A Tutorial on Vectors, Tensors and All That

Ammar H. Hakim

Contents

1 Vector Space

A vector can be interpreted geometrically as a *directed line segment*. For example, $\nearrow$. We will represent the set of all vectors as $\mathcal{V}$ and denote vectors in this space with bold letters. The key property of vector space is that they can be rescaled by a real scalar, and added using the usual geometrical method. If $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, then $\alpha\mathbf{a} \in \mathcal{V}$ and $\mathbf{a}+\mathbf{b} \in \mathcal{V}$, where α is a real scalar. We will also assume that we can compute the *inner-product* or *dot-product* $\mathbf{a} \cdot \mathbf{b}$ of two vectors. Using the dot-product we can define the *norm* of the vector as $\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}}$. In fact, the process of computing dot-products is so important that we will also represent it as the bilinear function $\mathbf{g}(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$. This function $\mathbf{g}$ is called the *metric tensor*. We will discuss *tensors* later in this tutorial, however, for now it suffices to say that tensors are multi-linear functions of vectors. This tutorial shows how to formulate physics in a *geometrical* fashion, using vectors and other object built from them. The goal is to write physical laws in geometrical language and leave the introduction of coordinates till the last possible moment, just before we need to solve a specific problem at hand. The geometrical formulation of physics is fundamental: after all, our physical world is independent of our *representation* of it using coordinates.

2 Coordinate Transforms

Consider a smooth, invertible mapping from point $z^1, \ldots, z^n$ in $\mathbb{R}^n$ to a position-vector $\mathbf{x}$:

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

From this coordinate mapping we can define a set of basis vectors

$$\mathbf{e}_i \equiv \frac{\partial \mathbf{x}}{\partial z^i}. \tag{2}$$

Such a set of vectors generated from the coordinate mappings are called *coordinate basis*. Sometimes, coordinate basis are denoted by $\mathbf{e}_i(\mathbf{x})$ to indicate that they are a *vector fields* defined over a patch of space.

We can represent any vector in $\mathcal{V}$ in terms of these coordinate basis. Of course, the basis vectors at one point p will not be the same as the basis vectors at another point q. Given any $\mathbf{a} \in \mathcal{V}$

we can compute the components of this vector in the chosen basis by using our dot-product (or metric-tensor) as $a_i = \mathbf{a} \cdot \mathbf{e}_i = \mathbf{g}(\mathbf{a}, \mathbf{e}_i)$.

Given an arbitrary basis set $\mathbf{e}_i$ at p, we can construct their *reciprocal basis-vectors*, $\mathbf{e}^i$, using the implicit definition

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta^i{}_j, \tag{3}$$

where $\delta^i{}_j$ is the Kronecker delta. Notice that we have again used the dot-product/metric in constructing the reciprocal basis vectors. Also, observe that the the reciprocal basis-vectors span the *same vector space* $\mathcal{V}$ as the basis-vectors themselves. Hence, for any $\mathbf{a} \in \mathcal{V}$ can also compute its components in the reciprocal basis as $a^i = \mathbf{a} \cdot \mathbf{e}^i = \mathbf{g}(\mathbf{a}, \mathbf{e}^i)$. Given the components of a vector in a given basis or reciprocal basis vectors, we can write

$$\mathbf{a} = a^i \mathbf{e}_i = a_i \mathbf{e}^i. \tag{4}$$

Here, and throughout the tutorial we will assume summation over repeated indices.

For coordinate basis, there is another way to express the reciprocal basis as gradients of the *inverse* mapping

$$\mathbf{e}^i = \nabla_{\mathbf{x}} z^i(\mathbf{x}). \tag{5}$$

In general, though, the reciprocal basis are computed most easily using the definition Eq. (3).

From the coordinate basis we can construct other basis sets also. For example, the components of a vector $\mathbf{u}$ in the basis $\mathbf{e}^i$ and $\mathbf{e}_i$ may not have the same units as the physical quantity (say velocity) the vector represents. To remedy this, we can instead work with *normalized basis-vectors*, $\hat{\mathbf{e}}_i$ defined as

$$\hat{\mathbf{e}}_i \equiv \frac{\mathbf{e}_{\underline{i}}}{\|\mathbf{e}_{\underline{i}}\|}. \tag{6}$$

No summation is implied over underlined indices: we will treat underlined repeated indices as being the same, without sums[1]. A vector $\mathbf{u}$ can be written in these basis as

$$\mathbf{u} = \hat{u}^i \hat{\mathbf{e}}_i \tag{7}$$

and the components $\hat{u}^i = u^{\underline{i}} \|\mathbf{e}_{\underline{i}}\|$ will have the same units as the underlying physical quantity. Geometrically, the components $\hat{u}^i$ are the components of $\mathbf{u}$ along unit vectors parallel (locally) to the coordinate lines. Further useful representations can also be constructed. In particular, we can construct a *locally orthonormal* basis set from the set of vectors $\mathbf{e}_i$ by a Gram-Schmidt process. This set of vector can often be useful as they are locally Cartesian and physics equations take particularly simple form in these coordinates.

For now, we will assume that the space we are dealing with is Euclidean. In this space we will denote unit vectors in this coordinate system as $\boldsymbol{\sigma}_i$. At each point p in Euclidean space the tangent vector space $\mathcal{V}$ can also use the same basis-vectors $\boldsymbol{\sigma}_i$, which also serve as the reciprocals basis. Consider the following example.

[1] For example $a^{ij} b_{\underline{j}} c_{\underline{j}} = \sum_j a^{ij} b_j c_j$.

Example 1. In 2D, consider the mapping from polar coordinates:

$$\mathbf{x} = r\cos\theta\,\boldsymbol{\sigma}_1 + r\sin\theta\,\boldsymbol{\sigma}_2. \tag{8}$$

From this the tangent vectors (coordinate basis) are

$$\mathbf{e}_r = \frac{\partial \mathbf{x}}{\partial r} = \cos\theta\,\boldsymbol{\sigma}_1 + \sin\theta\,\boldsymbol{\sigma}_2 \tag{9}$$
$$\mathbf{e}_\theta = \frac{\partial \mathbf{x}}{\partial \theta} = -r\sin\theta\,\boldsymbol{\sigma}_1 + r\cos\theta\,\boldsymbol{\sigma}_2. \tag{10}$$

Note this mapping breaks down at $r = 0$. The metric tensor components are $\mathbf{g}(\mathbf{e}_r, \mathbf{e}_r) = \mathbf{e}_r \cdot \mathbf{e}_r = 1$, $\mathbf{g}(\mathbf{e}_r, \mathbf{e}_\theta) = \mathbf{e}_r \cdot \mathbf{e}_\theta = 0$ and $\mathbf{g}(\mathbf{e}_\theta, \mathbf{e}_\theta) = \mathbf{e}_\theta \cdot \mathbf{e}_\theta = r^2$. The reciprocal vectors can be now be computed as

$$\mathbf{e}^r = \cos\theta\,\boldsymbol{\sigma}_1 + \sin\theta\,\boldsymbol{\sigma}_2 \tag{11}$$
$$\mathbf{e}^\theta = -\frac{1}{r}\sin\theta\,\boldsymbol{\sigma}_1 + \frac{1}{r}\cos\theta\,\boldsymbol{\sigma}_2 \tag{12}$$

