldsymbol{\kappa}^{(0)}-\mu_0\left(\boldsymbol{\xi}_\perp\cdot\nabla B^{(0)}+B_\parallel^{(1)}\right)-Ze\big(\phi^{(1)}+\boldsymbol{\xi}_\perp\cdot\nabla\phi^{(0)}\big),\end{aligned} \tag{160}$$

Equation (160) agrees with Eq. (55) in Porcelli's paper[1]. The perturbed electrical field is

$$\mathbf{E}^{(1)}=-\frac{1}{c}\mathbf{u}^{(1)}\times\mathbf{B}^{(0)}=\frac{i\omega}{c}\boldsymbol{\xi}_\perp\times\mathbf{B}^{(0)}. \tag{161}$$

On the other hand, $\mathbf{E}^{(1)}$ can be expressed as

$$\begin{aligned}\mathbf{E}^{(1)} &= -\frac{1}{c}\frac{\partial\mathbf{A}^{(1)}}{\partial t}-\nabla\phi^{(1)}\\ &= -\frac{1}{c}\frac{\partial\big(\boldsymbol{\xi}_\perp\times\mathbf{B}^{(0)}\big)}{\partial t}-\nabla\phi^{(1)}\\ &= \frac{i\omega}{c}\boldsymbol{\xi}_\perp\times\mathbf{B}^{(0)}-\nabla\phi^{(1)}\end{aligned} \tag{162}$$

Comparing Eqs. (161) and (162), we obtain

$$\nabla\phi^{(1)}=0. \tag{163}$$

which indicates the perturbed scalar potential is a constant. We can choose $\phi^{(1)}=0$. We consider the case that there is no electrical field in the equilibrium, i.e., $\phi^{(0)}=0$. Using $\phi^{(0)}=0$ and $\phi^{(1)}=0$, the Lagrangian in Eq. (160) is written as

$$\mathcal{L}^{(1)}=-mv_\parallel^2\boldsymbol{\xi}_\perp\cdot\boldsymbol{\kappa}-\left(B_\parallel^{(1)}+\boldsymbol{\xi}_\perp\cdot\nabla B_0\right)\mu_0. \tag{164}$$

Substituting the expression of $B_{\parallel}^{(1)}$ in Eq. (157) into the above equation, we obtain

$$\begin{aligned}\mathcal{L}^{(1)} &= -mv_{\parallel}^2\boldsymbol{\xi}_{\perp}\cdot\boldsymbol{\kappa} - (-B_0(\nabla\cdot\boldsymbol{\xi}_{\perp}) - B_0\boldsymbol{\kappa}\cdot\boldsymbol{\xi}_{\perp} - \boldsymbol{\xi}_{\perp}\cdot\nabla B_0 + \boldsymbol{\xi}_{\perp}\cdot\nabla B_0)\mu_0 \\ &= -mv_{\parallel}^2\boldsymbol{\xi}_{\perp}\cdot\boldsymbol{\kappa} - (-B_0(\nabla\cdot\boldsymbol{\xi}_{\perp}) - B_0\boldsymbol{\kappa}\cdot\boldsymbol{\xi}_{\perp})\mu_0 \\ &= -\big(mv_{\parallel}^2 - \mu_0 B_0\big)\boldsymbol{\xi}_{\perp}\cdot\boldsymbol{\kappa} + \mu_0 B_0 \nabla\cdot\boldsymbol{\xi}_{\perp} \end{aligned} \tag{165}$$

