# 1 Drift kinetic equation in $(\varepsilon, \mu, \sigma)$ coordinates

Kinetic equation is written as

$$\frac{df}{dt} = C(f) \tag{1}$$

In $(\boldsymbol{X}, \varepsilon, \mu, \sigma)$ coordinates, the above equation is written as

$$\frac{df}{dt} \equiv \frac{\partial f}{\partial t} + \dot{\boldsymbol{X}} \cdot \nabla f + \dot{\varepsilon}\frac{\partial f}{\partial \varepsilon} + \dot{\mu}\frac{\partial f}{\partial \mu} + \dot{\sigma}\frac{\partial f}{\partial \sigma} = C(f) \tag{2}$$

where $f = f(\boldsymbol{X}, \varepsilon, \mu, \sigma)$, and the space gradient $\nabla$ is taken by holding $\varepsilon$, $\mu$, and $\sigma$ constant. Using the fact that $\dot{\mu} = 0$ and $\dot{\varepsilon} = 0$ (question here: energy is generally not conserved, why do we assume $\dot{\varepsilon} = 0$?), the above equation is written as

$$\frac{\partial f}{\partial t} + \dot{\boldsymbol{X}} \cdot \nabla f + \dot{\sigma}\frac{\partial f}{\partial \sigma} = C(f). \tag{3}$$

Now we consider the last term on the left-hand side of the above equation. For passing particles, since $v_\parallel$ does not change sign, we have $\dot{\sigma} = 0$. Thus the last term for passing particles drops. For trapped particles, $\dot{\sigma} \neq 0$. In the transport problem, the time scale (transport time scale) we are interested in is much larger than the bounce period of trapped particles. Therefore, in the transport time scale, the value of $\sigma$ of trapped particles will have almost equal time to be $+1$ and $-1$. This results in that the distribution function $f$ should be equal for $\sigma = +1$ and $\sigma = -1$, i.e., $f$ is actually independent of $\sigma$, i.e., $\partial f / \partial \sigma = 0$ for trapped electrons. Using these results, we obtain that

$$\dot{\sigma}\frac{\partial f}{\partial \sigma} = 0,$$

for both passing and trapped particles. Therefore Eq. (3) is finally written as

$$\frac{\partial f}{\partial t} + \dot{\boldsymbol{X}} \cdot \nabla f = C(f), \tag{4}$$

which agrees equation (8.8) in White's book[1].

# 2 Parallel current density

The parallel current density is defined by

$$j_\parallel = q \int v_\parallel f(\boldsymbol{X}, \varepsilon, \mu, \sigma) d\Gamma,$$

where $d\Gamma$ is the volume element of velocity space

$$d\Gamma = \frac{2\pi}{m^2} \sum_{\sigma=\pm 1} \frac{B(\boldsymbol{r})}{|v_\parallel|} d\mu d\varepsilon. \tag{5}$$

Using Eq. (5), the parallel current density is written as

$$\begin{aligned} j_\parallel &= q\frac{2\pi}{m^2} \sum_{\sigma=\pm 1} \int v_\parallel f(\boldsymbol{X}, \varepsilon, \mu, \sigma) \frac{B}{|v_\parallel|} d\mu d\varepsilon. \\ &= q\frac{2\pi}{m^2} \sum_{\sigma=\pm 1} \int \sigma f(\boldsymbol{X}, \varepsilon, \mu, \sigma) B d\mu d\varepsilon. \\ &= q\frac{2\pi}{m^2} \sum_{\sigma=\pm 1} \sigma \int f(\boldsymbol{X}, \varepsilon, \mu, \sigma) B d\mu d\varepsilon. \end{aligned} \tag{6}$$

Now we divide the velocity space ($\varepsilon$, $\mu$, $\sigma$) into passing region (PR) and trapped region (TR), then Eq. (6) can be written as

$$j_{\|}=q\frac{2\pi}{m^2}\left[\sum_{\sigma=\pm 1}\sigma\int_{\mathrm{PR}}f(\boldsymbol{X},\varepsilon,\mu,\sigma)Bd\mu d\varepsilon+\sum_{\sigma=\pm 1}\sigma\int_{\mathrm{TR}}f(\boldsymbol{X},\varepsilon,\mu,\sigma)Bd\mu d\varepsilon\right] \tag{7}$$

In the trapped region, $f$ is indepndent of $\sigma$, thus the integration in the trapped region is written as

$$\begin{aligned}\sum_{\sigma=\pm 1}\sigma\int_{\mathrm{TR}}f(\boldsymbol{X},\varepsilon,\mu,\sigma)Bd\mu d\varepsilon &= +1\int_{\mathrm{TR}}f(\boldsymbol{X},\varepsilon,\mu)Bd\mu d\varepsilon-1\int_{\mathrm{TR}}f(\boldsymbol{X},\varepsilon,\mu)Bd\mu d\varepsilon\\ &= 0.\end{aligned}$$

i.e. the trapped particles does not contribute to the parallel current density. The parallel current density is written as

$$j_{\|}=q\frac{2\pi}{m^2}\sum_{\sigma=\pm 1}\sigma\int_{\mathrm{PR}}f(\boldsymbol{X},\varepsilon,\mu,\sigma)Bd\mu d\varepsilon. \tag{8}$$