Example 2. With polar coordinates defined in the above example, let $\mathbf{u} = u_x\boldsymbol{\sigma}_1 + u_y\boldsymbol{\sigma}_2$ be a vector. Then we can get the components $u^r = \mathbf{u} \cdot \mathbf{e}^r$ and $u^\theta = \mathbf{u} \cdot \mathbf{e}^\theta$ as

$$u^r = u_x\cos\theta + u_y\sin\theta \tag{13}$$
$$u^\theta = -\frac{u_x}{r}\sin\theta + \frac{u_y}{r}\cos\theta. \tag{14}$$

Similarly, $u_r = \mathbf{u} \cdot \mathbf{e}_r$ and $u_\theta = \mathbf{u} \cdot \mathbf{e}_\theta$ to get

$$u_r = u_x\cos\theta + u_y\sin\theta \tag{15}$$
$$u_\theta = -u_x r\sin\theta + u_y r\cos\theta. \tag{16}$$

Finally

$$u^2 = \left(u^r\right)^2 + r^2\left(u^\theta\right)^2 = u_r^2 + \frac{1}{r^2}u_\theta^2. \tag{17}$$

Normalizing the tangent vectors we get

$$\hat{\mathbf{e}}_r = \cos\theta\,\boldsymbol{\sigma}_1 + \sin\theta\,\boldsymbol{\sigma}_2 \tag{18}$$
$$\hat{\mathbf{e}}_\theta = -\sin\theta\,\boldsymbol{\sigma}_1 + \cos\theta\,\boldsymbol{\sigma}_2. \tag{19}$$

In terms of these, the components of a vector $\mathbf{u}$ can be written as

$$\hat{u}^r = u_x\cos\theta + u_y\sin\theta \tag{20}$$
$$\hat{u}^\theta = -u_x\sin\theta + u_y\cos\theta. \tag{21}$$

Notice that $\hat{u}^r$ and $\hat{u}^\theta$ have the same units as the physical vectors they represent. ■

Exercise 1. In 3D consider the mapping from spherical coordinates

$$\mathbf{x} = r\sin\theta\cos\phi\,\boldsymbol{\sigma}_1 + r\sin\theta\sin\phi\,\boldsymbol{\sigma}_2 + r\cos\theta\boldsymbol{\sigma}_3. \tag{22}$$

Compute the tangent vectors, their reciprocals and the metric tensor for this mapping.

Exercise 2. Consider the 2D non-orthogonal but linear mapping

$$\mathbf{x} = \left(z^1 + z^2\cos\alpha\right)\boldsymbol{\sigma}_1 + z^2\sin\alpha\,\boldsymbol{\sigma}_2, \tag{23}$$

where α is a constant angle $0 < \alpha < \pi$. Compute the tangent vectors, their reciprocals and the metric tensor for this mapping.

3 Basis Vectors and Dual Basis in Three Dimensions

In three dimensions, given a basis set $\mathbf{e}_i$, $i = 1, 2, 3$, we can explicitly compute the duals. To do this first define

$$\mathcal{J} \equiv \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)\,. \tag{24}$$

Geometrically, $\mathcal{J}$ is the volume (modulo a sign) of the parallelpiped formed by $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$. We will assume that the basis are arranged such that $\mathcal{J} > 0$. Now, let us guess that the dual basis $\mathbf{e}^1$ is given by

$$\mathbf{e}^1 = C\mathbf{e}_2 \times \mathbf{e}_3. \tag{25}$$

Take the dot-product of this expression with $\mathbf{e}_1$ to see that $\mathbf{e}^1 \cdot \mathbf{e}_1 = 1 = C\mathcal{J}$. This process allows us to compute the explicit expressions for the duals as

$$\mathbf{e}^1 = \frac{\mathbf{e}_2 \times \mathbf{e}_3}{\mathcal{J}} \tag{26a}$$

$$\mathbf{e}^2 = \frac{\mathbf{e}_3 \times \mathbf{e}_1}{\mathcal{J}} \tag{26b}$$

$$\mathbf{e}^3 = \frac{\mathbf{e}_1 \times \mathbf{e}_2}{\mathcal{J}}. \tag{26c}$$

As is asked to prove in the next exercise, we can show that

$$\mathbf{e}^1 \cdot \left(\mathbf{e}^2 \times \mathbf{e}^3\right) = \frac{1}{\mathcal{J}}. \tag{27}$$

Using this, we can also show that

$$\mathbf{e}_1 = \mathcal{J}\mathbf{e}^2 \times \mathbf{e}^3 \tag{28a}$$

$$\mathbf{e}_2 = \mathcal{J}\mathbf{e}^3 \times \mathbf{e}^1 \tag{28b}$$

$$\mathbf{e}_3 = \mathcal{J}\mathbf{e}^1 \times \mathbf{e}^2. \tag{28c}$$

Exercise 3. Show that Eq. (27) is indeed true [2]

4 Tensor Products and Tensors

Given two or more vectors in $\mathcal{V}$, we will denote their *tensor product* with the $\otimes$ symbol. For example, given $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ we can write their tensor-product as $\mathbf{u} \otimes \mathbf{v}$ The tensor-product creates a *multilinear mapping* from n vectors, where n is the number of vectors in the product, to a real number. Given $\mathbf{u} \otimes \mathbf{v}$, the mapping is $\mathbf{u} \otimes \mathbf{v} : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, and can evaluate it for the vectors $\mathbf{a}$ and $\mathbf{b}$ as follows

$$(\mathbf{u} \otimes \mathbf{v})(\mathbf{a}, \mathbf{b}) = (\mathbf{u} \cdot \mathbf{a})(\mathbf{v} \cdot \mathbf{b}). \tag{29}$$

A tensor of rank n is a *multilinear function* that takes n vectors and maps them to a real number. For example, a second order tensor $\mathbf{T}(\mathbf{a}, \mathbf{b})$ will take two input vectors ($\mathbf{a}$ and $\mathbf{b}$ in this case) and produce a single scalar. As the mapping is multilinear we have

$$\mathbf{T}(\alpha \mathbf{a} + \delta \mathbf{d}, \mathbf{b}) = \alpha \mathbf{T}(\mathbf{a}, \mathbf{b}) + \delta \mathbf{T}(\mathbf{d}, \mathbf{b}). \tag{30}$$

In this sense, a scalar is a rank-0 tensor, that simply evaluates to itself. A vector $\mathbf{u}$ is a rank-1 tensor mapping an input vector $\mathbf{a}$ to

$$\mathbf{u}(\mathbf{a}) = \mathbf{u} \cdot \mathbf{a}. \tag{31}$$

Defined in this manner, rank-n tensors (including vectors) are *geometric* quantities, hence independent of the basis vectors used to represent them.

