Some Notes on Geometry: Metric and Symplectic

Ammar H. Hakim

Contents

1 Coordinate Transforms and Metric Tensor

We will denote physical Cartesian coordinates as $x^1, x^2, \ldots, x^n$ and computational space coordinates as $z^1, z^2, \ldots, z^n$. The mapping between computation space coordinates and physical coordinates will be denoted by

$$x^i = x^i(z^1, \ldots, z^n) \tag{1}$$

or $\mathbf{x} = \mathbf{x}(z^1, \ldots, z^n)$, for $i = 1, \ldots, n$. We will assume this mapping is invertible.

The *metric* tensor $\mathbf{g}$ is a fundamental geometrical object, and measures the length and angle between vectors. In Cartesian coordinates the length of a vector connecting two close-by points, $d\mathbf{x}$ is $ds^2 = d\mathbf{x} \cdot d\mathbf{x}$. Using chain rule in the z^i coordinates

$$d\mathbf{x} = \frac{\partial \mathbf{x}}{\partial z^i} dz^i = \mathbf{e}_i dz^i \tag{2}$$

where summation over repeated indices is assumed. The *tangent vectors* $\mathbf{e}_i$ and their *duals* (defined below) play a fundamental role in geometry. In fact, all of the geometrical information we need is contained in them. In terms of the tangent vector the length of the infinitesimal vector $d\mathbf{x}$ is

$$ds^2 = d\mathbf{x} \cdot d\mathbf{x} = \mathbf{e}_i \cdot \mathbf{e}_j dz^i dz^j = g_{ij} dz^i dz^j \tag{3}$$

where we have defined the *metric tensor* as $g_{ij} \equiv \mathbf{e}_i \cdot \mathbf{e}_j$. In the language of Thorne and Blanford[5] the metric tensor is an object that takes two vectors and pops out a number: $\mathbf{g}(\_,\_)$. Note that written like this, the metric is *coordinate free*. When fed the tangent vectors this object produces the components of the metric in that coordinate system, $\mathbf{g}(\mathbf{e}_i, \mathbf{e}_j) = g_{ij}$.

Given the set of tangent vectors we can construct their *duals*, which are simply the coordinate gradients, i.e. $\mathbf{e}^i = \nabla z^i$. The chain rule shows that

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta^i{}_j \tag{4}$$

This is a fundamental relation that shows that the tangent vectors and their duals form an orthonormal system of vectors. Multiplying by g_{ki} one gets $g_{ki}\mathbf{e}^i \cdot \mathbf{e}_j = g_{ki}\delta^i{}_j = g_{kj}$, implying, from the definition of the metric tensor:

$$\mathbf{e}_k = g_{ki}\mathbf{e}^i. \tag{5}$$

Taking the dot product of this with $\mathbf{e}^j$ leads to

$$\mathbf{e}_k \cdot \mathbf{e}^j = \delta_k{}^j = g_{ki}\mathbf{e}^i \cdot \mathbf{e}^j \tag{6}$$

Hence, the object $g^{ij} \equiv \mathbf{e}^i \cdot \mathbf{e}^j$ is the *inverse* of the metric tensor, that is:

$$g_{ki}g^{ij} = \delta_k{}^j. \tag{7}$$

In a similar manner we can show that the dual vectors can also be computed from the inverse of the metric tensor as

$$\mathbf{e}^i = g^{ij}\mathbf{e}_j. \tag{8}$$

It may appear at first sight that to compute the dual vectors we need to invert the mapping, i.e. express the z^i in terms of the x^i and then compute ∇z^i. However that is not so: once the metric and its inverse are determined, the dual vectors can be computed from Eq. (8). Another, though less general approach, is to first compute the Jacobian matrix of the transform:

$$D^i{}_j = \frac{\partial x^i}{\partial z^j}. \tag{9}$$

Note that the tangent vectors form the *columns* of the $\mathbf{D}$ matrix. Now, the *inverse* of this matrix can be computed $\mathbf{M} = \mathbf{D}^{-1}$. Hence as $\mathbf{MD} = \mathbf{I}$, the *rows* of this matrix are the coordinate gradients or the dual vectors. (This follows from the defining equation Eq. (4)).