## 3 The forms of drift kinetic equation in $(\boldsymbol{X},\ \boldsymbol{v},\ \boldsymbol{\theta})$ coordinates

Neglecting the perpendicular drift terms, the steady drift kinetic equation (DKE) in terms of ($\varepsilon$, $\mu,\sigma$) variables is written as

$$v_{\|}\hat{\boldsymbol{b}}\cdot\nabla f(\varepsilon,\mu,\sigma,\boldsymbol{r},t)=C(f)+Q(f) \tag{9}$$

where $f=f(\varepsilon,\mu,\sigma,\boldsymbol{r})$ is the distribution function of guiding-center variables (note that $\boldsymbol{r}$ is the location of the guiding-center), $\hat{\boldsymbol{b}}$ is a unit vector along the equilibrium magnetic field, $v_{\|}\equiv\boldsymbol{v}\cdot\hat{\boldsymbol{b}}=v_{\|}(\varepsilon,\ \mu,\ \sigma,\ \boldsymbol{r})$, $\sigma=\mathrm{sgn}(v_{\|})=\pm 1$, The $\nabla$ operates on the 4th variable (i.e., guiding-center location) of the distribution function $f(\varepsilon,\ \mu,\sigma,\boldsymbol{r},t)$; $C(f)$ and $Q(f)$ denotes, respectively, collision terms and wave-induced diffusion in velocity space. Eq. (9) can be further written as

$$v_{\|}\frac{\partial f(\varepsilon,\mu,\sigma,\rho,l)}{\partial l}=C(f)+Q(f) \tag{10}$$

This form of DKE is expressed in the ($\varepsilon$, $\mu$, $\sigma$, $\boldsymbol{r}$) coordinates (i.e., guiding-center coordinates). We now want to derive its form in ($v$, $\theta$, $\boldsymbol{r}$) coordinates (i.e. using particle's $v$ and $\theta$, instead of $\varepsilon$, $\mu$, and $\sigma$), where $\theta$ is the pitch angle in velocity space, which is the included angle between particles' velocity and the equilibrium magnetic field. (Note that $(v,\theta,\boldsymbol{r})$ are not purely particle variables, since here $\boldsymbol{r}$ is still a guiding-center variable, the location of the guiding-center.)

I use $g(v,\theta,l)$ to denote the distribution function in $(v,\theta,l)$ coordinates. Then the two functions, $f$ and $g$, are related to each other by

$$f(\varepsilon,\mu,\sigma,l)=g\left(\sqrt{2\varepsilon/m},\arccos^{-1}\left(\sigma\sqrt{1-\frac{B(l)\mu}{\varepsilon}}\right),l\right). \tag{11}$$

Then, what is the equation that $g(v,\ \theta,\ l)$ must satisfy? Now I derive this equation. Using Eq. (11) in Eq. (10) gives

$$v_{\|}\frac{\partial}{\partial l}g\left(\sqrt{2\varepsilon/m},\arccos^{-1}\left(\sigma\sqrt{1-\frac{B(l)\mu}{\varepsilon}}\right),l\right)=C(f)+Q(f) \tag{12}$$

$$\Rightarrow v_{\|}\left\{\frac{\partial g(v,\theta,l)}{\partial l}+\frac{\partial g(v,\theta,l)}{\partial\theta}\frac{d}{dl}\left[\arccos^{-1}\left(\sigma\sqrt{1-\frac{B(l)\mu}{\varepsilon}}\right)\right]\right\}=C(f)+Q(f) \tag{13}$$

The factor of the second term can be further written as

$$
\begin{aligned}
\frac{d}{dl}\left[\arccos^{-1}\left(\sigma\sqrt{1-\frac{B(l)\mu}{\varepsilon}}\right)\right] &= \left(-\frac{1}{\sqrt{1-v_{\|}^2/v^2}}\right)\sigma\frac{-\frac{\mu}{\varepsilon}\frac{dB(l)}{dl}}{2\sqrt{1-\frac{B\mu}{\varepsilon}}}\\
&= \left(-\frac{1}{\sqrt{1-v_{\|}^2/v^2}}\right)\frac{-\frac{\mu}{\varepsilon}\frac{dB(l)}{dl}}{2\sigma\sqrt{1-\frac{B\mu}{\varepsilon}}}\\
&= \left(-\frac{1}{\sin\theta}\right)\frac{-\frac{\sin^2\theta}{B}\frac{dB(l)}{dl}}{2\cos\theta}\\
&= \frac{\frac{\sin\theta}{B}\frac{dB(l)}{dl}}{2\cos\theta}\\
&= \frac{1}{2}\frac{v_{\perp}}{v_{\|}}\frac{d\ln B(l)}{dl}
\end{aligned}
\tag{14}
$$

Using Eq. (14) in Eq. (13), we obtain

$$
v_{\|}\frac{\partial g(v,\theta,l)}{\partial l}+\frac{1}{2}v_{\perp}\frac{d\ln B(l)}{dl}\frac{\partial g(v,\theta,l)}{\partial\theta}=C(f)+Q(f)
\tag{15}
$$

Eq. (15) agrees with equation in Westerhof's paper[2]. Note this equation contains additional term which does not appear in the original DKE in ($\varepsilon$, $\mu$, $\sigma$, $l$) coordinates. We can further rewrite Eq. (15) as

$$
v\cos\theta\,\hat{\boldsymbol{b}}\cdot\nabla g(v,\theta,\boldsymbol{r})-\frac{1}{2}v\sin\theta\nabla\cdot\left(\hat{\boldsymbol{b}}\right)\frac{\partial g(v,\theta,\boldsymbol{r})}{\partial\theta}=C(f)+Q(f)
\tag{16}
$$

In obtaining Eq. (16), use has been made of

$$
\nabla\cdot\left(\hat{\boldsymbol{b}}\right)=\nabla\cdot\left(\frac{\boldsymbol{B}}{B}\right)=-\frac{d\ln B(l)}{dl}.
\tag{17}
$$

Eq. (16) agrees with Eq. (5) in Kerbel's paper[3].