Now, if we feed a vector with one of the basis $\mathbf{e}_i$ we will get

$$\mathbf{u}(\mathbf{e}_i) = \mathbf{u} \cdot \mathbf{e}_i = u_i \tag{32}$$

and

$$\mathbf{u}(\mathbf{e}^i) = \mathbf{u} \cdot \mathbf{e}^i = u^i. \tag{33}$$

Hence, with the basis as an input, the vector mapping produces the *component* of the vector along that basis. Analogously, we will *define* the components of a higher-rank tensor as the real numbers produced when the input vectors are the basis. So

$$T_{ij} \equiv \mathbf{T}(\mathbf{e}_i, \mathbf{e}_j) \tag{34}$$

[2] You will need to use the identities

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c})$$
$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{d}]\mathbf{c} - [(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}]\mathbf{d}.$$

and

$$T^{ij} \equiv \mathbf{T}(\mathbf{e}^i, \mathbf{e}^j). \tag{35}$$

As tensors are multilinear mapping, in a specific basis we can write them as linear combinations of the tensor products of the selected basis. For example,

$$\mathbf{T} = T^{ij}\mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{T}(\mathbf{e}^i, \mathbf{e}^j)\mathbf{e}_i \otimes \mathbf{e}_j = T_{ij}\mathbf{e}^i \otimes \mathbf{e}^j = \mathbf{T}(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}^i \otimes \mathbf{e}^j. \tag{36}$$

Hence, using the components, we can write an explicit formula for the evaluation of the multilinear form in terms of its components as

$$\mathbf{T}(\mathbf{a}, \mathbf{b}) = T^{ij}(\mathbf{a} \cdot \mathbf{e}_i)(\mathbf{b} \cdot \mathbf{e}_j) = T_{ij}(\mathbf{a} \cdot \mathbf{e}^i)(\mathbf{b} \cdot \mathbf{e}^j). \tag{37}$$

We can also compute the *partial* evaluation of the tensor by filling up one or more of its slots with vectors. The resulting function (taking fewer input parameters) is also a tensor, however of lower rank. For example $\mathbf{T}(\mathbf{a}, \_)$ results in a vector (rank-1 tensor). In a specific representation

$$\mathbf{T}(\mathbf{a}, \_) = T^{ij}\mathbf{a} \cdot (\breve{\mathbf{e}}_i \otimes \mathbf{e}_j) = T^{ij}(\mathbf{a} \cdot \mathbf{e}_i)\mathbf{e}_j \tag{38}$$

where we have used the "breve" marker on the $\mathbf{e}_i$ to indicate which of the vectors making up the tensor product the dot-product must be taken with. Similarly,

$$\mathbf{T}(\_, \mathbf{a}) = T^{ij}\mathbf{a} \cdot (\mathbf{e}_i \otimes \breve{\mathbf{e}}_j) = T^{ij}\mathbf{e}_i(\mathbf{a} \cdot \mathbf{e}_j). \tag{39}$$

Of course, we do not need to use the tangent or their reciprocals to represent tensors. As they are geometric objects, any basis will do, for example, the normalized tangent vectors. Once the components are known in one set of basis, we can simply compute them by evaluating, for example,

$$\mathbf{T}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) \equiv \hat{T}_{ij} = T_{mn}(\hat{\mathbf{e}}_i \cdot \mathbf{e}^m)(\hat{\mathbf{e}}_j \cdot \mathbf{e}^n). \tag{40}$$

Exercise 4. A second-order tensor $\mathbf{S}$ is called *symmetric* if $\mathbf{S}(\mathbf{a}, \mathbf{b}) = \mathbf{S}(\mathbf{b}, \mathbf{a})$ and *anti-symmetric* if $\mathbf{S}(\mathbf{a}, \mathbf{b}) = -\mathbf{S}(\mathbf{b}, \mathbf{a})$. Show that the tensors $\mathbf{a} \otimes \mathbf{b} + \mathbf{b} \otimes \mathbf{a}$ is symmetric and $\mathbf{a} \otimes \mathbf{b} - \mathbf{b} \otimes \mathbf{a}$ is anti-symmetric.

The *trace* of a tensor product is a linear operator defined as, for example

$$\mathrm{Tr}(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}. \tag{41}$$

For products with more than two vectors, we need to indicate the pair of vectors on which the trace operator acts. For example,

$$\mathrm{Tr}(\breve{\mathbf{u}} \otimes \mathbf{v} \otimes \breve{\mathbf{w}}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v}. \tag{42}$$

Here we used the "breve" marker on the vectors we wish to participate in the trace. Just like partial evaluation, trace operator also is a rank-reducing operation: the resulting object has two ranks lower than the original tensor.

The dot-product operator can be extended to act on a pair of tensors. To define this, consider we want to compute the dot-product between a vector $\mathbf{u}$ and a second-order tensor $\mathbf{T}$. We first need to select the slot of the tensor we wish to take the product with. For example, we will denote the dot-product with the first slot as $\mathbf{u} \cdot \mathbf{T}(\breve{\_}, \_)$ and define this to be just the partial evaluation $\mathbf{T}(\mathbf{u}, \_)$. Similarly, $\mathbf{u} \cdot \mathbf{T}(\_, \breve{\_}) = \mathbf{T}(\_, \mathbf{u})$.

We have already defined the metric tensor is a special bilinear mapping

$$\mathbf{g}(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}. \tag{43}$$

From this definition we see that the partial evaluation of the metric tensor is particularly simple:

$$\mathbf{a} \cdot \mathbf{g}(\breve{\_}, \_) = \mathbf{g}(\mathbf{a}, \_) = \mathbf{a}. \tag{44}$$

We can compute the *components* of the metric tensor as

$$\mathbf{g}(\mathbf{e}_i, \mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j \equiv g_{ij}. \tag{45}$$

$$\mathbf{g}(\mathbf{e}^i, \mathbf{e}^j) = \mathbf{e}^i \cdot \mathbf{e}^j \equiv g^{ij}. \tag{46}$$

Now consider any vector $\mathbf{u}$ and write

$$u^i = \mathbf{u} \cdot \mathbf{e}^i = (\mathbf{u} \cdot \mathbf{e}_j)(\mathbf{e}^j \cdot \mathbf{e}^i) = g^{ij} u_j. \tag{47}$$

Similarly, we have

$$u_i = \mathbf{u} \cdot \mathbf{e}_i = (\mathbf{u} \cdot \mathbf{e}^j)(\mathbf{e}_j \cdot \mathbf{e}_i) = g_{ij} u^j. \tag{48}$$

This process is sometimes called *raising* and *lowering* of indices, and extends to tensors of any rank. In fact, we can also easily show that

$$\mathbf{e}^i = g^{ij} \mathbf{e}_j \tag{49}$$

$$\mathbf{e}_i = g_{ij} \mathbf{e}^j. \tag{50}$$

These expressions are useful to replace the basis vectors for their reciprocals (and vice-versa).

Exercise 5. Show that $\mathbf{g} = \mathbf{e}^j \otimes \mathbf{e}_j = \mathbf{e}_j \otimes \mathbf{e}^j$.

Notice that we must distinguish between a tensor $\mathbf{T}$, its *definition*, for example $\mathbf{T} = \mathbf{u} \otimes \mathbf{v}$ and it's *evaluation* $\mathbf{T}(\mathbf{a}, \mathbf{b})$. It is helpful to think of tensors as functions in a programming language: there also one must distinguish between the *name*, the *definition* and it's *evaluation*.

Finally, we remark that tensors are a very special, but important, class amongst general scalar-valued functions. In general, an arbitrary function $f : \mathcal{V} \to \mathbb{R}$ need not be linear. For example, $f(\mathbf{u}) = \mathbf{u} \cdot \mathbf{u}$ is a quadratic function of it's input vector, and hence is *not a tensor*.

NOTE: Need to bring out the key to the whole text: we treat tensors and other geometric objects as functions in a programming language.

