Derivation of the Kinetic Drift Equation

Steven Balbus

1 Intro

These notes are intended to provide a detailed, explicit derivation of the angle-averaged drift kinetic equation, which is the starting point of many astrophysical plasma problems. It is based on the development of Kulsrud 1983[1]. The equation is

$$\frac{\partial f_0}{\partial t} + (\boldsymbol{U_d} + v_\parallel \boldsymbol{b})\boldsymbol{\cdot}\boldsymbol{\nabla} f_0 - \left(\boldsymbol{b}\boldsymbol{\cdot}\frac{D\boldsymbol{U_d}}{Dt} - \mu B \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{b} - \frac{e}{m}E_\parallel\right)\frac{\partial f_0}{\partial v_\parallel} = 0 \qquad (1)$$

The notation is standard: $v_\parallel$ is the velocity along the field lines, μ is the magnetic moment, and $\boldsymbol{U_d}$ (what Kulsrud calls $\boldsymbol{U_E}$) is the drift velocity $c\boldsymbol{E} \times \boldsymbol{B}/B^2$.) The derivative D/Dt is $\partial/\partial t + (\boldsymbol{U_d} + \boldsymbol{b}v_\parallel)\boldsymbol{\cdot}\boldsymbol{\nabla}$.

For reference we also list Kulsrud's equation (37), needed as an intermediate point of the derivation. This is a Vlasov-type equation written in terms of $v_\perp$, the velocity perpendicular to the magnetic field, instead of μ. It is

$$\frac{\partial f_0}{\partial t} + (\boldsymbol{U_d} + v_\parallel \boldsymbol{b})\boldsymbol{\cdot}\boldsymbol{\nabla} f_0 - \frac{v_\perp}{2}(\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{U_d} - [(\boldsymbol{b}\boldsymbol{\cdot}\boldsymbol{\nabla})\boldsymbol{U_d}]\boldsymbol{\cdot}\boldsymbol{b} + v_\parallel \boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{b})\frac{\partial f_0}{\partial v_\perp}$$
$$- \left(\boldsymbol{b}\boldsymbol{\cdot}\frac{D\boldsymbol{U_d}}{Dt} - \frac{v_\perp^2}{2}\boldsymbol{\nabla}\boldsymbol{\cdot}\boldsymbol{b} - \frac{e}{m}E_\parallel\right)\frac{\partial f_0}{\partial v_\parallel} = 0 \qquad (2)$$

The outline of the derivation is that one starts with the standard Vlasov equation in a readily recognizable form, transforms to (2), then to (1).

2 Derivation of Kulsrud equation 37

The Vlasov equation for the distribution function f in standard form in standard notation in cgs units is

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla} f + \frac{e}{m}\left(\boldsymbol{E} + \frac{\boldsymbol{v} \times \boldsymbol{B}}{c}\right)\boldsymbol{\cdot}\boldsymbol{\nabla}_{\boldsymbol{v}} f = 0 \qquad (3)$$

[1] Handbook of Plasma Physics 1983, eds. M.N. Rosenbluth and R.Z. Sagdeev, Vol. 1: Basic Plasma Physics I (North Holland Publishing)

The final term is dominant, since it is associated with a (very rapid) cyclotron frequency, $\omega_c = eB/mc$. To effect a perturbation analysis, we therefore insert a small parameter ϵ into the equation:

$$\frac{\partial f}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} f + \left(\frac{1}{\epsilon}\right)\frac{e}{m}\left(\boldsymbol{E} + \frac{\boldsymbol{v}\times\boldsymbol{B}}{c}\right)\cdot\boldsymbol{\nabla}_{\boldsymbol{v}} f = 0, \tag{4}$$

and seek solutions in the form of $f = f_0 + \epsilon f_1 + ...$

Inserting the series expansion for f into (4) and retaining the term of order $1/\epsilon$ we find

$$\left(\boldsymbol{E} + \frac{\boldsymbol{v}\times\boldsymbol{B}}{c}\right)\cdot\boldsymbol{\nabla}_{\boldsymbol{v}} f_0 = 0. \tag{5}$$