[Now we derive the DKE in $(v_{\perp},v_{\|},l)$ coordinates. We use $h(v_{\perp},v_{\|},l)$ to denote distribution function in $(v_{\perp},v_{\|},l)$ coordinates. Then we have the relation

$$
g(v,\theta,l)=h(v\sin\theta,v\cos\theta,l)
$$

Then the $\partial g(v,\theta,l)/\partial\theta$ term in Eq.(15) can be written as

$$
\begin{aligned}
\frac{\partial g(v,\theta,l)}{\partial\theta} &= \frac{\partial h(v\sin\theta,v\cos\theta,l)}{\partial\theta}\\
&= h_1 v\cos\theta-h_2\,v\sin\theta
\end{aligned}
\tag{18}
$$

Using Eq.(18) in Eq.(15) gives

$$
-v_{\|}\frac{\partial h(v_{\perp},v_{\|},l)}{\partial l}-\frac{1}{2}v_{\perp}\frac{d\ln B(l)}{dl}\left[\frac{\partial h(v_{\perp},v_{\|},l)}{\partial v_{\perp}}v_{\|}-\frac{\partial h(v_{\perp},v_{\|},l)}{\partial v_{\|}}v_{\perp}\right]=C(f)+Q(f)
\tag{19}
$$

This is the DKE in $(v_{\perp},v_{\|},l)$ coordinates.]

# 4 Use Eq. (16) to prove $\nabla\cdot\left(\boldsymbol{j}_{\|}\hat{\boldsymbol{b}}\right)=0$

Now I use Eq. (16)

$$
v\cos\theta\,\hat{\boldsymbol{b}}\cdot\nabla g(v,\theta,\boldsymbol{r})-\frac{1}{2}v\sin\theta\nabla\cdot\left(\hat{\boldsymbol{b}}\right)\frac{\partial g(v,\theta,\boldsymbol{r})}{\partial\theta}=C(f)+Q(f)
$$

to prove that

$$\nabla\cdot\left(\hat{\boldsymbol{b}}j_{\|}\right)=0 \tag{20}$$

where

$$j_{\|}=\int d\Gamma v_{\|}g(v,\theta,l) \tag{21}$$

and $d\Gamma$ is the volume element in velocity space.

Proof: Integrating Eq. (16) in velocity space, we obtain

$$\hat{\boldsymbol{b}}\cdot\nabla j_{\|}-\nabla\cdot\left(\hat{\boldsymbol{b}}\right)\frac{1}{2}\int d\Gamma v\sin\theta\frac{\partial g(v,\theta,\boldsymbol{r})}{\partial\theta}=0, \tag{22}$$

where use has been made of the fact that collision term conserves particle number, i.e., $\int d\Gamma C(f)=0$, and quasilinear diffusion operator also conserves particle number, i.e., $\int d\Gamma Q(f)=0$. The second term of Eq. (22) can be written as

$$\begin{aligned}\int d\Gamma v\sin\theta\frac{\partial g(v,\theta,\boldsymbol{r})}{\partial\theta} &= \int d\Gamma v\left(\frac{\partial g\sin\theta}{\partial\theta}-\cos\theta g\right)\\ &= -j_{\|}+\int d\Gamma v\frac{\partial g\sin\theta}{\partial\theta}\\ &= -j_{\|}+2\pi\int v^3\,dv\int_0^{\pi}\sin\theta\frac{\partial g\sin\theta}{\partial\theta}d\theta\end{aligned} \tag{23}$$

Integrating by parts, the above equation is written as

$$\begin{aligned}\int d\Gamma v\sin\theta\frac{\partial g(v,\theta,\boldsymbol{r})}{\partial\theta} &= -j_{\|}-2\pi\int v^3dv\int_0^{\pi}g\sin\theta\cos\theta d\theta\\ &= -j_{\|}-2\pi\int v^2dv\int_0^{\pi}gv_{\|}\sin\theta d\theta\\ &= -j_{\|}-j_{\|}\\ &= -2j_{\|}\end{aligned}$$

Using this in Eq. (22) gives

$$\hat{\boldsymbol{b}}\cdot\nabla j_{\|}+\nabla\cdot\left(\hat{\boldsymbol{b}}\right)j_{\|}=0 \tag{24}$$

$$\Longrightarrow\nabla\cdot\left(\hat{\boldsymbol{b}}j_{\|}\right)=0 \tag{25}$$

## 5 Use Eq. (9) to prove $\boldsymbol{\nabla}\cdot\left(\boldsymbol{j}_{\|}\hat{\boldsymbol{b}}\right)=\mathbf{0}$

If the distribution function $f$ satisfies the steady state drift kinetic equation

$$v_{\|}\hat{\boldsymbol{b}}\cdot\nabla f=C(f)+Q(f), \tag{26}$$

where $v_{\|}=v_{\|}(\varepsilon,\mu,\sigma,\boldsymbol{r})$, $f=f(\varepsilon,\mu,\sigma,\boldsymbol{r})$ and $\nabla$ operates on the 4th variable of $(\varepsilon,\mu,\sigma,\boldsymbol{r})$, it can be proved that

$$\nabla\cdot\left(j_{\|}\hat{\boldsymbol{b}}\right)=0 \tag{27}$$

where

$$j_{\|}=\int d\Gamma fv_{\|}, \tag{28}$$

with $d\Gamma$ the velocity space volume element.