We note an important fact that magnetic curvature $\boldsymbol{\kappa}$ is perpendicular to $\mathbf{b}$. This is because that $\boldsymbol{\kappa}\equiv\mathbf{b}\cdot\nabla\mathbf{b} = -\mathbf{b}\times\nabla\times\mathbf{b}$. The last equality is due to the vector identity $\nabla(\mathbf{A}\cdot\mathbf{B}) = \mathbf{A}\times\nabla\times\mathbf{B} + \mathbf{B}\times\nabla\times\mathbf{A} + \mathbf{B}\cdot\nabla\mathbf{A} + \mathbf{B}\cdot\nabla\mathbf{A}$. Using this fact, Eq. (165) can also be written as

$$\mathcal{L}^{(1)} = -\big(mv_{\parallel}^2 - \mu_0 B_0\big)\boldsymbol{\xi}\cdot\boldsymbol{\kappa} + \mu_0 B_0\nabla\cdot\boldsymbol{\xi}_{\perp}. \tag{166}$$

Eq. (166) agrees with Eq. (58) in Porcelli's paper[1].

4 Proof of equilavenc between Eq. (8) and Euler-Lagrange equation (168)

Equation (8) is repeated here, i.e.,

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}}\right) = \frac{\partial\mathcal{L}}{\partial\mathbf{X}}, \tag{167}$$

where $\partial/\partial\dot{\mathbf{X}}$ and $\partial/\partial\mathbf{X}$ are considered as gradient operators. Note that, in Cartisian coordinates, the components of Eq. (8) are obviously equivalent to the respective Euler-Lagrange equations. We now check whether the component equations in arbitrary coordinate system obtained by evaluating the gradient of the Lagrangian $\mathcal{L}$ in Eq. (8) are equivalent to the respective Euler-Lagrange equations in that coordinates system. The Euler-Lagrange equation in any coordinates is given by

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\right) = \frac{\partial\mathcal{L}}{\partial q_i}, \tag{168}$$

which is expressed in terms of a single coordinate component, and is coordinate independent, i..e, it takes the same form for every components of any coordinate systems.

In this section, I prove that the components of Eq. (8) are equivalent to the Euler-Lagrange equation (168) in cylindrical coordinates (R, Z, φ). First, let us consider the term $\partial\mathcal{L}/\partial\dot{\mathbf{X}}$, which, in Cartesian coordinators, is written as

$$\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}} = \frac{\partial\mathcal{L}}{\partial\dot{x}}\hat{\mathbf{e}}_x + \frac{\partial\mathcal{L}}{\partial\dot{y}}\hat{\mathbf{e}}_y + \frac{\partial\mathcal{L}}{\partial\dot{z}}\hat{\mathbf{e}}_z. \tag{169}$$

Using the chain rule, the above equation is written as

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}} = &\left(\frac{\partial\mathcal{L}}{\partial R}\frac{\partial R}{\partial\dot{x}} + \frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial\dot{x}} + \frac{\partial\mathcal{L}}{\partial Z}\frac{\partial Z}{\partial\dot{x}} + \frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{x}} + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{x}} + \frac{\partial\mathcal{L}}{\partial\dot{Z}}\frac{\partial\dot{Z}}{\partial\dot{x}}\right)\hat{\mathbf{e}}_x \\ + &\left(\frac{\partial\mathcal{L}}{\partial R}\frac{\partial R}{\partial\dot{y}} + \frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial\dot{y}} + \frac{\partial\mathcal{L}}{\partial Z}\frac{\partial Z}{\partial\dot{y}} + \frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{y}} + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{y}} + \frac{\partial\mathcal{L}}{\partial\dot{Z}}\frac{\partial\dot{Z}}{\partial\dot{y}}\right)\hat{\mathbf{e}}_y \\ + &\left(\frac{\partial\mathcal{L}}{\partial R}\frac{\partial R}{\partial\dot{z}} + \frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial\dot{z}} + \frac{\partial\mathcal{L}}{\partial Z}\frac{\partial Z}{\partial\dot{z}} + \frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{z}} + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{z}} + \frac{\partial\mathcal{L}}{\partial\dot{Z}}\frac{\partial\dot{Z}}{\partial\dot{z}}\right)\hat{\mathbf{e}}_z\end{aligned}$$