We shall work with the velocity $\boldsymbol{v'}$ relative to the $E\times B$ drift velocity $\boldsymbol{U_d}$, defined by

$$\boldsymbol{U_d} = c\frac{\boldsymbol{E}\times\boldsymbol{B}}{B^2}. \tag{6}$$

We next divide $\boldsymbol{E}$ into components along $\boldsymbol{B}$ (unit vector $\boldsymbol{b}$) and perpendicular to $\boldsymbol{B}$:

$$\boldsymbol{E} = \boldsymbol{E}_\perp + E_\parallel \boldsymbol{b} \tag{7}$$

and define

$$\boldsymbol{v'} = \boldsymbol{v} - \boldsymbol{U_d}. \tag{8}$$

Then, direction substitution using (8) yields

$$\boldsymbol{E} + \frac{\boldsymbol{v}\times\boldsymbol{B}}{c} = E_\parallel \boldsymbol{b} + \frac{\boldsymbol{v'}\times\boldsymbol{B}}{c} \tag{9}$$

Hence

$$\left(\boldsymbol{E} + \frac{\boldsymbol{v}\times\boldsymbol{B}}{c}\right)\cdot\boldsymbol{\nabla}_{\boldsymbol{v}} f_0 = \left[\frac{\boldsymbol{v'}\times\boldsymbol{B}}{c}\right]\cdot\boldsymbol{\nabla}_{\boldsymbol{v'}} + E_\parallel \boldsymbol{b}\cdot\boldsymbol{\nabla}_{\boldsymbol{v'}} = 0, \tag{10}$$

since $\boldsymbol{\nabla}_{\boldsymbol{v}} = \boldsymbol{\nabla}_{\boldsymbol{v'}}$.

Our next step is to introduce the velocity coordinate variables, $v_\perp$, ϕ, and $v_\parallel$. They are related to $\boldsymbol{v'}$ by

$$\boldsymbol{v'} = \hat{\boldsymbol{x}} v_\perp \cos\phi + \hat{\boldsymbol{y}} v_\perp \sin\phi + \boldsymbol{b} v_\parallel \tag{11}$$

Notice that the direction associated with $\boldsymbol{v}_\perp$ is *radial.* Thus,

$$\frac{\boldsymbol{v'}\times\boldsymbol{B}}{c} = -\frac{v_\perp B}{c}\hat{\boldsymbol{\phi}}. \tag{12}$$

With

$$\hat{\boldsymbol{\phi}}\boldsymbol{\cdot}\boldsymbol{\nabla}_{\boldsymbol{v'}} = \frac{1}{v_\perp}\frac{\partial}{\partial\phi} \tag{13}$$

equation (10) becomes

$$-\frac{B}{c}\frac{\partial f_0}{\partial\phi} + E_\parallel \frac{\partial f_0}{\partial v_\parallel} = 0 \tag{14}$$

But this is a ridiculous equation as it stands. All instinct suggests that f should be phase independent, while (14) states that f_0 is constant along the helix

$$\frac{dv_\parallel}{d\phi} = -\frac{B}{cE_\parallel} \tag{15}$$

up to, apparently, $v_\parallel = \infty$! Obviously, the assumption that $E_\parallel$ is present at leading order is incorrect. Instead we have

$$\boldsymbol{E} = \boldsymbol{E}_\perp + \epsilon \boldsymbol{b} E_\parallel + ... \tag{16}$$

and to leading order,

$$\frac{\partial f_0}{\partial\phi} = 0 \tag{17}$$

Going back to the Vlasov equation (3), we retain terms of order unity in the ϵ expansion. This immediately gives

$$\frac{\partial f_0}{\partial t} + \boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla} f_0 + \frac{e}{m}\left(\boldsymbol{E}_\perp + \frac{\boldsymbol{v}\times\boldsymbol{B}}{c}\right)\boldsymbol{\cdot}\boldsymbol{\nabla}_{\boldsymbol{v}} f_1 + \frac{e}{m}E_\parallel \frac{\partial f_0}{\partial v_\parallel} = 0 \tag{18}$$

where we have retained the $E_\parallel$ term, now entering at order unity in ϵ expansion, in terms of its $v_\parallel$ form. Proceeding identically as we did for the $1/\epsilon$ term, expressing the velocity gradients in $v_\perp$, $v_\parallel$, ϕ variables, the equation becomes

$$\frac{eB}{mc}\frac{\partial f_1}{\partial\phi} = (\frac{\partial f_0}{\partial t} + \boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla})^* f_0 + \frac{eE_\parallel}{m}\frac{\partial f_0}{\partial v_\parallel}. \tag{19}$$