Computing derivatives of vectors and other tensor quantities is subtle in the z^i coordinates. We must ensure that we only compute derivatives of geometric quantities and not directly of the components themselves. Consider a vector written using the tangent vectors as $\mathbf{A} = A^i\mathbf{e}_i$. This is a geometric quantity and we can compute it's gradient as follows

$$\nabla_j \mathbf{A} = \frac{\partial A^i}{\partial z^j}\mathbf{e}_i + A^i\frac{\partial \mathbf{e}_i}{\partial z^j}. \tag{10}$$

Note that the first term in the vector gradient are just the partial derivatives of the components A^i but *those are not enough* as one must also account for the fact that as one moves from point-to-point, the tangent vectors will, in general, also *change direction*. As we have the mappings, we can compute their gradients directly. However, it is more conventional to express the gradient in terms of the tangent vectors themselves as

$$\frac{\partial \mathbf{e}_i}{\partial z^j} \equiv \Gamma^k{}_{ij}\mathbf{e}_k. \tag{11}$$

The quantities $\Gamma^k{}_{ij}$ are known as the *Christoffel symbols*. Using these in Eq. (10) we get

$$\nabla_j \mathbf{A} = \left(\frac{\partial A^k}{\partial z^j} + A^i \Gamma^k{}_{ij}\right)\mathbf{e}_k \equiv A^k{}_{;j}\mathbf{e}_k \tag{12}$$

The quantities $A^i{}_{;j}$ are the components along the tangent vector of the vector gradient:

$$A^k{}_{;j} = \frac{\partial A^k}{\partial z^j} + A^i \Gamma^k{}_{ij}. \tag{13}$$

In the literature this is called the *covariant derivative* of $\mathbf{A}$. Note that the Christoffel symbols can be computed by taking the dot product of Eq. (11) by $\mathbf{e}^m$ to get

$$\Gamma^m{}_{ij} = \mathbf{e}^m \cdot \frac{\partial \mathbf{e}_i}{\partial z^j}. \tag{14}$$

The Christoffel symbols are symmetric in the lower indices as

$$\frac{\partial \mathbf{e}_i}{\partial z^j} = \frac{\partial}{\partial z^j}\left(\frac{\partial \mathbf{x}}{\partial z^i}\right) = \frac{\partial}{\partial z^i}\left(\frac{\partial \mathbf{x}}{\partial z^j}\right) = \frac{\partial \mathbf{e}_j}{\partial z^i} \tag{15}$$

showing that

$$\Gamma^m{}_{ij} = \Gamma^m{}_{ji}. \tag{16}$$

Note that taking the derivative of $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$, cyclically permuting indices and using the symmetry property Eq. (15) we can express the Christoffel symbols in terms of the metric tensor as

$$\Gamma^n{}_{kj} = \frac{1}{2} g^{nm}\left(\frac{\partial g_{mk}}{\partial z^j} + \frac{\partial g_{jm}}{\partial z^k} - \frac{\partial g_{kj}}{\partial z^m}\right). \tag{17}$$

This is left as an (tedious) exercise for the reader.

If we instead used the dual vectors to express the vector $\mathbf{A} = A_i \mathbf{e}^i$ the gradient would be[1]

$$\nabla_j \mathbf{A} = \frac{\partial A_i}{\partial z^j}\mathbf{e}^i + A_i \frac{\partial \mathbf{e}^i}{\partial z^j} = \left(\frac{\partial A_k}{\partial z^j} + A_i \mathbf{e}_k \cdot \frac{\partial \mathbf{e}^i}{\partial z^j}\right)\mathbf{e}^k \tag{18}$$

[1] Consider a vector written in terms of the duals: $\mathbf{a} = a_k \mathbf{e}^k$. From the orthonormality relation we have $a_k = \mathbf{a} \cdot \mathbf{e}_k$. Hence we can write $\mathbf{a} = (\mathbf{a} \cdot \mathbf{e}_k)\mathbf{e}^k$.