Using the transformation relation

$$\begin{cases} R = \sqrt{x^2+y^2} \\ \varphi = \mathrm{ArcSin}\left(\dfrac{y}{\sqrt{x^2+y^2}}\right) \\ Z = z \end{cases} \tag{170}$$

we obtain $\partial R/\partial\dot{x}=0$, $\partial R/\partial\dot{y}=0$, etc. Thus $\partial\mathcal{L}/\partial\dot{\mathbf{X}}$ reduces to

$$\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}}=\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{x}}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{x}}+\frac{\partial\mathcal{L}}{\partial\dot{Z}}\frac{\partial\dot{Z}}{\partial\dot{x}}\right)\hat{\mathbf{e}}_x+\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{y}}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{y}}+\frac{\partial\mathcal{L}}{\partial\dot{Z}}\frac{\partial\dot{Z}}{\partial\dot{y}}\right)\hat{\mathbf{e}}_y+\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{z}}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{z}}+\frac{\partial\mathcal{L}}{\partial\dot{Z}}\frac{\partial\dot{Z}}{\partial\dot{z}}\right)\hat{\mathbf{e}}_z \tag{171}$$

Using the transformation relation Eq. (170), we obtain

$$\dot{R}=\frac{x\,\dot{x}}{\sqrt{x^2+y^2}}+\frac{y\dot{y}}{\sqrt{x^2+y^2}}, \tag{172}$$

$$\dot{\varphi}=-\frac{y}{x^2+y^2}\dot{x}+\frac{x}{x^2+y^2}\dot{y}, \tag{173}$$

and

$$\dot{Z}=\dot{z} \tag{174}$$

Noting that $\partial\dot{Z}/\partial\dot{x}=0$, $\partial\dot{Z}/\partial\dot{y}=0$, $\partial\dot{R}/\partial\dot{z}=0$, $\partial\dot{\varphi}/\partial\dot{z}=0$, Eq. (171) is further written as

$$\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}}=\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{x}}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{x}}\right)\hat{\mathbf{e}}_x+\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial\dot{y}}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial\dot{y}}\right)\hat{\mathbf{e}}_y+\left(\frac{\partial\mathcal{L}}{\partial\dot{Z}}\right)\hat{\mathbf{e}}_z \tag{175}$$

Using Eqs. (172) and (173), we obtain

$$\frac{\partial\dot{R}}{\partial\dot{x}}=\frac{x}{R}=\cos\varphi \tag{176}$$

$$\frac{\partial\dot{R}}{\partial\dot{y}}=\frac{y}{R}=\sin\varphi \tag{177}$$

$$\frac{\partial\dot{\varphi}}{\partial\dot{x}}=-\frac{y}{R^2} \tag{178}$$

$$\frac{\partial\dot{\varphi}}{\partial\dot{y}}=\frac{x}{R^2} \tag{179}$$

Using these, Eq. (175) is written as

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}} &= \left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\cos\varphi-\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\sin\varphi\frac{1}{R}\right)\hat{\mathbf{e}}_x+\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\sin\varphi+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\cos\varphi\frac{1}{R}\right)\hat{\mathbf{e}}_y+\left(\frac{\partial\mathcal{L}}{\partial\dot{Z}}\right)\hat{\mathbf{e}}_z \\ &= \frac{\partial\mathcal{L}}{\partial\dot{R}}(\cos\varphi\hat{\mathbf{e}}_x+\sin\varphi\hat{\mathbf{e}}_y)+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{1}{R}(-\sin\varphi\hat{\mathbf{e}}_x+\cos\varphi\hat{\mathbf{e}}_y)+\left(\frac{\partial\mathcal{L}}{\partial\dot{Z}}\right)\hat{\mathbf{e}}_z \\ &= \frac{\partial\mathcal{L}}{\partial\dot{R}}\hat{\mathbf{e}}_R+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{1}{R}\hat{\mathbf{e}}_\varphi+\frac{\partial\mathcal{L}}{\partial\dot{Z}}\hat{\mathbf{e}}_z\end{aligned} \tag{180}$$