At this point, we first note that the $*$'ed spatial gradients are untransformed, taken at constant $\boldsymbol{v}$. To be mathematically useful, these gradients need to be at constant cylindrical velocity. The problem is that $v_\parallel$ and $v_\perp$ depend upon position through $\boldsymbol{b}$ and $\boldsymbol{U_d}$ (and of course $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$). Hence the spatial gradients will generate $v_\parallel$ and $v_\perp$ gradients upon coordinate transformation.

We next note that if we average this equation over ϕ, the right side must vanish since the left side does so trivially (B clearly does not depend on ϕ).

This means that many terms will vanish once the averaging is effected, only the even powers of $\cos\phi$ and $\sin\phi$ will survive.

Let us see how the coordinate transformation works in detail. The formal transformation between our velocity coordinates is given by

$$\boldsymbol{v} = \boldsymbol{U_d} + \boldsymbol{b}v_\parallel + \hat{\boldsymbol{x}}v_\perp\cos\phi + \hat{\boldsymbol{y}}v_\perp\sin\phi \tag{20}$$

where $\boldsymbol{U_d}$ and the unit vectors $\boldsymbol{b}$, $\hat{\boldsymbol{x}}$, and $\hat{\boldsymbol{y}}$ are all coordinate dependent. (We will ignore a possible time dependence as it is a trivial formal complication that causes only more writing.) The reverse transformation is

$$v_\parallel = \boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{b}, \qquad {v_\perp}^2 = (\boldsymbol{v}\boldsymbol{\cdot}\hat{\boldsymbol{x}} - U_{dx})^2 + (\boldsymbol{v}\boldsymbol{\cdot}\hat{\boldsymbol{y}} - U_{dy})^2 \tag{21}$$

where U_{dx} and U_{dy} are the x and y components of $\boldsymbol{U_d}$.

The spatial derivatives transform as

$$\frac{\partial}{\partial x} = \left[\frac{\partial}{\partial x}\right] + \frac{\partial v_\perp}{\partial x}\frac{\partial}{\partial v_\perp} + \frac{\partial v_\parallel}{\partial x}\frac{\partial}{\partial v_\parallel} \tag{22}$$

and similarly for y and z (the field direction). We use square brackets for the new spatial gradients on the right, which appear in the same form as the original spatial gradients, BUT with $v_\perp$, $v_\parallel$ and ϕ held constant. Therefore,

$$\boldsymbol{v}\boldsymbol{\cdot}\boldsymbol{\nabla} = \Big[(\boldsymbol{b}v_\parallel + \boldsymbol{U_d})\boldsymbol{\cdot}\boldsymbol{\nabla}\Big] + \boldsymbol{v}\boldsymbol{\cdot}\hat{\boldsymbol{\alpha}}\frac{\partial v_\perp}{\partial\alpha}\frac{\partial}{\partial v_\perp} + \boldsymbol{v}\boldsymbol{\cdot}\hat{\boldsymbol{\alpha}}\frac{\partial v_\parallel}{\partial\alpha}\frac{\partial}{\partial v_\parallel} \tag{23}$$

where the dummy variable α is summed over x, y and z. Inside the square brackets on the right, we have not written out the $v_\perp$ terms, as these $\cos\phi$ and $\sin\phi$ terms will vanish when ϕ is averaged.