Now take derivative of $\mathbf{e}_k \cdot \mathbf{e}^i = \delta_k{}^i$ to get:

$$\mathbf{e}_k \cdot \frac{\partial \mathbf{e}^i}{\partial z^j} + \frac{\partial \mathbf{e}_k}{\partial z^j} \cdot \mathbf{e}^i = 0. \tag{19}$$

This implies that

$$\mathbf{e}_k \cdot \frac{\partial \mathbf{e}^i}{\partial z^j} = -\mathbf{e}^i \cdot \frac{\partial \mathbf{e}_k}{\partial z^j} = -\Gamma^i{}_{kj}. \tag{20}$$

Finally, using this in Eq. (18):

$$\nabla_j \mathbf{A} = \left(\frac{\partial A_k}{\partial z^j} - A_i \Gamma^i{}_{kj} \right) \mathbf{e}^k \equiv A_{k;j} \mathbf{e}^k. \tag{21}$$

The quantities $A_{i;j}$ are the components along the dual vectors of the vector gradient:

$$A_{k;j} = \frac{\partial A_k}{\partial z^j} - A_i \Gamma^i{}_{kj}. \tag{22}$$

For higher-order tensors we can take the derivatives by simply applying the rules Eq. (13) and Eq. (22) for each "upstairs" and "downstairs" indices in turn. For example,

$$T^i{}_{j;k} = \frac{\partial T^i{}_j}{\partial z^k} + T^m{}_j \Gamma^i{}_{mk} - T^i{}_m \Gamma^m{}_{jk}. \tag{23}$$

Curiously, we can easily show that the covariant derivative of the metric tensor is zero. To show this we use the identity Eq. (14) to write

$$g_{mk} \Gamma^m{}_{ij} = \mathbf{e}_k \cdot \frac{\partial \mathbf{e}_i}{\partial z^j}. \tag{24}$$

Swapping k and i and adding the resulting equation to the above we get

$$g_{mk} \Gamma^m{}_{ij} + g_{mi} \Gamma^m{}_{kj} = \frac{\partial g_{ik}}{\partial z^j} \tag{25}$$

which shows $g_{ik;j} = 0$.

For several differential operators we do not need the Christoffel symbols directly. In particular, one can show that the divergence of a vector can be written as

$$A^j{}_{;j} = \frac{1}{\sqrt{g}} \frac{\partial}{\partial z^j} \left(\sqrt{g} A^j \right) \tag{26}$$

where $\sqrt{g} = \sqrt{\det(g_{ij})}$. Using this the Laplacian of a scalar can be written as as the divergence of the gradient:

$$\nabla^2 \phi = \frac{1}{\sqrt{g}} \frac{\partial}{\partial z^j} \left(\sqrt{g} g^{jk} \frac{\partial \phi}{\partial z^k} \right). \tag{27}$$

For other operators, in particular the divergences of general second-order tensors needed in writing down fluid equations, we need the Christoffel symbols.

2 Field-line following coordinates

We will now focus on 3D space and assume we have the mapping $\mathbf{x} = \mathbf{x}(z^1, z^2, z^3)$. Now, once we have computed the tangent vectors and their duals, consider the vector field

$$\mathbf{B} = C(\mathbf{x})\mathbf{e}^1 \times \mathbf{e}^2 \tag{28}$$

where $C(\mathbf{x})$ is a scalar field, as yet undefined. This form may appear very arbitrary and peculiar, but to see its utility, now take the divergence of $\mathbf{B}$ to get:

$$\nabla \cdot \mathbf{B} = \nabla \cdot \big(C(\mathbf{x})\mathbf{e}^1 \times \mathbf{e}^2\big) = C\big(\mathbf{e}^2 \cdot \underbrace{\nabla \times \mathbf{e}^1}_{\nabla\times\nabla z^1 = 0} - \mathbf{e}^1 \cdot \underbrace{\nabla \times \mathbf{e}^2}_{\nabla\times\nabla z^2=0}\big) + \mathbf{e}^1 \times \mathbf{e}^2 \cdot \nabla C. \tag{29}$$

We are now left with

$$\nabla \cdot \mathbf{B} = \frac{1}{C}\mathbf{B} \cdot \nabla C \tag{30}$$

where we have assumed $C(\mathbf{x}) \neq 0$. Hence, if we choose $C(\mathbf{x})$ such that $\mathbf{B} \cdot \nabla C = 0$, then the vector field $\mathbf{B}$ is *divergence free*[2]. Hence, in a magnetized plasma problem we can *choose* the coordinate system (at least locally) such that the form of the magnetic field is given by Eq. (28). Such a coordinate system is called a *field-line following coordinate*.