5 Second-Order Tensors and Dyadic Products

Second-order tensors, also called *dyads*, appear frequently in mathematical physics. We shall define a *dyadic product* as follows. Let $\mathbf{T}$ be a second-order tensor and $\mathbf{u}$ and $\mathbf{v}$ be vectors. Then the dyadic product is denoted by the : symbol and is defined as

$$\mathbf{T} : \mathbf{u} \otimes \mathbf{v} \equiv \mathbf{T}(\mathbf{u}, \mathbf{v}). \tag{51}$$

In particular, if $\mathbf{T} = \mathbf{a} \otimes \mathbf{b}$ then

$$\mathbf{a} \otimes \mathbf{b} : \mathbf{u} \otimes \mathbf{v} = (\mathbf{a} \cdot \mathbf{u})(\mathbf{b} \cdot \mathbf{v}). \tag{52}$$

Now let $\mathbf{T}$ and $\mathbf{G}$ be two dyads. Then the above definitions can be used to write the dyadic product in term of the dyad representation in a particular basis as

$$\mathbf{T} : \mathbf{G} = \mathbf{T} : G_{ij}\mathbf{e}^i \otimes \mathbf{e}^j = \mathbf{T}(\mathbf{e}^i, \mathbf{e}^j)G_{ij} = T^{ij}G_{ij}. \tag{53}$$

This also shows that $\mathbf{T} : \mathbf{G} = \mathbf{G} : \mathbf{T}$. From the definitions it is also easy to see that the dyadic product with the metric-tensor is particularly simple

$$\mathbf{g} : \mathbf{T} = \mathrm{Tr}(\mathbf{T}). \tag{54}$$

From this it follows that $\mathbf{g} : \mathbf{g} = d$, where d is the dimension of the space.

The dyadic product is a *rank reducing* operator: it takes two second-order tensors and produces a scalar.

The dyadic product is bilinear, just like our definitions of tensors. In fact, the product can be instead recast as a special type of *"high-order" tensor* analogous to the metric tensor, if we allow tensor inputs to be tensors and output a tensor, instead of a scalar. Denote the dyad-product tensor as $\mathfrak{D}$. Then its signature is

$$\mathfrak{D} : (\mathcal{V} \times \mathcal{V} \to \mathbb{R}) \times (\mathcal{V} \times \mathcal{V} \to \mathbb{R}) \to \mathbb{R}. \tag{55}$$

This expression merely states that the $\mathfrak{D}$ tensor takes two dyads (themselves bilinear functions) and produces a scalar. Its definition can be constructed from the fundamental rule

$$\mathfrak{D}(\mathbf{a} \otimes \mathbf{b}, \mathbf{u} \otimes \mathbf{v}) \equiv (\mathbf{a} \cdot \mathbf{u})(\mathbf{b} \cdot \mathbf{v}). \tag{56}$$

Given multilinearity of $\mathfrak{D}$ this rule can be used to compute the dyadic product of any two tensors. The signature of $\mathfrak{D}$ is not as peculiar as it might look at first sight. For example, a second order tensor $\mathbf{T}$, that has a signature $\mathcal{V} \times \mathcal{V} \to \mathbb{R}$ has an equivalent signature

$$\mathbf{T} : (\mathcal{V} \to \mathbb{R}) \times (\mathcal{V} \to \mathbb{R}) \to \mathbb{R} \tag{57}$$

as the vector inputs in each of the tensor's slots are linear functions themselves.

Given a dyad we can construct from it a *linear transformation* that maps vectors to vectors. Consider the function $\mathtt{t}$ defined as as partial evaluation of a dyad:

$$\mathtt{t}(\mathbf{a}) = \mathbf{T}(\_, \mathbf{a}). \tag{58}$$

Clearly, this function has the signature

$$\mathtt{t} : \mathcal{V} \to \mathcal{V} \tag{59}$$

that is, it maps a vector to a vector. Due to the linearity of the dyad $\mathbf{T}$, the function $\mathtt{t}$ is also linear, and hence a tensor in our definition. Such linear transforms that map a vector to a vector are, in a sense, a coordinate-free way to represent matrices and hence play a central role in the development of linear algebra.

6 The Fundamental Derivative Operator

Given a coordinate system $(z^1, \ldots, z^n)$ and the corresponding coordinate basis, we can define the fundamental derivative operator as

$$\nabla = \mathbf{e}^i \frac{\partial}{\partial z^i}. \tag{60}$$

This operator allows taking various derivatives needed in writing down equations that arise in mathematical physics. In fact, everything one needs to do calculus is fully encoded in this operator. The use of the coordinate basis, as they are smooth, ensures that we can compute their partial derivatives with respect to the coordinates.

The operator itself is a *geometrical* quantity and can be represented in several different basis. For example, we can instead write

$$\nabla = g^{ij}\mathbf{e}_j \frac{\partial}{\partial z^i} = g^{ij}\|\mathbf{e}_{\underline{j}}\|\hat{\mathbf{e}}_{\underline{j}} \frac{\partial}{\partial z^i} \tag{61}$$

As with vectors, the last representation in normalized tangent vectors allows interpreting the gradient as having units of inverse length. Often, in many formularies (like the NRL Plasma Formulary), the components of the gradient and other operators are given in the *normalized* tangent vectors.

Example 3. Consider computing the gradient of a scalar function f in polar coordinates. Using the derivative operator we get

$$\nabla f = \mathbf{e}^r \frac{\partial f}{\partial r} + \mathbf{e}^\theta \frac{\partial f}{\partial \theta} = \mathbf{e}_r \frac{\partial f}{\partial r} + \frac{\mathbf{e}_\theta}{r^2} \frac{\partial f}{\partial \theta} = \hat{\mathbf{e}}_r \frac{\partial f}{\partial r} + \frac{\hat{\mathbf{e}}_\theta}{r} \frac{\partial f}{\partial \theta}. \tag{62}$$

The last form in normalized tangent-vectors is the one listed in the NRL Formulary.

Exercise 6. Compute the gradient of a scalar function f in the spherical coordinates defined in Exercise 1.

As the fundamental derivative operator behaves like a *vector*, we can use it as we would in place of vectors in various operations. We must take care, however, as in general, the product of the basis and partial derivative in Eq. (60) does not commute. For example, given a vector $\mathbf{u}$ we can compute its *divergence* $\nabla \cdot \mathbf{u}$ and the second-order tensor $\nabla \otimes \mathbf{u}$. Note that we can also define the somewhat odd-looking operator $\mathbf{u} \otimes \overleftarrow{\nabla}$, the left-pointing arrow indicating that the partial derivatives acts on the vector to the left:

$$\mathbf{u} \otimes \overleftarrow{\nabla} \equiv \frac{\partial \mathbf{u}}{\partial z^i} \otimes \mathbf{e}^i. \tag{63}$$

As one would expect, $\nabla \otimes \mathbf{u} + \mathbf{u} \otimes \overleftarrow{\nabla}$ is symmetric.

The tensor product with the gradient operator, and the divergence are connected via the trace operator

$$\nabla \cdot \mathbf{u} = \mathrm{Tr}(\nabla \otimes \mathbf{u}) = \mathrm{Tr}\Big(\mathbf{u} \otimes \overleftarrow{\nabla}\Big). \tag{64}$$

We can also compute the tensor-product of the gradient operator with a second or higher-order tensor by specifying which slot the tensor product acts on. For example $\nabla \otimes \mathbf{T}(\breve{\_}, \_)$. Similarly, the divergence of a second or higher-order tensor can be computed by specifying which slot the divergence acts on. For example the operations $\nabla \cdot \mathbf{T}(\breve{\_}, \_)$ and $\nabla \cdot \mathbf{T}(\_, \breve{\_})$ will result in different vectors, unless $\mathbf{T}$ is symmetric. For symmetric tensors, often the breve-marker on the slot is dropped, and one merely writes $\nabla \cdot \mathbf{T}$. However, in general, we need to be careful to indicate the slot of the tensor.