Next, we compute the gradients of $v_\perp$,

$$\frac{\partial v_\perp}{\partial x} = \cos\phi\left(\boldsymbol{v}\boldsymbol{\cdot}\frac{\partial\hat{\boldsymbol{x}}}{\partial x} - \frac{\partial U_{dx}}{\partial x}\right) + \sin\phi\left(\boldsymbol{v}\boldsymbol{\cdot}\frac{\partial\hat{\boldsymbol{y}}}{\partial x} - \frac{\partial U_{dy}}{\partial x}\right) \tag{24}$$

$$\frac{\partial v_\perp}{\partial y} = \cos\phi\left(\boldsymbol{v}\boldsymbol{\cdot}\frac{\partial\hat{\boldsymbol{x}}}{\partial y} - \frac{\partial U_{dx}}{\partial y}\right) + \sin\phi\left(\boldsymbol{v}\boldsymbol{\cdot}\frac{\partial\hat{\boldsymbol{y}}}{\partial y} - \frac{\partial U_{dy}}{\partial y}\right) \tag{25}$$

$$\frac{\partial v_\perp}{\partial z} = \cos\phi\left(\boldsymbol{v}\boldsymbol{\cdot}\frac{\partial\hat{\boldsymbol{x}}}{\partial z} - \frac{\partial U_{dx}}{\partial z}\right) + \sin\phi\left(\boldsymbol{v}\boldsymbol{\cdot}\frac{\partial\hat{\boldsymbol{y}}}{\partial z} - \frac{\partial U_{dy}}{\partial z}\right) \tag{26}$$

It is important to note that the spatial gradients of $\hat{\boldsymbol{x}}$ or $\hat{\boldsymbol{y}}$ have only $\boldsymbol{b}$ components. This is because these unit vectors lie in the plane transverse to

$\boldsymbol{b}$, and only a change in the tilt of $\boldsymbol{b}$ through some angle can induce change in $\hat{\boldsymbol{x}}$ or $\hat{\boldsymbol{y}}$. No change in $\boldsymbol{b(x)}$ arising from changing location can cause $\hat{\boldsymbol{x}}$ or $\hat{\boldsymbol{y}}$ to change by remaining in their mutual plane and twisting around the $\boldsymbol{b}$ axis. Note as well the identities

$$\boldsymbol{b}\cdot\frac{\partial\hat{\boldsymbol{w}}}{\partial u} = -\hat{\boldsymbol{w}}\cdot\frac{\partial\boldsymbol{b}}{\partial u} \tag{27}$$

where u and w can be x, y, or z (in which case both sides vanish).

Consider now

$$\boldsymbol{v}\cdot\hat{\boldsymbol{x}}\frac{\partial v_\perp}{\partial x} = (U_{dx}+v_\perp\cos\phi)\left[\cos\phi\left(v_\parallel\boldsymbol{b}\cdot\frac{\partial\hat{\boldsymbol{x}}}{\partial x} - \frac{\partial U_{dx}}{\partial x}\right) + \sin\phi\left(v_\parallel\boldsymbol{b}\cdot\frac{\partial\hat{\boldsymbol{y}}}{\partial x} - \frac{\partial U_{dy}}{\partial x}\right)\right] \tag{28}$$

Upon ϕ averaging, only the $\langle\cos^2\phi\rangle = 1/2$ term survives:

$$\boldsymbol{v}\cdot\hat{\boldsymbol{x}}\frac{\partial v_\perp}{\partial x} = \frac{v_\perp}{2}\left(v_\parallel\boldsymbol{b}\cdot\frac{\partial\hat{\boldsymbol{x}}}{\partial x} - \frac{\partial U_{dx}}{\partial x}\right) = -\frac{v_\perp}{2}\left(v_\parallel\hat{\boldsymbol{x}}\cdot\frac{\partial\boldsymbol{b}}{\partial x} + \frac{\partial U_{dx}}{\partial x}\right) \tag{29}$$

Similarly,

$$\boldsymbol{v}\cdot\hat{\boldsymbol{y}}\frac{\partial v_\perp}{\partial y} = -\frac{v_\perp}{2}\left(v_\parallel\hat{\boldsymbol{y}}\cdot\frac{\partial\boldsymbol{b}}{\partial y} + \frac{\partial U_{dy}}{\partial y}\right), \tag{30}$$

but

$$\boldsymbol{v}\cdot\hat{\boldsymbol{z}}\frac{\partial v_\perp}{\partial z} = 0. \tag{31}$$