Proof: Integrating Eq. (26) in velocity space, we obtain

$$\int d\Gamma v_{\|}\hat{\boldsymbol{b}}\cdot\nabla f=\int d\Gamma C(f)+\int d\Gamma Q(f) \tag{29}$$

Using the fact that collision term conserves particle number, $\int d\Gamma C(f)=0$, and quasilinear diffusion operator also conserves particle number, $\int d\Gamma Q(f)=0$, Eq. (29) is reduced to

$$\int d\Gamma v_{\|}\hat{\boldsymbol{b}}\cdot\nabla f(\varepsilon,\mu,\sigma,\boldsymbol{r})=0 \tag{30}$$

Using Jacobian, one can get the velocity element $d\Gamma$ in $(\varepsilon,\mu,\sigma,\boldsymbol{r})$ coordinates (Refer to another note)

$$d\Gamma=\frac{2\pi}{m^2}\sum_{\sigma=\pm1}\frac{B(\boldsymbol{r})}{|v_{\|}|}d\mu d\varepsilon \tag{31}$$

Using this in Eq. (30), we obtain

$$\frac{2\pi}{m^2}\sum_{\sigma=\pm1}\int\frac{B}{|v_{\|}|}d\mu d\varepsilon v_{\|}\hat{\boldsymbol{b}}\cdot\nabla f=0 \tag{32}$$

$$\Longrightarrow\frac{2\pi}{m^2}\sum_{\sigma=\pm1}\int d\mu d\varepsilon\sigma\boldsymbol{B}\cdot\nabla f=0 \tag{33}$$

where $\sigma=\operatorname{sgn}(v_{\|})$.

$$\Longrightarrow\frac{2\pi}{m^2}\sum_{\sigma=\pm1}\int d\mu d\varepsilon\sigma[\nabla\cdot(\boldsymbol{B}f)-f\nabla\cdot\boldsymbol{B}]=0 \tag{34}$$

Using $\nabla\cdot\boldsymbol{B}=\boldsymbol{0}$, Eq. (34) is reduced to

$$\frac{2\pi}{m^2}\sum_{\sigma=\pm1}\int d\mu d\varepsilon\sigma\nabla\cdot[\boldsymbol{B}f(\varepsilon,\mu,\sigma,\boldsymbol{r})]=0 \tag{35}$$

Note the $\nabla$ operates on the 4th variable of $(\varepsilon,\mu,\sigma,\boldsymbol{r})$. Thus the $d\mu d\varepsilon$ integration along with the summation over $\sigma$ can be exchanged with the $\nabla$ operator to give

$$\frac{2\pi}{m^2}\nabla\cdot\left[\sum_{\sigma=\pm1}\int d\mu d\varepsilon\sigma B\hat{\boldsymbol{b}}f(\varepsilon,\mu,\sigma,\boldsymbol{r})\right]=0 \tag{36}$$

Noting that $\hat{\boldsymbol{b}}(\boldsymbol{r})$ is independent of $\varepsilon$, $\mu$, and $\sigma$ in the ($\varepsilon$, $\mu$, $\sigma$, $\boldsymbol{r}$) coordinates, $\hat{\boldsymbol{b}}$ thus can be taken out of the $d\mu d\varepsilon$ integration and the summation over $\sigma$ to give

$$\frac{2\pi}{m^2}\nabla\cdot\left[\hat{\boldsymbol{b}}\sum_{\sigma=\pm1}\int d\mu d\varepsilon\sigma Bf(\varepsilon,\mu,\boldsymbol{r})\right]=0 \tag{37}$$

Using Eq. (31) again in Eq. (37), we obtain

$$\frac{2\pi}{m^2}\nabla\cdot\left[\hat{\boldsymbol{b}}\int d\Gamma\left(\frac{m^2}{2\pi}\frac{|v_{\|}|}{B}\right)\sigma Bf(\varepsilon,\mu,\boldsymbol{r})\right]=0 \tag{38}$$

$$\Longrightarrow\nabla\cdot\left[\hat{\boldsymbol{b}}\int d\Gamma v_{\|}f(\varepsilon,\mu,\boldsymbol{r})\right]=0 \tag{39}$$

Therefore we finally obtain

$$\nabla\cdot\left(j_{\|}\hat{\boldsymbol{b}}\right)=0. \tag{40}$$

Using the above equation, we can further prove that $j_{\|}/B$ is a function of the flux surface. The proof is as follows. Eq. (40) is written as

$$\nabla\cdot\left(\frac{j_{\|}}{B}\boldsymbol{B}\right)=0.$$

$$\Rightarrow\boldsymbol{B}\cdot\nabla\frac{j_{\|}}{B}+\frac{j_{\|}}{B}\nabla\cdot\boldsymbol{B}=0$$

$$\Rightarrow\boldsymbol{B}\cdot\nabla\frac{j_{\|}}{B}=0,$$

which indicates $j_{\|}/B$ is a function of the flux surface.