Note that the tensor-product of the gradient operator and a tensor is a *rank increasing* operation: the resulting tensor has one higher rank. For notational consistency, for a scalar function f we will write $\nabla \otimes f \equiv \nabla f$. This is also consistent with the rank increasing property of tensor-products: f is a scalar and ∇f is rank 1.

We can compute higher-order derivatives also. For example, we can compute the Laplacian of a scalar as

$$\nabla \cdot (\nabla f) = (\nabla \cdot \nabla) f \equiv \nabla^2 f. \tag{65}$$

Applying this to $\nabla \otimes \mathbf{u}$ we can define the Laplacian of a vector as

$$\nabla \cdot (\breve{\nabla} \otimes \mathbf{u}) = (\nabla \cdot \nabla)\mathbf{u} \equiv \nabla^2 \mathbf{u}. \tag{66}$$

An important result is that the divergence of the metric tensor is zero. To show this, we write $\mathbf{g} = \mathbf{e}_i \otimes \mathbf{e}^i$ and compute

$$\nabla \cdot \mathbf{g}(\breve{\_}, \_) = \mathbf{e}^k \cdot \frac{\partial}{\partial z^k}(\breve{\mathbf{e}}_i \otimes \mathbf{e}^i) = (\mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial z^k})\mathbf{e}^i + \underbrace{\mathbf{e}^k \cdot \mathbf{e}_i}_{\delta^k{}_i} \frac{\partial \mathbf{e}_i}{\partial z^k}. \tag{67}$$

Now using the chain-rule

$$\mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial z^k} = \underbrace{\frac{\partial}{\partial z^k}(\mathbf{e}^k \cdot \mathbf{e}_i)}_{0} - \mathbf{e}_i \cdot \frac{\partial \mathbf{e}_k}{\partial z^k}. \tag{68}$$

Hence,

$$\nabla \cdot \mathbf{g} = -(\mathbf{e}_i \cdot \frac{\partial \mathbf{e}_k}{\partial z^k})\mathbf{e}^i + \frac{\partial \mathbf{e}_k}{\partial z^k} = 0. \tag{69}$$

As we can also write $\mathbf{g} = \nabla \otimes \mathbf{x}$ the divergence-free condition also means that $\nabla^2 \mathbf{x} = 0$.

The dyadic product can also be applied to expressions involving the derivative operator. Consider $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be vectors. Then, in consistency with the definition of the dyadic product, we define

$$\mathbf{a} \otimes \mathbf{b} : \nabla \otimes \mathbf{c} = \mathbf{b} \cdot [\mathbf{a} \cdot \nabla \mathbf{c}]. \tag{70}$$

Note the order of the operations above which are chosen to be consistent with the order of the dot-products in Eq. (52). Also we have the useful relationship

$$\mathbf{g} : \nabla \otimes \mathbf{c} = \nabla \cdot \mathbf{c}. \tag{71}$$

Exercise 7. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors, and f be a scalar. Use the definition Eq. (60) to prove the following

$$\nabla \otimes (\mathbf{a} \otimes \mathbf{b}) = (\nabla \otimes \mathbf{a}) \otimes \mathbf{b} + \mathbf{a} \otimes (\nabla \otimes \mathbf{b}) \tag{72}$$

$$\nabla \otimes (f\mathbf{a}) = \nabla \otimes (\mathbf{a}f) = \nabla f \otimes \mathbf{a} + f\nabla \otimes \mathbf{a} \tag{73}$$

$$\nabla \cdot (\breve{\mathbf{a}} \otimes \mathbf{b}) = \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{a} \cdot \nabla)\mathbf{b} \tag{74}$$

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\nabla \otimes \breve{\mathbf{b}}) + \mathbf{b} \cdot (\nabla \otimes \breve{\mathbf{a}}). \tag{75}$$

7 Cross Product and Curl in Three Dimensions

The operations constructed from the tensor product of the derivative operator ∇ with tensors, as well as the trace of these apply to spaces of all dimensions. However, the *curl* (and the cross-product) operator is only defined in 3D space. A deep reason is that vectors are not enough for a complete description of geometry: in fact, just armed with vectors we cannot define plane segments or higher-dimensional objects. The expressions in this section are *restricted only to 3D space*.

The cross-product of two vectors is denoted by $\mathbf{b} \times \mathbf{c}$ and results in a vector. Here we will consider a different approach to the cross product by defining a *third*-order tensor $\boldsymbol{\varepsilon}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ as

$$\boldsymbol{\varepsilon}(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}). \tag{76}$$

The partial evaluation of this tensor with two of its slots filled in gives

$$\boldsymbol{\varepsilon}(\_, \mathbf{b}, \mathbf{c}) = \mathbf{b} \times \mathbf{c}, \tag{77}$$

that is, the resulting vector has the same components as the cross-product of $\mathbf{b}$ and $\mathbf{c}$. Hence, the second-order tensor that results from filling in the last slot, $\boldsymbol{\varepsilon}(\_, \_, \mathbf{c})$, has the property that

$$\mathbf{b} \cdot \boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{c}) = \boldsymbol{\varepsilon}(\_, \mathbf{b}, \mathbf{c}) = \mathbf{b} \times \mathbf{c}. \tag{78}$$

Although these manipulations all seem strange and esoteric, we have achieved something quite interesting: we have written the cross-product as a *dot*-product between a vector and a special second-order tensor.

Exercise 8. Show that $\boldsymbol{\varepsilon}$ is anti-symmetric in each pair of slots. Such tensors are called *totally anti-symmetric* tensors.

Exercise 9. Consider the components of $\boldsymbol{\varepsilon}$ along a basis, for example, $\boldsymbol{\varepsilon}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$. Show that all non-zero components are identical, except for a sign.

Exercise 10. Show that

$$\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B}) \cdot \nabla = -\mathbf{B} \times \nabla. \tag{79}$$

As with dot-product, we can take the cross-product between tensors. For example, consider the cross-product between a vector $\mathbf{u}$ and a second-order tensor $\mathbf{P}$. The cross-product is now defined as

$$\mathbf{u} \times \mathbf{P}(\breve{\_}, \_) = P^{mn} (\mathbf{u} \times \mathbf{e}_m) \otimes \mathbf{e}_n. \tag{80}$$

Notice that the resulting object is a second-order tensor.

Exercise 11. Let $\mathbf{P}$ be a symmetric second-order tensor. Show that $\mathbf{u} \times \mathbf{P}(\breve{\_}, \_) + \mathbf{u} \times \mathbf{P}(\_, \breve{\_})$ is a also symmetric second-order tensor.

Given a vector field $\mathbf{B}$ we denote its curl as usual as $\nabla \times \mathbf{B}$. Using the $\boldsymbol{\varepsilon}$ tensor above, we can write the curl instead as a *divergence* as

$$\nabla \times \mathbf{B} = \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B})]. \tag{81}$$

Exercise 12. Show that this is indeed true.