Now the $\boldsymbol{U_d}$ derivatives in (29) and (30) have the appearance of divergence terms, but remember that our unit vectors are not strictly Cartesian, they depend upon position. So, for example,

$$\frac{\partial U_{dx}}{\partial x} = \hat{\boldsymbol{x}}\cdot\boldsymbol{\nabla}(\boldsymbol{U_d}\cdot\hat{\boldsymbol{x}}), \tag{32}$$

whereas the correct form of the divergence term is

$$[(\hat{\boldsymbol{x}}\cdot\boldsymbol{\nabla})\boldsymbol{U_d}]\cdot\hat{\boldsymbol{x}}. \tag{33}$$

The difference between (32) and (33), however, is

$$[(\hat{\boldsymbol{x}}\cdot\boldsymbol{\nabla})\hat{\boldsymbol{x}}]\cdot\boldsymbol{U_d}, \tag{34}$$

and this vanishes because $\boldsymbol{U_d}$ has only $\boldsymbol{x}$ and $\boldsymbol{y}$ components, while any spatial gradient of $\hat{\boldsymbol{x}}$ has, as noted, only a $\boldsymbol{b}$ component.

Collecting all the terms in (29) and (30) proportional to $\partial/\partial v_\perp$, we obtain

$$-\frac{v_\perp}{2}\left(v_\parallel \boldsymbol{\nabla\cdot b} + \boldsymbol{\nabla\cdot U_d} - [(\boldsymbol{b\cdot\nabla})\boldsymbol{U_d}]\boldsymbol{\cdot b}\right)\frac{\partial}{\partial v_\perp} \tag{35}$$

(The final term in square brackets has been added and subtracted "by hand" to complete the $\boldsymbol{U_d}$ divergence.)

On to the the $v_\parallel$ terms. We have

$$\frac{\partial v_\parallel}{\partial x} = \boldsymbol{v\cdot}\frac{\partial \boldsymbol{b}}{\partial x} = (\boldsymbol{U_d} + v_\perp \cos\phi\, \boldsymbol{\hat{x}} + v_\perp \sin\phi\, \boldsymbol{\hat{y}})\boldsymbol{\cdot}\frac{\partial \boldsymbol{b}}{\partial x} \tag{36}$$

Then,

$$(\boldsymbol{v\cdot\hat{x}})\frac{\partial v_\parallel}{\partial x} = (U_{dx} + v_\perp\cos\phi)\left(\boldsymbol{U_d\cdot}\frac{\partial \boldsymbol{b}}{\partial x} + v_\perp\cos\phi\,\boldsymbol{\hat{x}\cdot}\frac{\partial \boldsymbol{b}}{\partial x} + ...\right) \tag{37}$$

Averaging over ϕ:

$$(\boldsymbol{v\cdot\hat{x}})\frac{\partial v_\parallel}{\partial x} = U_{dx}\boldsymbol{U_d\cdot}\frac{\partial \boldsymbol{b}}{\partial x} + \frac{v_\perp^2}{2}\boldsymbol{\hat{x}\cdot}\frac{\partial \boldsymbol{b}}{\partial x} \tag{38}$$

Exactly the same calculation produces

$$(\boldsymbol{v\cdot\hat{y}})\frac{\partial v_\parallel}{\partial y} = U_{dy}\boldsymbol{U_d\cdot}\frac{\partial \boldsymbol{b}}{\partial y} + \frac{v_\perp^2}{2}\boldsymbol{\hat{y}\cdot}\frac{\partial \boldsymbol{b}}{\partial y} \tag{39}$$

The z calculation is

$$(\boldsymbol{v\cdot\hat{z}})\frac{\partial v_\parallel}{\partial z} = v_\parallel(\boldsymbol{U_d} + v_\perp\cos\phi\,\boldsymbol{\hat{x}} + v_\perp\sin\phi\,\boldsymbol{\hat{y}})\boldsymbol{\cdot}\frac{\partial \boldsymbol{b}}{\partial z} \to v_\parallel \boldsymbol{U_d\cdot}\frac{\partial \boldsymbol{b}}{\partial z} \tag{40}$$

upon averaging. Summing all the terms, and integrating the $\boldsymbol{U_d\cdot}\partial\boldsymbol{b}$ terms by parts produces

$$-\boldsymbol{b\cdot}(\boldsymbol{U_d\cdot\nabla} + v_\parallel \boldsymbol{b\cdot\nabla})\boldsymbol{U_d} + \frac{v_\perp^2}{2}\boldsymbol{\nabla\cdot b} \tag{41}$$

These are precisely the terms in addition to the $E_\parallel$ term in Kulsrud (37).