## 6 To prove $\left\langle \int d\Gamma f v_{\parallel} \hat{\boldsymbol{b}} \cdot \nabla g \right\rangle = -\left\langle \int d\Gamma g v_{\parallel} \hat{\boldsymbol{b}} \cdot \nabla f \right\rangle$

Try to prove that

$$\left\langle \int d\Gamma f v_{\parallel} \hat{b} \cdot \nabla g \right\rangle = -\left\langle \int d\Gamma g v_{\parallel} \hat{b} \cdot \nabla f \right\rangle \tag{41}$$

where $f$ and $g$ are arbitray distribution functions, $\langle ... \rangle$ is the flux average operator. Eq. (41) is equivalent to

$$\left\langle \int d\Gamma \left[ f v_{\parallel} \hat{b} \cdot \nabla g + g v_{\parallel} \hat{b} \cdot \nabla f \right] \right\rangle = 0 \tag{42}$$

$$\Longleftrightarrow \left\langle \int d\Gamma v_{\parallel} \hat{b} \cdot [f \nabla g + g \nabla f] \right\rangle = 0 \tag{43}$$

$$\Longleftrightarrow \left\langle \int d\Gamma v_{\parallel} \hat{b} \cdot \nabla (fg) \right\rangle = 0 \tag{44}$$

Using the same method in Sec. 5, Eq. (44) can be reduced to

$$\Longleftrightarrow \left\langle \nabla \cdot \left( \hat{\boldsymbol{b}} \int d\Gamma v_{\parallel} f g \right) \right\rangle = 0 \tag{45}$$

Using the definition of the flux average, the left-hand side of the above equation is written as

$$\left\langle \nabla \cdot \left( \hat{\boldsymbol{b}} \int d\Gamma v_{\parallel} f g \right) \right\rangle = \frac{\int_V \nabla \cdot \left( \hat{\boldsymbol{b}} \int d\Gamma g f v_{\parallel} \right) dV}{\int_V dV} \tag{46}$$

Using the Gauss theorem to transformt the above volume integration to surface integration, we obtain

$$\begin{aligned} \left\langle \nabla \cdot \left( \hat{\boldsymbol{b}} \int d\Gamma v_{\parallel} f g \right) \right\rangle &= \frac{\int_S \left( \hat{\boldsymbol{b}} \int d\Gamma g f v_{\parallel} \right) \cdot d\boldsymbol{S}}{\int_V dV} \\ &= 0 \end{aligned}$$

The last equality is due to that $\hat{\boldsymbol{b}} \cdot d\boldsymbol{S} = 0$ since $d\boldsymbol{S}$ is along the direction of $\nabla \psi$, which is perpendicular to $\hat{\boldsymbol{b}}$. Therefore Eq. (45) is proved.

Some remarks on the results obtained are in order. In the above section, we prove that

$$\left\langle \int d\Gamma f v_{\parallel} \hat{b} \cdot \nabla g \right\rangle = -\left\langle \int d\Gamma g v_{\parallel} \hat{b} \cdot \nabla f \right\rangle \tag{47}$$

If the flux averaging is removed, we generally has

$$\int d\Gamma f v_{\parallel} \hat{b} \cdot \nabla g \neq -\int d\Gamma g v_{\parallel} \hat{b} \cdot \nabla f \tag{48}$$

This indicates the flux averaging is crucial to the success of the adjoint method in solving the current drive equation in toroidal geometry, since, without the flux averaging, we can not easily obtain the adjoint operator of $v_{\parallel} \hat{\boldsymbol{b}} \cdot \nabla$.

## 7 Note on bounce-averaged drift kinetic equation

The drift kinetic equation (DKE) neglecting cross-field drift takes the form

$$\frac{\partial f(w, \mu, \sigma, l, t)}{\partial t} + v_{\parallel} \frac{\partial f(w, \mu, \sigma, l, t)}{\partial l} + q E_{\parallel} v_{\parallel} \frac{\partial f(w, \mu, \sigma, l, t)}{\partial w} - C(f) = 0 \tag{49}$$

where $f = f(w, \mu, \sigma, l, t)$, $\sigma = \pm 1$, and $v_\parallel = \sigma\sqrt{\frac{2}{m}[w - \mu B(l)]}$. Now we derive the bounce-averaged form of the DKE.