In a field-line following coordinate the $\mathbf{B}$ vector is everywhere parallel to $\mathbf{e}^3$. This follows form the fact that in writing the $\mathbf{B}$ as Eq. (28) we have ensured that $\mathbf{B} \cdot \mathbf{e}^1 = \mathbf{B} \cdot \mathbf{e}^2 = 0$. Hence, we can also write

$$\mathbf{B} = \big(\mathbf{B} \cdot \mathbf{e}^3\big)\mathbf{e}_3. \tag{31}$$

Using Eq. (28) in this we get[3]

$$\mathbf{B} = C \underbrace{(\mathbf{e}^1 \times \mathbf{e}^2) \cdot \mathbf{e}^3}_{J^{-1}} \mathbf{e}_3. \tag{32}$$

Hence, we must have

$$B^2 = \mathbf{B} \cdot \mathbf{B} = \frac{C^2}{J^2} \underbrace{\mathbf{e}_3 \cdot \mathbf{e}_3}_{g_{33}} \tag{33}$$

and hence

$$C = \frac{BJ}{\sqrt{g_{33}}}. \tag{34}$$

[2]This form of a divergence-free field is often known as the *Clebsch* form.

[3]In 3D the Jacobian of the transform is $1/(\mathbf{e}^1 \times \mathbf{e}^2) \cdot \mathbf{e}^3$ as can be seen from the definition of the Jacobian matrix and its inverse.

Now we can *choose* the *length along the field-line* as the parallel coordinate z^3. In this case, we must have $g_{33} = 1$, as this represents the line-element along the z^3 direction (i.e. when $dz^1 = dz^2 = 0$). Hence, on a given field-line, the condition for divergence-free field must be

$$\mathbf{B} \cdot \nabla(BJ) = 0 \tag{35}$$

or that BJ must be a *constant* on a field-line. Hence, on a field-line we must have $J = \mathrm{const}/B$, *independent* of the specific form of the field-line following coordinate transform and as long as $g_{zz} = 1$, i.e. we use the length along the field-line as the z^3 coordinate.

3 Symplectic Geometry and Transforming Hamiltonian Systems

Hamiltonian systems can be written in the generic form

$$\frac{\partial f}{\partial t} + \{f, H\} = 0 \tag{36}$$

where f is the distribution function, H is the Hamiltonian and $\{f, g\}$ is the Poisson bracket operator. Hamiltonian dynamics live in $2n$ dimensional canonical phase-space $(\mathbf{x}, \mathbf{v})$. The Poisson bracket operator is defined as

$$\{f, g\} = \nabla_{\mathbf{x}} f \cdot \nabla_{\mathbf{v}} g - \nabla_{\mathbf{v}} f \cdot \nabla_{\mathbf{x}} g \tag{37}$$

Transforming these equations to general coordinate system is not trivial. See Chapter II of Cary and Brizard[3] for details on how to do this[4]. Here I only state the final result. The mapping to a general coordinate system z^i, $i = 1, \ldots, 2n$ is written as

$$x^i = x^i(z^1, \ldots, z^{2n}) \tag{38}$$

$$v^i = v^i(z^1, \ldots, z^{2n}) \tag{39}$$

for $i = 1, \ldots, n$. Often, we will (confusingly) let the functions x^i run from $1, \ldots, 2n$ with the identification $x^{n+i} = v^i$, for $i = 1, \ldots, n$. In these generalized coordinates, the Poisson bracket transforms to

$$\{f, g\} = \frac{\partial f}{\partial z^i} \Pi^{ij} \frac{\partial g}{\partial z^j} \tag{40}$$

The anti-symmetric Poisson tensor Π^{ij} plays a fundamental role in the theory of Hamiltonian systems and has an analogous importance to the metric tensor $\mathbf{g}$. The geometry governed by the

[4]The notation in Cary and Brizard[3] is different than used here, and a potential source of confusion. What I am calling $\mathbf{M} = \mathbf{D}^{-1}$ they call $\mathbf{D}$ instead. Hence, in their notation the volume element is $\det(\mathbf{D}^{-1})$ instead of $\det(\mathbf{D})$. I guess the reason they do this is they are only interested in symplectic geometry, in which the dual vectors (rows of $\mathbf{M}$) play an important role, rather than Riemannian geometry, in which the metric plays an important role.