3 Derivation of Kulsrud 51

The next step is switch from $v_\perp$ to $\mu = v_\perp^2/2B$. It will help in our calculation to note the following identity ahead of time:

$$\boldsymbol{\nabla\cdot b} = \partial_i(B_i/B) = -(1/B^2)B_i\partial_i B = -\partial_z \ln B. \tag{42}$$

Now,

$$\frac{\partial \mu}{\partial x} = -\frac{v_\perp^2}{2B^2}\frac{\partial B}{\partial x} = -\mu \frac{\partial \ln B}{\partial x} \tag{43}$$

and similarly for x and y. The transformation between $v_\perp$ and μ gradients is

$$\frac{\partial}{\partial v_\perp} = \frac{\partial \mu}{\partial v_\perp}\frac{\partial}{\partial \mu} = \frac{v_\perp}{B}\frac{\partial}{\partial \mu}, \tag{44}$$

and

$$\frac{v_\perp}{2}\frac{\partial}{\partial v_\perp} = \mu \frac{\partial}{\partial \mu}. \tag{45}$$

The coordinate transformation is

$$\frac{\partial}{\partial x} = \left[\frac{\partial}{\partial x}\right] - \mu \frac{\partial \ln B}{\partial x}\frac{\partial}{\partial \mu} \tag{46}$$

and similarly for y and z, but for z we use the identity (42) to write

$$\frac{\partial}{\partial z} = \left[\frac{\partial}{\partial z}\right] + \mu \boldsymbol{\nabla}\cdot\boldsymbol{b}\frac{\partial}{\partial \mu} \tag{47}$$

The term $(\boldsymbol{U_d} + v_\parallel \boldsymbol{b})\cdot\boldsymbol{\nabla}$ in equation (2) therefore generates, upon transformation of spatial derivatives, the follwoing additional terms proportional to $\partial/\partial \ln \mu$:

$$-\boldsymbol{U_d}\cdot\boldsymbol{\nabla} \ln B + v_\parallel \boldsymbol{\nabla}\cdot\boldsymbol{b} \tag{48}$$

To these, we must add the existing terms from the $(v_\perp/2)(\partial/\partial v_\perp)$ derivative of equation (37):

$$-\boldsymbol{\nabla}\cdot\boldsymbol{U_d} + [(\boldsymbol{b}\cdot\boldsymbol{\nabla})\boldsymbol{U_d}]\cdot\boldsymbol{b} - v_\parallel \boldsymbol{\nabla}\cdot\boldsymbol{b}. \tag{49}$$

The $v_\parallel$ terms cancel, leaving

$$-\boldsymbol{U_d}\cdot\boldsymbol{\nabla} \ln B - \boldsymbol{\nabla}\cdot\boldsymbol{U_d} + [(\boldsymbol{b}\cdot\boldsymbol{\nabla})\boldsymbol{U_d}]\cdot\boldsymbol{b} \tag{50}$$

But these three terms may be combined into

$$\frac{1}{B^2}\boldsymbol{B}\cdot\boldsymbol{\nabla}\times(\boldsymbol{U_d}\times\boldsymbol{B}) = \frac{1}{B^2}\boldsymbol{B}\cdot\boldsymbol{\nabla}\times(\boldsymbol{U}\times\boldsymbol{B}) \tag{51}$$

which vanishes within our scheme of ignoring the time dependence. (Including the time dependent terms would just reproduce this simplification via the full induction equation.) Hence all the terms proportional to the μ partial derivative vanish in the final equation. This is as expected, since the terms must involve the time derivative of the magnetic moment, and adiabatic *constant*.

This completes are derivation of the kinetic drift equation.