If Eq. (49) describes a physical process at frequency $\omega_l$, then we have

$$\frac{\partial}{\partial t} \sim \omega_l \tag{50}$$

and we further assume that

$$v_\parallel \frac{\partial}{\partial l} \sim \omega_h \tag{51}$$

and $\omega_h \gg \omega_l$. We expand $f$ as $f = f_0 + f_1$, where $f_i \sim O[(\omega_l/\omega_h)^i]$. Using this expansion in Eq. (49), we get

$$\frac{\partial f_0}{\partial t} + \frac{\partial f_1}{\partial t} + v_\parallel \frac{\partial f_0}{\partial l} + v_\parallel \frac{\partial f_1}{\partial l} + qE_\parallel v_\parallel \frac{\partial f_0}{\partial w} + qE_\parallel v_\parallel \frac{\partial f_1}{\partial w} - C(f_0 + f_1) = 0 \tag{52}$$

Dividing both sides of the above equation by $\omega_h$,

$$\frac{\partial f_0}{\partial t}\frac{1}{\omega_h} + \frac{\partial f_1}{\partial t}\frac{1}{\omega_h} + v_\parallel \frac{\partial f_0}{\partial l}\frac{1}{\omega_h} + v_\parallel \frac{\partial f_1}{\partial l}\frac{1}{\omega_h} + qE_\parallel v_\parallel \frac{\partial f_0}{\partial w}\frac{1}{\omega_h} + qE_\parallel v_\parallel \frac{\partial f_1}{\partial w}\frac{1}{\omega_h} - C(f_0)\frac{1}{\omega_h} - C(f_1)\frac{1}{\omega_h} = 0 \tag{53}$$

Now we use the above assumptions to separate Eq.(53) into several equations according to the orders of the various terms. We get the order of the terms in Eq.(53) as

$$\begin{aligned}
&\frac{\partial f_0}{\partial t}\frac{1}{\omega_h} \sim 1 \cdot \omega_l/\omega_h \\
&\frac{\partial f_1}{\partial t}\frac{1}{\omega_h} \sim (\omega_l/\omega_h)^2 \\
&v_\parallel \frac{\partial f_0}{\partial l}\frac{1}{\omega_h} \sim 1 \\
&v_\parallel \frac{\partial f_1}{\partial l}\frac{1}{\omega_h} \sim (\omega_l/\omega_h)^1 \\
&qE_\parallel v_\parallel \frac{\partial f_0}{\partial w}\frac{1}{\omega_h} \sim (\omega_l/\omega_h)^1 && (54)\\
&qE_\parallel v_\parallel \frac{\partial f_1}{\partial w}\frac{1}{\omega_h} \sim (\omega_l/\omega_h)^2 && (55)\\
&C(f_0)\frac{1}{\omega_h} \sim \omega_l/\omega_h \\
&C(f_1)\frac{1}{\omega_h} \sim (\omega_l/\omega_h)^2
\end{aligned}$$

In obtaining Eqs.(54)(55), $qE_\parallel v_\parallel \partial/\partial w \sim \omega_l$ and $C(f) \sim f\omega_l$ is assumed.

Then the terms at the order of $(\omega_l/\omega_h)^0$ is separated to form a equation

$$v_\parallel \frac{\partial f_0}{\partial l} = 0 \tag{56}$$

and the terms at the order of $(\omega_l/\omega_h)^1$ is separated to form another equation

$$\frac{\partial f_0}{\partial t} + v_\parallel \frac{\partial f_1}{\partial l} + qE_\parallel v_\parallel \frac{\partial f_0}{\partial w} - C(f_0) = 0 \tag{57}$$

The terms at the order of $(\omega_l/\omega_h)^2$ form the third equation

$$\frac{\partial f_1}{\partial t} + qE_\parallel v_\parallel \frac{\partial f_1}{\partial w} - C(f_1) = 0 \tag{58}$$

Eq. (56) indicates $f_0$ is independent of $l$, i.e., $f_0 = f_0(w, \mu, \sigma, t)$. Eq. (57) determines how $f_0$ evolves with time. Note that there is a term involving $f_1$ in Eq. (57). Is it possible to eliminate this term so that the time development of $f_0$ can be determined? The answer to this question is yes. Before we give the details, we note that $f_0$ in the trapped region should be independent of $\sigma$ (you can ask why).

First, we require $l$ in Eq. (57) is a function of $\tau$, which satisfies the equation

$$\frac{dl(\tau)}{d\tau} = v_{\parallel} = \sigma\sqrt{\frac{2}{m}[w - \mu B(l(\tau))]}. \tag{59}$$

where $w$, $\mu$ is a constant independent of $\tau$. The second term in Eq. (57) can be integrated over $\tau$ to give

$$\int_{\tau_1}^{\tau_2} d\tau v_{\parallel}\frac{\partial f_1}{\partial l} = \int_{l_1}^{l_2} dl\frac{\partial f_1}{\partial l} = f_1(w, \mu, \sigma, l_2, t) - f_1(w, \mu, \sigma, l_1, t), \tag{60}$$

where $l_1 = l(\tau_1)$, $l_2 = l(\tau_2)$. If $f_1(w, \mu, \sigma, l_1, t) = f_1(w, \mu, \sigma, l_2, t)$, Eq. (60) reduces to zero. Thus the second term in Eq.(57) is eliminated from the equation. Now our task is to choose convenient starting and ending time, $\tau_1$ and $\tau_2$, to ensure that $f_1(w, \mu, \sigma, l_1, t) = f_1(w, \mu, \sigma, l_2, t)$. First, we choose $\tau_1 = 0$ and $l(\tau_1) = 0$. Then how to choose $\tau_2$ to ensure $f_1(w, \mu, \sigma, l_1 = 0, t) = f_1(w, \mu, \sigma, l_2, t)$? This requires we examine the equation that $l(\tau)$ satisfies, Eq.(59). For passing particles ($w > \mu B_{\max}$), it is obvious that the choice is to set $\tau_2$ so that $l(\tau_2) = L$, where $L$ is the length of the magnetic field line when it travels a full circle around the poloidal direction.