8 Select Equations of Mathematical Physics

In this section I list some common equations of mathematical physics. The goal is not a comprehensive catalog, but simply a selection of equations that have different types of terms that commonly appear in fluid and plasma applications. Using these it should be straightforward to write other equations in similar, coordinate independent forms. In manipulating terms if we only use geometric identities the resulting equations will also be geometric and coordinate independent. An attempt should be made to delay the introduction of coordinates to the last possible moments.

In the following, the geometric forms of equations are listed first, followed by a discussion on representation in specific coordinate systems. Also keep in mind that equations that involve the cross-product and curl operators are only restricted to 3D space.

8.1 The Advection-Diffusion Equation

The advection-diffusion equation is easy to write in a coordinate independent way:

$$\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{u} f) = \nu \nabla^2 f \tag{82}$$

where $\mathbf{u}$ is a specified (possibly time-dependent) velocity vector, and ν is the diffusion coefficient. If the diffusion is anisotropic it is described by the (usually symmetric) second-order *diffusion tensor* $\mathbf{D}$. In this case, the diffusion operator is instead $\nabla \cdot [\mathbf{D}(\nabla f, \_)]$.

$$\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{u} f) = \nabla \cdot [\mathbf{D}(\nabla f, \_)] \tag{83}$$

Note the order of operations here: we must *first* evaluate $\mathbf{D}(\nabla f, \_)$ and then take the divergence of the resulting vector.

Exercise 13. Show that if $\mathbf{D} = \nu \mathbf{g}$, where $\mathbf{g}$ is the metric tensor, then $\nabla \cdot [\mathbf{D}(\nabla f, \_)] = \nu \nabla^2 f$.

Exercise 14. Consider the diffusion tensor for heat-flux in a magnetized plasma. This is given by

$$\mathbf{D} = \kappa_\perp (\mathbf{g} - \mathbf{b} \otimes \mathbf{b}) + \kappa_\parallel \mathbf{b} \otimes \mathbf{b} + \kappa_\wedge \boldsymbol{\varepsilon}(\_, \mathbf{b}, \_) \tag{84}$$

where $\mathbf{b}$ is the unit vector along the magnetic-field, and $\kappa_{\parallel,\perp,\wedge}$ are the constant parallel, perpendicular and cross diffusion coefficients, respectively. Though $\kappa_{\parallel,\perp} \geq 0$, $\kappa_\wedge$ can be of any sign. Note that the first two term are symmetric, while the last term is antisymmetric. Show that

$$\nabla \cdot [\mathbf{D}(\nabla f, \_)] = \kappa_\perp \nabla^2 f + (\kappa_\parallel - \kappa_\perp) \nabla \cdot [\mathbf{b}\,(\mathbf{b} \cdot \nabla f)] - \kappa_\wedge \nabla \cdot (\mathbf{b} \times \nabla f). \tag{85}$$

8.2 Maxwell Equations

Maxwell equations were written in vector notation in the 19th century as

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0 \tag{86}$$

$$\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} = -\mu_0 \mathbf{J} \tag{87}$$

with the divergence relations $\nabla \cdot \mathbf{B} = 0$ and $\nabla \cdot \mathbf{E} = \varrho/\epsilon_0$. We can also write them in the somewhat less familiar divergence form instead

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{E})] = 0 \tag{88}$$

$$\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} - \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B})] = -\mu_0 \mathbf{J} \tag{89}$$

Taking the dot-product of Eq. (86) by $\mathbf{B}/\mu_0$ and Eq. (87) by $\mathbf{E}/\mu_0$ and adding the two resulting equations gives the evolution equation for electromagnetic energy

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{B}^2}{2\mu_0} + \frac{\epsilon_0 \mathbf{E}^2}{2}\right) + \frac{1}{\mu_0} \nabla \cdot (\mathbf{E} \times \mathbf{B}) = -\mathbf{E} \cdot \mathbf{J}. \tag{90}$$

In deriving this expression we have used the identity

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{B}. \tag{91}$$

Clearly, $\mathbf{S} = \mathbf{E} \times \mathbf{B}/\mu_0$ is the energy flux of the electromagnetic field.

We can also derive an equation for the evolution of electromagnetic momentum. To do this, take the cross-product of Eq. (86) by $\epsilon_0 \mathbf{E}$ and Eq. (87) by $\mathbf{B}/\mu_0$ and subtract the two resulting equations to get

$$\epsilon_0 \frac{\partial}{\partial t} (\mathbf{E} \times \mathbf{B}) + \epsilon_0 \mathbf{E} \times (\nabla \times \mathbf{E}) + \frac{1}{\mu_0} \mathbf{B} \times (\nabla \times \mathbf{B}) = -\mathbf{J} \times \mathbf{B}. \tag{92}$$

Now, for any vector $\mathbf{A}$ we have the identity

$$\mathbf{A} \times (\nabla \times \mathbf{A}) = \nabla \cdot \left[\frac{1}{2} \mathbf{A}^2 \mathbf{g}(\_, \breve{\_}) - \mathbf{A} \otimes \breve{\mathbf{A}}\right] + \mathbf{A}(\nabla \cdot \mathbf{A}). \tag{93}$$

Using this and the divergence relations for the fields we get the momentum evolution equation as

$$\epsilon_0 \frac{\partial}{\partial t} (\mathbf{E} \times \mathbf{B}) + \nabla \cdot \mathbf{T}(\_, \breve{\_}) = -\varrho_c \mathbf{E} - \mathbf{J} \times \mathbf{B}, \tag{94}$$

where $\mathbf{T}$ is the stress-energy tensor for the electromagnetic field defined by

$$\mathbf{T} = \left(\frac{\mathbf{B}^2}{2\mu_0} + \frac{\epsilon_0 \mathbf{E}^2}{2}\right) \mathbf{g} - \left(\epsilon_0 \mathbf{E} \otimes \mathbf{E} + \frac{1}{\mu_0} \mathbf{B} \otimes \mathbf{B}\right). \tag{95}$$

Exercise 15. Show that Eq. (93) is indeed true.

8.3 The Ideal MHD Equations

Ideal MHD is a pre-Maxwell theory in the sense that the displacement currents are ignored (hence, no electromagnetic waves) and the plasma is treated as a conducting fluid. The equations consist of the *continuity* equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{96}$$

and the *momentum equation*

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{\nabla p}{\rho} = \frac{1}{\mu_0 \rho} (\nabla \times \mathbf{B}) \times \mathbf{B} \tag{97}$$

where $\mathbf{B}$ is the magnetic field. The evolution of the field is determined by the *induction equation*

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0 \tag{98}$$

where $\mathbf{E} = -\mathbf{u} \times \mathbf{B}$ (ideal Ohm's Law). Finally, for an ideal plasma, the pressure evolves according to

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p = -\gamma p \nabla \cdot \mathbf{u}. \tag{99}$$

The momentum equation can be written in conservative form as follows. We use the identity Eq. (93) and the fact that $\nabla \cdot \mathbf{B} = 0$ to write

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = \nabla \cdot \left[-\frac{1}{2} \mathbf{B}^2 \mathbf{g}(\_, \breve{\_}) + \mathbf{B} \otimes \breve{\mathbf{B}} \right] \tag{100}$$

Using this expression and the continuity equation we can write the momentum equation in conservative form as

$$\frac{\partial}{\partial t} (\rho \mathbf{u}) + \nabla \cdot \mathbf{T}(\_, \breve{\_}) = 0 \tag{101}$$

where

$$\mathbf{T} = \rho \mathbf{u} \otimes \mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \mathbf{g} - \frac{1}{\mu_0} \mathbf{B} \otimes \mathbf{B}. \tag{102}$$

Note that $\mathbf{T}$ is a *symmetric* second-order tensor.