Using this, the left-hand side of Eq. (8) is written as

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{X}}}\right) &= \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\right)\hat{\mathbf{e}}_R+\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right)\frac{1}{R}\hat{\mathbf{e}}_\varphi+\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{Z}}\right)\hat{\mathbf{e}}_z+\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{d\hat{\mathbf{e}}_R}{dt}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{d}{dt}\left(\frac{1}{R}\hat{\mathbf{e}}_\varphi\right) \\ &= \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\right)\hat{\mathbf{e}}_R+\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right)\frac{1}{R}\hat{\mathbf{e}}_\varphi+\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{Z}}\right)\hat{\mathbf{e}}_z+\frac{\partial\mathcal{L}}{\partial\dot{R}}\dot{\varphi}\hat{\mathbf{e}}_\varphi+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left[-\dot{\varphi}\frac{1}{R}\hat{\mathbf{e}}_R-\hat{\mathbf{e}}_\varphi\frac{\dot{R}}{R^2}\right]\end{aligned} \tag{181}$$

Next, consider the right-hand side of Eq. (8), which is the space gradient of Lagrangian $\mathcal{L}$. When I at first considered this problem, I took it for granted that the space gradient in cylindrical coordinates should be

$$\frac{\partial\mathcal{L}}{\partial\mathbf{X}}=\frac{\partial\mathcal{L}}{\partial R}\hat{\mathbf{e}}_R+\frac{\partial\mathcal{L}}{\partial Z}\hat{\mathbf{e}}_Z+\frac{\partial\mathcal{L}}{\partial\varphi}\frac{1}{R}\hat{\mathbf{e}}_\varphi, \tag{182}$$

which turns out to be wrong because this formula does not take into account that $\mathcal{L}$ depends on $\dot{\mathbf{X}}$ which in turn depends on the spatial coordinates. The correct way to calculate the space gradient is as follows:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\mathbf{X}} &= \frac{\partial\mathcal{L}}{\partial x}\hat{\mathbf{e}}_x+\frac{\partial\mathcal{L}}{\partial y}\hat{\mathbf{e}}_y+\frac{\partial\mathcal{L}}{\partial z}\hat{\mathbf{e}}_z\\ &= \left[\frac{\partial\mathcal{L}}{\partial R}\frac{\partial R}{\partial x}+\frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial x}+\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial x}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial x}\right]\hat{\mathbf{e}}_x+\left[\frac{\partial\mathcal{L}}{\partial R}\frac{\partial R}{\partial y}+\frac{\partial\mathcal{L}}{\partial\varphi}\frac{\partial\varphi}{\partial y}+\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial y}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial y}\right]\hat{\mathbf{e}}_y+\\ &\quad\frac{\partial\mathcal{L}}{\partial Z}\hat{\mathbf{e}}_z\end{aligned} \tag{183}$$

It is important to note that $\mathcal{L}$ depends on $\dot{R}$, $\dot{\varphi}$ which in turn depend on x and y. As a result, there exist additional terms (the last two terms in both the brackets), which would be missed if we used the formula in Eq. (182). Equation (183) is further written as