Poisson tensor is called *symplectic geometry* in contrast to the geometry governed by the metric tensor, called *Riemannian geometry*. The Poisson tensor is given by

$$\mathbf{\Pi} = \mathbf{M}\boldsymbol{\sigma}\mathbf{M}^T \tag{41}$$

where recall from the previous section that $\mathbf{M} = \mathbf{D}^{-1}$, where $\mathbf{D}$ is the Jacobian matrix of the transform, is made up the dual vectors as its rows. The matrix $\boldsymbol{\sigma}$ is the fundamental symplectic matrix (playing a role analogous to the Cartesian metric δ_{ij} in Riemannian geometry):

$$\boldsymbol{\sigma} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \tag{42}$$

where $\mathbf{I}$ is $n \times n$ unit matrix. If a transform is such that $\mathbf{\Pi} = \boldsymbol{\sigma}$ then it is called a *canonical transform*. In general, though, not all useful transforms are canonical and often it is beneficial to use non-canonical transforms instead. A different way of writing Eq. (41) is

$$\Pi^{ij} = \nabla_{\mathbf{x}} z^i \cdot \nabla_{\mathbf{v}} z^j - \nabla_{\mathbf{v}} z^i \cdot \nabla_{\mathbf{x}} z^j. \tag{43}$$

This follows from the fact that the dual vectors (coordinate gradients) are the rows of the $\mathbf{M}$ matrix. We can write this expression in yet another way by introducing the set of dual vectors

$$\mathbf{e}^i = \nabla_{\mathbf{x}} z^i \tag{44}$$

$$\tilde{\mathbf{e}}^i = \nabla_{\mathbf{v}} z^i. \tag{45}$$

In terms of these, the Poisson tensor can be written as

$$\Pi^{ij} = \mathbf{e}^i \cdot \tilde{\mathbf{e}}^j - \mathbf{e}^j \cdot \tilde{\mathbf{e}}^i. \tag{46}$$

Note that these dual vectors can be simply read off the rows of the $\mathbf{M}$ matrix. If we now introduce the vector gradient operators

$$\partial \equiv \mathbf{e}^i \frac{\partial}{\partial z^i} \tag{47}$$

$$\tilde{\partial} \equiv \tilde{\mathbf{e}}^i \frac{\partial}{\partial z^i} \tag{48}$$

then the Poisson bracket can be represented in the compact form

$$\{f, g\} = \partial f \cdot \tilde{\partial} g - \partial g \cdot \tilde{\partial} f. \tag{49}$$

Notice that this has the same structure as the canonical Poisson bracket defined in Eq. (37), with $\nabla_{\mathbf{x}}$ replaced by ∂ and $\nabla_{\mathbf{v}}$ replaced by $\tilde{\partial}$. In terms of these new derivative operators, the Hamiltonian system Eq. (36) can be written in the elegant form

$$\frac{\partial f}{\partial t} + \tilde{\partial} H \cdot \partial f - \partial H \cdot \tilde{\partial} f = 0. \tag{50}$$

Notice that this form of the equation is completely *coordinate independent*. Once the gradient operators are determined in any desired coordinate system, this expression can be used to immediately write down the equations of motion.

We can write the Hamiltonian system Eq. (36) in conservative form as follows:

$$\frac{\partial f}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z_j}\left(\mathcal{J}\alpha^j f\right) = 0 \tag{51}$$

where α^i is the *advection velocity* in phase-space:

$$\alpha^i \equiv \dot{z}^i = \{z^i, H\} = \Pi^{ij}\frac{\partial H}{\partial z^j}. \tag{52}$$

In terms of the gradient operators defined above, the advection velocity in phase-space can instead be written as

$$\alpha^i = \mathbf{e}^i \cdot \tilde{\partial} H - \tilde{\mathbf{e}}^i \cdot \partial H. \tag{53}$$

Note that the advection is phase-space is *incompressible*, that is

$$\frac{1}{\mathcal{J}}\frac{\partial}{\partial z_j}\left(\mathcal{J}\alpha^j\right) = 0. \tag{54}$$

Using Eq. (40) we can also write Eq. (36) in the non-conservative form:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z_j}\Pi^{jk}\frac{\partial H}{\partial z_k} = \frac{\partial f}{\partial t} + \alpha^j \frac{\partial f}{\partial z_j} = 0. \tag{55}$$

For example, in Binney and Termaine[2] they usually write the collisionless Boltzmann equation (as Vlasov equation is referred to in the self-gravitating systems literature) in this form, as it is convenient for analytical work.