**Not correct**[ For trapped particles ($w < \mu B_{\max}$), how to choose $\tau_2$ is a little subtle. In this case, examining the motion equation,

$$\frac{dl(\tau)}{d\tau} = \sigma\sqrt{\frac{2}{m}[w - \mu B(l(\tau))]}. \tag{61}$$

we find that if $\sigma$ is constant independent of $\tau$, the above equation does not give a bounce motion, instead it gives a motion in which $l$ approach the point where $B(l_c) = w/\mu$, then do a infinitesimal oscillation at this point (in the region $l < l_c$). Therefore, in order to construct a bounce motion using Eq.(61), we have to require $\sigma$ to be a function of $\tau$:

$$\sigma = \sigma(\tau) = \begin{cases} 1 & \text{before reaching } l_c \text{ point} \\ -1 & \text{after reaching } l_c \text{ point} \end{cases} \tag{62}$$

Using this in Eq.(61) gives a bounce motion whose turning point is at $l = l_c$. Then it is obvious that $\tau_2$ can be chosen to the value so that $l(\tau_2) = 0$. ]**Not correct**

Comments: Owing to $\sigma$ contained in $f_0$, the above averaging method can not give information about how the distribution in the trapped region evolves. That is the distribution in the phase space where $w < \mu B_{\max}$ can not be determined from only Eq. (59), (the $f_1$ term can not be eliminated from this equation, thus the third equation about $f_1$ will come to play).

In the adjoint method for current drive probelm, the averaging process does not need to be applied to the trapped particles, since thess particles do not contribute to the parallel current density. So we do not care the value of distribution function in this region of phase space. Due to that $f_0$ is a old funtion about $\sigma$ in the passing region and $f_0$ is even function about $\sigma$ in the trapped region, the distribution function $f_0$ must be zero on the boundary between trapped and passing particles in the phase space. So the trapped particles effect gives a boundary condition for $f_0$ in the passing region.

Comments: Approximate or inconsitent models can often be confusing and misleading.

# 8 Some remarks

$$v_{\parallel} = \sigma\sqrt{\frac{2}{m}(w - \mu B)}. \tag{63}$$

$$\frac{dl}{d\tau} = v_{\parallel} = \sigma\sqrt{\frac{2}{m}[w - \mu B(l)]} \tag{64}$$

The passing orbits satisfy the condition

$$w > \mu B_{\max} \tag{65}$$

The trapped orbits satisfy the condition

$$w < \mu B_{\max} \tag{66}$$

The turning points of the trapped orbit are given by

$$w - \mu B(l_c) = 0 \Longrightarrow B(l_c) = \frac{w}{\mu} \tag{67}$$

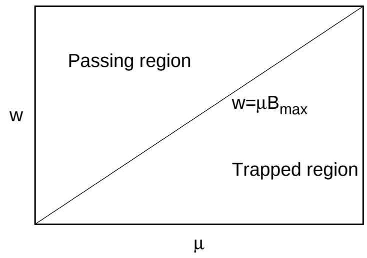


**Figure 1.** The line $w = \mu B_{\max}$ divides the $(w, \mu)$ plane into passing region and trapped region.

Question: In the phase space ($w_0$, $\mu_0$, $l_0$) where $w_0 < \mu_0 B_{\max}$ and $l_0 > l_c$ [$l_c$ is the turning point], the drift kinetic equation will become undefined since in this case the variable in the square root, $\frac{2}{m}[w_0 - \mu_0 B(l_0)]$, is negative. How to understand this?

# 9 Drift Kinetic Equation

In this section, we consider the cross field drift term, which, when included, will give rise to the bootstrapt current. In this case, the drift kinetic equation is

$$v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f + \boldsymbol{v}_d \cdot \nabla f = C(f), \tag{68}$$

where $f$ is the electron distribution function, $\hat{\boldsymbol{b}} = \boldsymbol{B}/B$, $\boldsymbol{v}_d$ is the drift velocity of guiding-centers perpendicular to magnetic field, $C(f)$ is the collision term.

We consider the case that the ratio of the typical Lamor radius to the minor radius of the device $\eta = \rho/a$ is a small parameter. We expand $f$ as $f = f_0 + f_1$ with $f_i \propto \eta^i$, then Eq. (68) is written as

$$v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f_0 + v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f_1 + \boldsymbol{v}_d \cdot \nabla f_0 + \boldsymbol{v}_d \cdot \nabla f_1 = C(f_0) + C(f_1). \tag{69}$$

The orders of the terms in the above equation are, respectively

$$v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f_0 \propto \varepsilon^0 \tag{70}$$

$$v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f_1 \propto \varepsilon^1 \tag{71}$$

$$\boldsymbol{v}_d \cdot \nabla f_0 \propto \varepsilon^1 \tag{72}$$

$$\boldsymbol{v}_d \cdot \nabla f_1 \propto \varepsilon^2 \tag{73}$$

$$C(f_0) \propto \varepsilon^0 \tag{74}$$

$$C(f_1) \propto \varepsilon^1 \tag{75}$$

Thus the zero order equation of Eq. (69) is

$$v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f_0 = C(f_0). \tag{76}$$

The first order equation of Eq. (69) is

$$v_{\parallel}\hat{\boldsymbol{b}} \cdot \nabla f_1 - C(f_1) = -\boldsymbol{v}_d \cdot \nabla f_0. \tag{77}$$