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\mathbf{X}} &= \left[\frac{\partial\mathcal{L}}{\partial R}\cos\varphi+\frac{\partial\mathcal{L}}{\partial\varphi}\left(-\frac{y}{R^2}\right)+\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial x}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial x}\right]\hat{\mathbf{e}}_x+\left[\frac{\partial\mathcal{L}}{\partial R}\sin\varphi+\frac{\partial\mathcal{L}}{\partial\varphi}\left(\frac{x}{R^2}\right)+\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial y}+\right.\\ &\quad\left.\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial y}\right]\hat{\mathbf{e}}_y+\frac{\partial\mathcal{L}}{\partial Z}\hat{\mathbf{e}}_z\\ &= \frac{\partial\mathcal{L}}{\partial R}\hat{\mathbf{e}}_R+\frac{\partial\mathcal{L}}{\partial\varphi}\frac{1}{R}\hat{\mathbf{e}}_\varphi+\frac{\partial\mathcal{L}}{\partial Z}\hat{\mathbf{e}}_z+\left[\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial x}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial x}\right]\hat{\mathbf{e}}_x+\left[\frac{\partial\mathcal{L}}{\partial\dot{R}}\frac{\partial\dot{R}}{\partial y}+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\frac{\partial\dot{\varphi}}{\partial y}\right]\hat{\mathbf{e}}_y\end{aligned} \tag{184}$$

Using

$$\begin{aligned}\frac{\partial\dot{R}}{\partial x} &= \dot{x}\frac{y^2}{R^3}-y\dot{y}\frac{x}{R^3}\\ &= \frac{\dot{x}}{R}\sin^2\varphi-\frac{\dot{y}}{R}\sin\varphi\cos\varphi\\ &= -\dot{\varphi}\sin\varphi\end{aligned} \tag{185}$$

$$\begin{aligned}\frac{\partial\dot{R}}{\partial y} &= \dot{y}\frac{x^2}{R^3}-x\dot{x}\frac{y}{R^3}\\ &= \frac{\dot{y}}{R}\cos^2\varphi-\frac{\dot{x}}{R}\sin\varphi\cos\varphi\\ &= \dot{\varphi}\cos\varphi,\end{aligned} \tag{186}$$

$$\begin{aligned}\frac{\partial\dot{\varphi}}{\partial x} &= -y\dot{x}\left(-\frac{2x}{R^4}\right)+\dot{y}\left(\frac{1}{R^2}-\frac{2x^2}{R^4}\right)\\ &= \frac{2\dot{x}}{R^2}\cos\varphi\sin\varphi+\frac{\dot{y}}{R^2}(1-2\cos^2\varphi)\\ &= \frac{2\dot{x}}{R^2}\cos\varphi\sin\varphi+\frac{\dot{y}}{R^2}(\sin^2\varphi-\cos^2\varphi)\\ &= -\frac{\cos\varphi}{R}\dot{\varphi}+\frac{\sin\varphi}{R^2}\dot{R},\end{aligned} \tag{187}$$

$$\begin{aligned}\frac{\partial\dot{\varphi}}{\partial y} &= -\dot{x}\left(-\frac{2y^2}{R^4}+\frac{1}{R^2}\right)-x\dot{y}\frac{2y}{R^4}\\ &= -\frac{2\dot{y}}{R^2}\cos\varphi\sin\varphi+\frac{\dot{x}}{R^2}(2\sin^2\varphi-1)\\ &= -\frac{2\dot{y}}{R^2}\cos\varphi\sin\varphi+\frac{\dot{x}}{R^2}(\sin^2\varphi-\cos^2\varphi)\\ &= -\frac{\dot{y}}{R^2}\cos\varphi\sin\varphi-\frac{\dot{x}}{R^2}\cos^2\varphi-\frac{\dot{y}}{R^2}\cos\varphi\sin\varphi+\frac{\dot{x}}{R^2}\sin^2\varphi\\ &= -\frac{\cos\varphi}{R^2}\dot{R}-\frac{\sin\varphi}{R}\dot{\varphi}\end{aligned} \tag{188}$$