3.1 Example I: 2D Incompressible Euler Equations

As a first example, consider the 2D incompressible Euler equations, for the evolution of $f(x, y, t)$ with the Poisson bracket

$$\{f, g\} = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x} \tag{56}$$

with the Hamiltonian given by $H(x, y) = \phi(x, y)$, where $\nabla^2\phi = -f$. We will transform this equation to (r, θ) coordinates:

$$x = r\cos\theta, \qquad y = r\sin\theta. \tag{57}$$

The Jacobian matrix of this transform and its inverse are:

$$\mathbf{D} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix} \tag{58}$$

The metric tensor can be read off from $ds^2 = dr^2 + r^2 d\theta^2$ and, $\mathcal{J} = r$. From Eq. (41) the Poisson tensor in these coordinates is

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 1/r \\ -1/r & 0 \end{pmatrix}. \tag{59}$$

From this, the characteristic velocity can be computed as

$$\dot{r} = \{r, H\} = \frac{1}{r}\frac{\partial \phi}{\partial \theta} \tag{60}$$

$$\dot{\theta} = \{\theta, H\} = -\frac{1}{r}\frac{\partial \phi}{\partial r}. \tag{61}$$

3.2 Example II: Vlasov-Poisson Equations

A much more interesting example is the Vlasov-Poisson equation for the evolution of $f(\mathbf{x}, \mathbf{v}, t)$ with the Poisson bracket given by Eq. (37), and the Hamiltonian:

$$H = \frac{1}{2}\mathbf{v}^2 + \phi(\mathbf{x}) \tag{62}$$

(I am dropping charge and mass factors here as they do not add anything to the discussion). The transformation to generalized coordinates z^i can be completely arbitrary and can mix the $(\mathbf{x}, \mathbf{v})$ in any invertible manner. For example, we can transform only the configuration-space, or only the velocity-space or both. Though the transformations can be completely arbitrary, often some choices of velocity-space transforms are suggested by the choice of the configuration-space transform. For example, let $\mathbf{e}_{c,i}$ be the tangent vectors in configuration-space and $\mathbf{e}^i_c$ be their duals. Then there are two "natural" choices of velocity coordinate transforms. The first is to choose w^i $(i = 1, 2, 3)$ such that

$$\mathbf{v} = \mathbf{e}_{c,i} w^i. \tag{63}$$

With this choice, the kinetic energy is

$$\frac{1}{2}\mathbf{v}^2 = \frac{1}{2}\mathbf{e}_{c,i} w^i \cdot \mathbf{e}_{c,j} w^j = \frac{1}{2} \underbrace{\mathbf{e}_{c,i} \cdot \mathbf{e}_{c,j}}_{h_{ij}} w^i w^j = \frac{1}{2} h_{ij} w^i w^j \tag{64}$$

where h_{ij} is the metric tensor of the configuration-space transform. We will call this "Choice NC" below.

The second choice is to choose instead w_i such that

$$\mathbf{v} = \mathbf{e}_c^i w_i. \tag{65}$$

With this choice, the kinetic energy is

$$\frac{1}{2}\mathbf{v}^2 = \frac{1}{2}\mathbf{e}_c^i w_i \cdot \mathbf{e}_c^j w_j = \frac{1}{2}\underbrace{\mathbf{e}_c^i \cdot \mathbf{e}_c^j}_{h^{ij}} w_i w_j = \frac{1}{2}h^{ij} w_i w_j \tag{66}$$

where h^{ij} is the inverse of metric tensor of the configuration-space transform. As it turns out, we can show that this latter choice results in a *canonical transform* and is hence particularly simple to deal with. We will call this "Choice C" below.

In the examples below we will use polar coordinates in configuration space. For this, the tangent vectors and their duals are

$$\mathbf{e}_r = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{67}$$

$$\mathbf{e}_\theta = -r\sin\theta\,\mathbf{e}_x + r\cos\theta\,\mathbf{e}_y \tag{68}$$

and

$$\mathbf{e}^r = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{69}$$

$$\mathbf{e}^\theta = -\frac{1}{r}\sin\theta\,\mathbf{e}_x + \frac{1}{r}\cos\theta\,\mathbf{e}_y \tag{70}$$

and the Jacobian of the configuration-space transform is $\mathcal{J}_c = r$. In all cases, the Poisson equation uses the Laplacian defined in Eq. (27), computed using the configuration-space metric.