A Maxwellian distribution solves Eq. (76). (The density and temperature of the Maxwellian can still be a function of the flux surface.) In the following $f_0$ is considered to be Maxwellian, i.e., $f_0 = f_M$. To simplify Eq. (77), we write $f_1$ in the form

$$f_1 = -\frac{Iv_\parallel}{\Omega}\frac{\partial f_M}{\partial \psi} + g, \tag{78}$$

where $\psi$ is the label of the flux surface (here it is chosen to be poloidal flux), $I = B_\varphi R$, which is a function of only $\psi$, $\Omega = Be/m_e c$. Using Eq. (78) in Eq. (77) gives

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla\left(-\frac{Iv_\parallel}{\Omega}\frac{\partial f_M}{\partial \psi}\right) + v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g - C(g) = -\boldsymbol{v}_d \cdot \nabla f_M + C\left(-\frac{Iv_\parallel}{\Omega}\frac{\partial f_M}{\partial \psi}\right) \tag{79}$$

Noting that $I\partial f_M/\partial\psi$ is constant along a magnetic field line, the above equation becomes

$$-I\frac{\partial f_M}{\partial \psi} v_\parallel \hat{\boldsymbol{b}} \cdot \nabla\left(\frac{v_\parallel}{\Omega}\right) + v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g - C(g) = -\boldsymbol{v}_d \cdot \nabla f_M + C\left(-\frac{Iv_\parallel}{\Omega}\frac{\partial f_M}{\partial \psi}\right) \tag{80}$$

Using

$$\begin{aligned} \boldsymbol{v}_d \cdot \nabla f_M &= \frac{\partial f_M}{\partial \psi}\boldsymbol{v}_d \cdot \nabla \psi \\ &= \frac{\partial f_M}{\partial \psi} I v_\parallel \hat{\boldsymbol{b}} \cdot \nabla\left(\frac{v_\parallel}{\Omega}\right) \end{aligned} \tag{81}$$

in Eq. (80), gives

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g - C(g) = C\left(-\frac{Iv_\parallel}{\Omega}\frac{\partial f_M}{\partial \psi}\right) \tag{82}$$

# 10 Banana regime

In the banana regime (i.e., the bounce frequency is larger than the collision frequency, $\omega_b \gg \omega_c$), $g$ is expanded as $g = g_0 + g_1$, where $g_i \propto (\omega_c/\omega_b)^i$. Using this expansion in Eq. (82), gives

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_0 + v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_1 - C(g_0) - C(g_1) = 0. \tag{83}$$

Dividing both sides of the above equation by $\omega_b$, gives

$$\frac{1}{\omega_b} v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_0 + \frac{1}{\omega_b} v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_1 - \frac{1}{\omega_b} C(g_0) - \frac{1}{\omega_b} C(g_1) = 0 \tag{84}$$

We assume that

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla \sim \omega_b, \tag{85}$$

and

$$C \sim \omega_c. \tag{86}$$

Using the above assumptions, Eq. (84) can be separated into several equations according to the orders of the various terms. The orders of the terms in Eq. (84) are, repectively

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_0 \frac{1}{\omega_b} \sim (\omega_c/\omega_b)^0, \tag{87}$$

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_1 \frac{1}{\omega_b} \sim (\omega_c/\omega_b)^1, \tag{88}$$

$$C(g_0)\frac{1}{\omega_b} \sim (\omega_c/\omega_b)^1, \tag{89}$$

$$C(g_1)\frac{1}{\omega_b} \sim (\omega_c/\omega_b)^2, \tag{90}$$

Thus the zeroth order equation is

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_0 = 0. \tag{91}$$

The first order equation is

$$v_\parallel \hat{\boldsymbol{b}} \cdot \nabla g_1 - C(g_0) = 0. \tag{92}$$

Taking the bounce average of the above equation, we obtain

$$\int v_{\|}\hat{\boldsymbol{b}}\cdot\nabla g_1\frac{dl}{v_{\|}}-\langle C(g_0)\rangle_b=0, \tag{93}$$

where the bounce average operator is defined by

$$\langle A\rangle_b=\int A\frac{dl}{v_{\|}}. \tag{94}$$

The first term of Eq. (93) is zero. Thus Eq. (93) is reduced to

$$\langle C(g_0)\rangle_b=0. \tag{95}$$

# Bibliography


**[1]** R. B. White. *The theory of toroidal confined plasmas*. Imperial College Press, 2001.

**[2]** E. Westerhof, G. Giruzzi, R. W. Harvey, M. R. O'Brien, and A. G. Peeters. Comments on "analysis of electron cyclotron current drive using neoclassical fokker–planck code without bounce-average approximation" [phys. plasmas 2, 4570 (1995)]. *Phys. Plasmas*, 3(7):2827–2828, 1996.

**[3]** G. D. Kerbel and M. G. McCoy. Kinetic theory and simulation of multispecies plasmas in tokamaks excited with electromagnetic waves in the ion-cyclotron range of frequencies. *Phys. Fluids*, 28(12):3629–3653, 1985.