Using these results, the last two terms in Eq. (184) is written as

$$\begin{aligned}&\left[-\frac{\partial\mathcal{L}}{\partial\dot{R}}\dot{\varphi}\sin\varphi+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left(-\frac{\cos\varphi}{R}\dot{\varphi}+\frac{\sin\varphi}{R^2}\dot{R}\right)\right]\hat{\mathbf{e}}_x+\left[\frac{\partial\mathcal{L}}{\partial\dot{R}}\dot{\varphi}\cos\varphi+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left(-\frac{\cos\varphi}{R^2}\dot{R}-\frac{\sin\varphi}{R}\dot{\varphi}\right)\right]\hat{\mathbf{e}}_y\\ =&\frac{\partial\mathcal{L}}{\partial\dot{R}}\dot{\varphi}\hat{\mathbf{e}}_\varphi+\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left(-\frac{\dot{\varphi}}{R}\hat{\mathbf{e}}_R-\frac{\dot{R}}{R^2}\hat{\mathbf{e}}_\varphi\right).\end{aligned} \tag{189}$$

Using this, Eq. (184) is written as

$$\frac{\partial\mathcal{L}}{\partial\mathbf{X}} = \frac{\partial\mathcal{L}}{\partial R}\hat{\mathbf{e}}_R + \frac{\partial\mathcal{L}}{\partial Z}\hat{\mathbf{e}}_Z + \frac{\partial\mathcal{L}}{\partial\varphi}\frac{1}{R}\hat{\mathbf{e}}_\varphi + \frac{\partial\mathcal{L}}{\partial\dot{R}}\dot{\varphi}\hat{\mathbf{e}}_\varphi + \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\left[-\frac{\dot{\varphi}}{R}\hat{\mathbf{e}}_R - \frac{\dot{R}}{R^2}\hat{\mathbf{e}}_\varphi\right]. \tag{190}$$

Note that, compared with Eq. (182), the above expression contains additional terms. The additional terms are the source of confusion when I at first tried to prove the equivalence between Eqs. (8) and (168). (I was confused for many days before I fininaly found the solution given here.) Using Eqs. (190) and (181) in Eq. (167), we recover the Euler-Lagrange equation in cylindrical coordinates, i.e.,

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{R}}\right) = \frac{\partial\mathcal{L}}{\partial R}, \tag{191}$$

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\right) = \frac{\partial\mathcal{L}}{\partial\varphi}, \tag{192}$$

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{Z}}\right) = \frac{\partial\mathcal{L}}{\partial Z}. \tag{193}$$

5 tmp (wrong! do not read this)

On the other hand we have

$$\dot{\varepsilon} = \frac{dH}{dt},$$

which H is the Hamiltonian. We note that it is the energy expressed in terms of generalized coordinate and momentum that can be called Hamilton. Using the Hamilton's equation, it can be easily proved that

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}.$$

Further we note that

$$\frac{\partial H}{\partial t} + \frac{\partial\mathcal{L}}{\partial t} = \frac{\partial}{\partial t}\sum_i p_i\dot{q}_i. \tag{194}$$

In usual case, p_i and $\dot{q}_i$ are not an explicit function of time, thus we have

$$\frac{\partial H}{\partial t} + \frac{\partial\mathcal{L}}{\partial t} = 0. \tag{195}$$

Using these results, we have

$$\dot{\varepsilon} = -\frac{\partial L}{\partial t}. \tag{196}$$

Then using Eq. (35) we obtain

$$-\frac{\partial L}{\partial t} = -\left(\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t} + mv_\parallel\frac{\partial\mathbf{b}}{\partial t}\right)\cdot\dot{\mathbf{X}} + \mu\frac{\partial B}{\partial t} + Ze\frac{\partial\phi}{\partial t}. \tag{197}$$

$$\mathcal{L} = \left(\frac{Ze}{c}\mathbf{A} + mv_\parallel\mathbf{b}\right)\cdot\dot{\mathbf{X}} + \frac{1}{\Omega}y\dot{\alpha} - \frac{1}{2}mv_\parallel^2 - y - Ze\phi. \tag{198}$$

$$\frac{\partial L}{\partial t} = \left(\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t} + mv_\parallel\frac{\partial\mathbf{b}}{\partial t}\right)\cdot\dot{\mathbf{X}} - \frac{1}{\Omega^2}\frac{\partial\Omega}{\partial t}y\dot{\alpha} - Ze\frac{\partial\phi}{\partial t}$$