Transforming only configuration space. As a first example, consider we only transform configuration space to polar coordinates (r, θ) (see Eq. (57) but keep the velocity space coordinates unchanged. The Jacobian matrix of the transform and its inverse are

$$\mathbf{D} = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 & 0 \\ \sin\theta & r\cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta/r & \cos\theta/r & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{71}$$

with the Jacobian $\mathcal{J} = r$. From this, we can write down the dual vectors as

$$\mathbf{e}^1 = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{72}$$

$$\mathbf{e}^2 = -\frac{1}{r}\sin\theta\,\mathbf{e}_x + \frac{1}{r}\cos\theta\,\mathbf{e}_y \tag{73}$$

and $\mathbf{e}^3 = \mathbf{e}^4 = 0$, $\tilde{\mathbf{e}}^1 = \tilde{\mathbf{e}}^2 = 0$, $\tilde{\mathbf{e}}^3 = \mathbf{e}_x$ and $\tilde{\mathbf{e}}^4 = \mathbf{e}_y$. From these, we can compute the gradient operators as

$$\partial = \mathbf{e}_x \left(\cos\theta \frac{\partial}{\partial r} - \frac{1}{r}\sin\theta \frac{\partial}{\partial \theta} \right) + \mathbf{e}_y \left(\sin\theta \frac{\partial}{\partial r} + \frac{1}{r}\cos\theta \frac{\partial}{\partial \theta} \right) \tag{74}$$

and

$$\tilde{\partial} = \mathbf{e}_x \frac{\partial}{\partial v_x} + \mathbf{e}_y \frac{\partial}{\partial v_y}. \tag{75}$$

Having obtained the gradient operators, we can write the Vlasov equation directly from Eq. (50). We can also, instead, compute first compute the Poisson tensor:

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta/r & \cos\theta/r \\ -\cos\theta & \sin\theta/r & 0 & 0 \\ -\sin\theta & -\cos\theta/r & 0 & 0 \end{pmatrix}. \tag{76}$$

Once we have the Poisson tensor we can write the characteristic velocities in configuration space as

$$\dot{r} = v_x \cos\theta + v_y \sin\theta \tag{77}$$

$$\dot{\theta} = -\frac{v_x}{r}\sin\theta + \frac{v_y}{r}\cos\theta. \tag{78}$$

and in velocity space as

$$\dot{v}_x = -\frac{\partial\phi}{\partial r}\cos\theta + \frac{1}{r}\frac{\partial\phi}{\partial\theta}\sin\theta \tag{79}$$

$$\dot{v}_y = -\frac{\partial\phi}{\partial r}\sin\theta - \frac{1}{r}\frac{\partial\phi}{\partial\theta}\cos\theta. \tag{80}$$

It is left as a (simple) exercise for the reader to show that using Eqns. 74 and 75 in Eq. (50) gives the same expressions as above.

Transforming phase-space, choice NC. In this example, we will use the velocity transform given by Eq. (63), which, for polar coordinates becomes:

$$v_x = w_r \cos\theta - w_\theta r \sin\theta \tag{81}$$

$$v_y = w_r \sin\theta + w_\theta r \cos\theta. \tag{82}$$

With this choice, the Hamiltonian becomes

$$H = \frac{1}{2}(w_r^2 + r^2 w_\theta^2) + \phi(r,\theta). \tag{83}$$

We can calculate the Jacobian of the transform and its inverse, and then compute the dual vectors as

$$\mathbf{e}^1 = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{84}$$

$$\mathbf{e}^2 = -\frac{1}{r}\sin\theta\,\mathbf{e}_x + \frac{1}{r}\cos\theta\,\mathbf{e}_y \tag{85}$$

$$\mathbf{e}^3 = -\sin\theta w_\theta\,\mathbf{e}_x + \cos\theta w_\theta\,\mathbf{e}_y \tag{86}$$

$$\mathbf{e}^4 = \left(\sin\theta\frac{w_r}{r^2} - \cos\theta\frac{w_\theta}{r}\right)\,\mathbf{e}_x - \left(\sin\theta\frac{w_\theta}{r} + \cos\theta\frac{w_r}{r^2}\right)\,\mathbf{e}_y. \tag{87}$$

and $\tilde{\mathbf{e}}^1 = \tilde{\mathbf{e}}^2 = 0$, $\tilde{\mathbf{e}}^3 = \mathbf{e}^1$ and $\tilde{\mathbf{e}}^4 = \mathbf{e}^2$. From this, using Eq. (46) we can compute