The induction equation can be written in divergence form as

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{u} \times \mathbf{B})] = 0. \tag{103}$$

To write this a little differently, use the identity $\mathbf{a}\times(\mathbf{u}\times\mathbf{B}) = (\mathbf{a}\cdot\mathbf{B})\mathbf{u} - (\mathbf{a}\cdot\mathbf{u})\mathbf{B}$, where $\mathbf{a}$ is an arbitrary vector to write

$$\boldsymbol{\varepsilon}(\_,\mathbf{a},\mathbf{u}\times\mathbf{B}) = \mathbf{g}(\mathbf{a},\mathbf{B})\mathbf{u} - \mathbf{g}(\mathbf{a},\mathbf{u})\mathbf{B} \tag{104}$$

and hence

$$\boldsymbol{\varepsilon}(\_,\breve{\_},\mathbf{u}\times\mathbf{B}) = \mathbf{u}\otimes\mathbf{g}(\breve{\_},\mathbf{B}) - \mathbf{B}\otimes\mathbf{g}(\breve{\_},\mathbf{u}) = \mathbf{u}\otimes\mathbf{B} - \mathbf{B}\otimes\mathbf{u} \tag{105}$$

(We need to put the metric-tensor term on the right of the tensor-product to ensure the slots on each side are filled in the right order). With this, the induction equation can be written as

$$\frac{\partial\mathbf{B}}{\partial t} - \nabla\cdot\left[\mathbf{u}\otimes\breve{\mathbf{B}} - \mathbf{B}\otimes\breve{\mathbf{u}}\right] = 0. \tag{106}$$

8.3.1 Evolution of Kinetic, Internal, Magnetic and Total Energy

Next we consider the evolution of the kinetic energy. To compute this take the dot-product of Eq. (101) with $\mathbf{u}$ to get

$$\underbrace{\mathbf{u}\cdot\frac{\partial}{\partial t}(\rho\mathbf{u})}_{\frac{\partial}{\partial t}(\frac{1}{2}\rho\mathbf{u}^2)+\frac{1}{2}\mathbf{u}^2\frac{\partial\rho}{\partial t}} + \mathbf{u}\cdot[\nabla\cdot\mathbf{T}(\_,\breve{\_})] = 0. \tag{107}$$

For any $\mathbf{u}$ we can prove the identity

$$\nabla\cdot\mathbf{T}(\mathbf{u},\_) = \mathbf{u}\cdot[\nabla\cdot\mathbf{T}(\_,\breve{\_})] + [\mathbf{T}(\_,\breve{\_})\cdot\nabla]\cdot\mathbf{u}. \tag{108}$$

Using the expression for $\mathbf{T}$ we can write

$$\mathbf{T}(\mathbf{u},\_) = \rho\mathbf{u}^2\mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0}\right)\mathbf{u} - \frac{1}{\mu_0}(\mathbf{B}\cdot\mathbf{u})\mathbf{B} \tag{109}$$

and

$$[\mathbf{T}(\_,\breve{\_})\cdot\nabla]\cdot\mathbf{u} = \underbrace{\rho\mathbf{u}\cdot[(\mathbf{u}\cdot\nabla)\mathbf{u}]}_{\rho\mathbf{u}\cdot\nabla\frac{1}{2}\mathbf{u}^2} + \left(p + \frac{\mathbf{B}^2}{2\mu_0}\right)\nabla\cdot\mathbf{u} - \frac{1}{\mu_0}\mathbf{B}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{u}]. \tag{110}$$

Now, we can show that

$$\mathbf{B}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{u}] = \nabla\cdot[\mathbf{B}(\mathbf{u}\cdot\mathbf{B})] - \mathbf{u}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{B}]\,. \tag{111}$$

Using these expressions and the continuity equation we get the evolution equation for the fluid kinetic-energy:

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho\mathbf{u}^2\right) + \nabla\cdot\left[\frac{1}{2}\rho\mathbf{u}^2\mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0}\right)\mathbf{u}\right] = \left(p + \frac{\mathbf{B}^2}{2\mu_0}\right)\nabla\cdot\mathbf{u} + \frac{1}{\mu_0}\mathbf{u}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{B}]\,. \tag{112}$$

The pressure evolution equation Eq. (99) can be rearranged as

$$\frac{\partial}{\partial t}\left(\frac{p}{\gamma-1}\right)+\nabla\cdot\left(\mathbf{u}\frac{p}{\gamma-1}\right)=-p\nabla\cdot\mathbf{u}. \tag{113}$$

Notice that $p/(\gamma-1)$ is the *internal energy* of the fluid. These equations essentially indicate that the rate of change of kinetic/internal energy in a volume of the fluid changes due to the flow of energy in/out of the fluid, but also due to mutual exchange: the compressibility of the plasma allows exchange between kinetic and internal energies. Note also that the kinetic energy in a volume also changes due to exchange with the magnetic field. There are two such terms: the term due to fluid compressibility and the second term that is present even if the fluid is incompressible.

The evolution equation of the magnetic field energy can be derived by taking the dot-product of the induction equation, Eq. (106), by $\mathbf{B}/\mu_0$. To simplify the second term we again use the identity Eq. (108) to write

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{B}^2}{2\mu_0}\right)\quad\underbrace{-\frac{1}{\mu_0}\nabla\cdot\left[(\mathbf{u}\cdot\mathbf{B})\mathbf{B}-\mathbf{B}^2\mathbf{u}\right]}_{-\nabla\cdot\left(\frac{1}{\mu_0}(\mathbf{u}\cdot\mathbf{B})\mathbf{B}-\frac{\mathbf{B}^2}{2\mu_0}\mathbf{u}\right)+\mathbf{u}\cdot\nabla\left(\frac{\mathbf{B}^2}{2\mu_0}\right)+\frac{\mathbf{B}^2}{2\mu_0}\nabla\cdot\mathbf{u}}\quad+\frac{1}{\mu_0}\mathbf{u}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{B}]\underbrace{-\frac{1}{\mu_0}\mathbf{B}\cdot[(\mathbf{u}\cdot\nabla)\mathbf{B}]}_{-\mathbf{u}\cdot\nabla\left(\frac{\mathbf{B}^2}{2\mu_0}\right)} \tag{114}$$

or

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{B}^2}{2\mu_0}\right)-\nabla\cdot\left(\frac{1}{\mu_0}(\mathbf{u}\cdot\mathbf{B})\mathbf{B}-\frac{\mathbf{B}^2}{2\mu_0}\mathbf{u}\right)=-\frac{\mathbf{B}^2}{2\mu_0}\nabla\cdot\mathbf{u}-\frac{1}{\mu_0}\mathbf{u}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{B}]. \tag{115}$$

Adding the equations for kinetic, internal and magnetic-field energies we finally get that the *total energy*

$$\mathcal{E}\equiv\frac{1}{2}\rho\mathbf{u}^2+\frac{p}{\gamma-1}+\frac{\mathbf{B}^2}{2\mu_0} \tag{116}$$

evolves as

$$\frac{\partial\mathcal{E}}{\partial t}+\nabla\cdot\left[(\mathcal{E}+p^*)\mathbf{u}-\frac{1}{\mu_0}(\mathbf{u}\cdot\mathbf{B})\mathbf{B}\right]=0 \tag{117}$$

where the total (fluid + magnetic-field) pressure is defined as

$$p^*\equiv p+\frac{\mathbf{B}^2}{2\mu_0}. \tag{118}$$

Exercise 16. Prove the identity Eq. (108).