$$\frac{\partial L}{\partial t} = \left(\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t} + mv_\parallel\frac{\partial\mathbf{b}}{\partial t}\right)\cdot\dot{\mathbf{X}} - \frac{1}{B\Omega}\frac{\partial B}{\partial t}y\dot{\alpha} - Ze\frac{\partial\phi}{\partial t}$$

$$\frac{\partial L}{\partial t} = \left(\frac{Ze}{c}\frac{\partial\mathbf{A}}{\partial t} + mv_\parallel\frac{\partial\mathbf{b}}{\partial t}\right)\cdot\dot{\mathbf{X}} - \frac{1}{\Omega}\frac{\partial B}{\partial t}\mu\dot{\alpha} - Ze\frac{\partial\phi}{\partial t}$$

The first order equation of Eq. (197) is (noting that $\mathbf{A}^{(0)}$, $\mathbf{b}^{(0)}$, and $B^{(0)}$ is time independent, thus $\partial\mathbf{A}^{(0)}/\partial t$, $\partial\mathbf{b}^{(0)}/\partial t$, and $\partial B^{(0)}/\partial t$ are all zeros)

$$-\left(\frac{\partial L}{\partial t}\right)^{(1)} = -\left(\frac{Ze}{c}\frac{\partial\mathbf{A}^{(1)}}{\partial t} + mv_{\|}\frac{\partial\mathbf{b}^{(1)}}{\partial t}\right)\cdot\dot{\mathbf{X}}^{(0)} + \mu_0\frac{\partial B^{(1)}}{\partial t} + Ze\frac{\partial\phi^{(1)}}{\partial t}. \tag{199}$$

In writing the above expression, we have used the fact that $\partial/\partial\varphi$ here is taken by holding constant $(\dot{\mathbf{X}}, \dot{v}_{\|}, \dot{y}, \dot{\alpha}; \psi, \theta, v_{\|}, y, \alpha)$, instead of holding constantnt $(\dot{\psi}, \dot{\theta}, \dot{\varphi}, \dot{v}_{\|}, \dot{y}, \dot{\alpha}; \psi, \theta, v_{\|}, y, \alpha)$. In this case obviously $\partial\dot{\mathbf{R}}/\partial\varphi = 0$. If we calculate in the second case, then we would have $\partial\dot{\mathbf{R}}/\partial\psi \neq 0$, since

$$\dot{\mathbf{X}} = \frac{\partial\mathbf{X}}{\partial\psi}\dot{\psi} + \frac{\partial\mathbf{X}}{\partial\theta}\dot{\theta} + \frac{\partial\mathbf{X}}{\partial\varphi}\dot{\varphi}, \tag{200}$$

in which the terms such as $\partial\mathbf{X}/\partial\psi$ would explicitly contain φ. The second term on the right-hand side of Eq. (74) can be further calculated as

$$\begin{aligned} y\dot{\alpha}\frac{\partial}{\partial\varphi}\left(\frac{1}{\Omega}\right) &= y\dot{\alpha}\left(-\frac{1}{\Omega^2}\right)\frac{\partial\Omega}{\partial\psi} \\ &= y\dot{\alpha}\left(-\frac{1}{B\Omega}\right)\frac{\partial B}{\partial\psi}. \end{aligned} \tag{201}$$

Then we can use $y = \mu B$ and $\dot{\alpha} = \Omega$ in the above equation, yielding

$$y\dot{\alpha}\frac{\partial}{\partial\varphi}\left(\frac{1}{\Omega}\right) = -\mu\frac{\partial B}{\partial\psi}. \tag{202}$$