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/r^2 \\ -1 & 0 & 0 & 2w_\theta/r \\ 0 & -1/r^2 & -2w_\theta/r & 0 \end{pmatrix}. \tag{88}$$

From the Poisson tensor we can calculate the characteristic velocities in configuration space as simply $\dot{r} = w_r$, $\dot{\theta} = w_\theta$ and in velocity space as

$$\dot{w}_r = r w_\theta^2 - \frac{\partial \phi}{\partial r} \tag{89}$$

$$\dot{w}_\theta = -\frac{2 w_r w_\theta}{r} - \frac{1}{r^2}\frac{\partial \phi}{\partial \theta}. \tag{90}$$

With these definitions, the Vlasov equation in these coordinates becomes

$$\frac{\partial f}{\partial t} + w_r \frac{\partial f}{\partial r} + w_\theta \frac{\partial f}{\partial \theta} + \left(r w_\theta^2 - \frac{\partial \phi}{\partial r} \right) \frac{\partial f}{\partial w_r} - \left(\frac{2 w_r w_\theta}{r} + \frac{1}{r^2}\frac{\partial \phi}{\partial \theta} \right) \frac{\partial f}{\partial w_\theta} = 0. \tag{91}$$

Transforming phase-space, choice C. Now consider the velocity transform given by Eq. (65), which, for polar coordinates becomes:

$$v_x = \cos\theta w_r - \frac{w_\theta}{r}\sin\theta \tag{92}$$

$$v_y = \sin\theta w_r + \frac{w_\theta}{r}\cos\theta. \tag{93}$$

With this choice, the Hamiltonian becomes

$$H = \frac{1}{2r^2}(r^2 w_r^2 + w_\theta^2) + \phi(r,\theta). \tag{94}$$

In these coordinates one can show that, though I skip the details, miraculously, the Poisson tensor is simply $\mathbf{\Pi} = \boldsymbol{\sigma}$. That is, the coordinate transform we have chosen is *canonical*. The characteristic velocity in configuration-space is now $\dot{r} = w_r$ and $\dot{\theta} = w_\theta/r^2$ and in velocity space is

$$\dot{w}_r = \frac{w_\theta^2}{r^3} - \frac{\partial \phi}{\partial r} \tag{95}$$

$$\dot{w}_\theta = -\frac{\partial \phi}{\partial \theta}. \tag{96}$$

Finally, using these definitions we can write the Vlasov equation as

$$\frac{\partial f}{\partial t} + w_r \frac{\partial f}{\partial r} + \frac{w_\theta}{r^2}\frac{\partial f}{\partial \theta} + \left(\frac{w_\theta^2}{r^3} - \frac{\partial \phi}{\partial r} \right) \frac{\partial f}{\partial w_r} - \frac{\partial \phi}{\partial \theta}\frac{\partial f}{\partial w_\theta} = 0. \tag{97}$$

As verification, this is identical to Eq. (4.12) in[2]. Note that Binney and Tremaine do not give the general transformation formalism presented here, but only deal with canonical coordinates.

Bibliographical Notes

There are a lot of books on tensor calculus and Riemannian geometry. The book by Lovelock and Rund[4] is very good and available in a cheap Dover edition. A primer on tensor calculus for engineers and physicists is Clarke[1] and is a good handy reference to have. Chapter 1 of Thorne and Blanford[5] is a good introduction to tensors without use of coordinates. Incidentally, this is an excellent text-book, with very good introduction to all aspects of classical physics (except classical mechanics), including plasma physics and MHD. I recommend you get it (though it is pricey). When you are not reading it you can use it to exercise or, in case of a break-in you can toss it on the intruder's head.

References

[1] A Primer on Tensor Calculus. `https://www.ap.smu.ca/~dclarke/home/documents/byDAC/tprimer.pdf`. Accessed: 2020-03-20.

[2] James Binney and Scott Tremaine. *Galactic Dynamics*. Princeton University Press, 2008.

[3] John Cary and Alain Brizard. Hamiltonian theory of guiding-center motion. *Reviews of Modern Physics*, 81(2):693–738, May 2009.

[4] David Lovelock and Hanno Rund. *Tensors, Differential Forms, and Variational Principles*. Dover, 1989.

[5] Kip S. Thorne and Roger D. Blanford. *Modern Classical Physics*. Princeton University Press, 2017.