Here I have some important comments. First, we note that $\dot{\alpha} = \Omega$ is one of the components of the Euler-Lagrange equation, thus of course can not be substituted into the original Lagrangian $\mathcal{L}$ (if we do this, we can no longer use the resulting Lagrangian as a correct Lagrangian to obtain correct Euler-Lagrange equation). In contrast to this, it is obvious we can use one component of the Euler-Lagrangian equation in another component equation. Thus we can substitute $\dot{\alpha} = \Omega$ into Eq. (201) to get Eq. (202). Second, we also substitute $y = \mu B$ into Eq. (201) in obtaining Eq. (202). This is trivial since what we do is only to rewrite the final result in a different form. However this kind of rewriting may be misleading to someone (including me) because the new form can be viewed as being written in terms of a new variable μ, instead of the original variable y. Of course, for this case, no matter which variable the right-hand side of Eq. (202) is understood to be written in terms of, the results are both correct. But it is crucial to understand correctly which variables Lagrangian $\mathcal{L}$ is written in terms of, since different choice of variables will give different forms of perturbed Lagrangian because the perturbed Lagrangian is obtained by keeping the independent variables constant.

Now I calculate the perturbed Lagrangian. The full Lagrangian is given by Eq. (1), i.e.,

$$\mathcal{L} = \left(\frac{Ze}{c}\mathbf{A} + mv_{\|}\mathbf{b}\right)\cdot\dot{\mathbf{X}} + \frac{1}{\Omega}y\dot{\alpha} - \frac{1}{2}mv_{\|}^2 - y - Ze\phi.$$

Then the perturbed and linearized version is (note that only the electromagnetic field is perturbed, the independent variables are keep constant)

$$\mathcal{L}^{(1)} = \left(\frac{Ze}{c}\mathbf{A}^{(1)} + mv_{\|}\mathbf{b}^{(1)}\right)\cdot\dot{\mathbf{X}} + \left(\frac{1}{\Omega}\right)^{(1)}y\dot{\alpha} - Ze\phi^{(1)}. \tag{203}$$

Using

$$\left(\frac{1}{B}\right)^{(1)} = -\frac{1}{B_0^2}B^{(1)}, \tag{204}$$

in Eq. (203), we obtain

$$\mathcal{L}^{(1)} = \left(\frac{Ze}{c}\mathbf{A}^{(1)} + mv_{\|}\mathbf{b}^{(1)}\right)\cdot\dot{\mathbf{X}} - \frac{mc}{Ze}\frac{1}{B_0^2}B^{(1)}y\dot{\alpha} - Ze\phi^{(1)}.$$

Using $y = \mu B$ and $\dot{\alpha} = \Omega$ in the above equation, we obtain

$$\mathcal{L}^{(1)} = \left(\frac{Ze}{c}\mathbf{A}^{(1)} + m v_{\parallel}\mathbf{b}^{(1)}\right) \cdot \dot{\mathbf{X}} - \mu B^{(1)} - Ze\phi^{(1)}. \tag{205}$$

My question is whether it is valid to substitute one of the Euler-Lagrangian equation $\dot{\alpha} = \Omega$ into the perturbed Lagrangian.**wrong!!**

Bibliography

[1] F. Porcelli, R. Stankiewicz, W. Kerner, and H. L. Berk. Solution of the drift-kinetic equation for global plasma modes and finite particle orbit widths. *Physics of Plasmas*, 1(3):470–480, 1994.

[2] Allen H. Boozer. Physics of magnetically confined plasmas. *Rev. Mod. Phys.*, 76:1071–1141, Jan 2005.

[3] Philipp Lauber. *Linear gyrokinetic descritpion of fast particle effects on the MHD stability in Tokamaks*. PhD thesis, Max-Planck institute for plasma physics, 2003.

[4] Y. Todo. Properties of energetic-particle continuum modes destabilized by energetic ions with beam-like velocity distributions. *Physics of Plasmas (1994-present)*, 13(8):–, 2006.

[5] John Wesson. *Tokamaks*. Oxford University Press, 2004.