Exercise 17. Prove that Eq. (111) is indeed true.

8.4 The Navier-Stokes Equations

The Navier-Stokes equations describe motion of viscous fluids in the continuum limit. These are the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{119}$$

the momentum equation

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \nabla \cdot \mathbf{T}(\_, \breve{\_}) = 0, \tag{120}$$

and the energy equation

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot [\mathbf{u}(\mathcal{E} + p) - \boldsymbol{\tau}(\mathbf{u}, \_) + \mathbf{q}] = 0. \tag{121}$$

The total energy $\mathcal{E}$ contains two contributions, one from the kinetic energy and the other from the internal energy:

$$\mathcal{E} = \frac{1}{2}\rho \mathbf{u} \cdot \mathbf{u} + \frac{p}{\gamma - 1} \tag{122}$$

where p is the scalar pressure. The tensor $\mathbf{T}$ is

$$\mathbf{T} = \rho \mathbf{u} \otimes \mathbf{u} + p\mathbf{g} - \boldsymbol{\tau} \tag{123}$$

where $\mathbf{g}$ is the metric tensor and $\boldsymbol{\tau}$ is the viscous stress-tensor defined as

$$\boldsymbol{\tau} = \mu \left(\nabla \otimes \mathbf{u} + \mathbf{u} \otimes \overleftarrow{\nabla} - \frac{2}{3}(\nabla \cdot \mathbf{u})\, \mathbf{g} \right) + \lambda (\nabla \cdot \mathbf{u})\, \mathbf{g} \tag{124}$$

and where μ and λ are the shear- and bulk-viscosity respectively. Finally, the heat-flux vector is given by

$$\mathbf{q} = -\kappa \nabla \varepsilon \tag{125}$$

where ε is the internal energy, determined from the equation of state, which, for an ideal-gas is $p = (\gamma - 1)\rho\varepsilon$.

9 Coordinate Representations, Choice of Basis

In this section we work out the various operators in generic coordinates. We will also discuss how the choice of representation of the vector/tensor quantities changes the form of the equations, some of which may be more suitable than others for different problems.

9.1 Divergence of Vector Field

Consider computing the divergence of a vector field, $\mathbf{u}$. We can write this as

$$\nabla \cdot \mathbf{u} = \mathbf{e}^i \cdot \frac{\partial \mathbf{u}}{\partial z^i}. \tag{126}$$

To make further progress we must choose the representation of the vector $\mathbf{u}$. Choosing $\mathbf{u} = u^j \mathbf{e}_j$ we can write

$$\nabla \cdot \mathbf{u} = \mathbf{e}^i \cdot \left(\frac{\partial u^j}{\partial z^i} \mathbf{e}_j + u^j \frac{\partial \mathbf{e}_j}{\partial z^i} \right) = \frac{\partial u^i}{\partial z^i} + \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_j}{\partial z^i} u^j. \tag{127}$$

The second term in the equation accounts for the change in the direction of the tangent vectors as one moves along a path. Although we can compute the derivatives of the tangent vectors directly, it is more conventional to represent them in basis expansion as follows

$$\frac{\partial \mathbf{e}_j}{\partial z^i} = \Gamma^m{}_{ji} \mathbf{e}_m \tag{128}$$

where $\Gamma^m{}_{ji}$ are the *Christoffel symbols*. Note that this equation merely states that the partial derivative of the tangent-vectors can be written as a linear combination of the tangent-vectors themselves. Using this expression we can write the divergence as

$$\nabla \cdot \mathbf{u} = \frac{\partial u^i}{\partial z^i} + \Gamma^i{}_{ji} u^j. \tag{129}$$

The Christoffel symbols can be computed by taking the dot-product of Eq. (128) by $\mathbf{e}^m$ to get

$$\Gamma^m{}_{ij} = \mathbf{e}^m \cdot \frac{\partial \mathbf{e}_i}{\partial z^j}. \tag{130}$$

Also notice that the Christoffel symbols are symmetric in the lower indices as

$$\frac{\partial \mathbf{e}_i}{\partial z^j} = \frac{\partial}{\partial z^j} \left(\frac{\partial \mathbf{x}}{\partial z^i} \right) = \frac{\partial}{\partial z^i} \left(\frac{\partial \mathbf{x}}{\partial z^j} \right) = \frac{\partial \mathbf{e}_j}{\partial z^i} \tag{131}$$

showing that

$$\Gamma^m{}_{ij} = \Gamma^m{}_{ji}. \tag{132}$$

The determinant of the metric, $g = \det(g_{ij})$ plays an important role in what follows. Define the Jacobian as $\mathcal{J} = \sqrt{g}$. We can derive the following very important expression for the derivative of the Jacobian

$$\nabla \cdot \mathbf{e}_j = \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_j}{\partial z^i} = \frac{1}{\mathcal{J}} \frac{\partial \mathcal{J}}{\partial z^j} \tag{133}$$

For now, we will use this important result without proof.

Now, in terms of the Christoffel symbols we can write the identity Eq. (133) as

$$\frac{1}{\mathcal{J}}\frac{\partial \mathcal{J}}{\partial z^j} = \Gamma^i{}_{ij} \tag{134}$$

Using this, we can write the divergence, Eq. (129), as

$$\nabla \cdot \mathbf{u} = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}u^i\right). \tag{135}$$

This is typically the formula for divergence that we will use in writing conservation laws.

Exercise 18. Use $\mathbf{e}_j \cdot \mathbf{e}^k = \delta_j{}^k$ to show that

$$\mathbf{e}_j \cdot \frac{\partial \mathbf{e}^k}{\partial z^i} = -\frac{\partial \mathbf{e}_j}{\partial z^i} \cdot \mathbf{e}^k = -\Gamma^k{}_{ji}. \tag{136}$$

Exercise 19. Take the derivative of $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$, cyclically permute indices and use the symmetry property Eq. (131) to express the Christoffel symbols in terms of the metric tensor as

$$\Gamma^n{}_{kj} = \frac{1}{2}g^{nm}\left(\frac{\partial g_{mk}}{\partial z^j} + \frac{\partial g_{jm}}{\partial z^k} - \frac{\partial g_{kj}}{\partial z^m}\right). \tag{137}$$

Exercise 20. If we choose a different representation of $\mathbf{u}$, we will get a different expression for the divergence (though, of course, the *value* of the divergence will be independent of the representation). Take $\mathbf{u} = u_j\mathbf{e}^j$ to show

$$\nabla \cdot \mathbf{u} = g^{ij}\left(\frac{\partial u_j}{\partial z_i} - u_k\Gamma^k{}_{ji}\right). \tag{138}$$

Exercise 21. Consider the polar coordinate system defined above. Show that when using the representation $\mathbf{u} = \hat{u}^r\hat{\mathbf{e}}_r + \hat{u}^\theta\hat{\mathbf{e}}_\theta$ the divergence is

$$\nabla \cdot \mathbf{u} = \frac{1}{r}\frac{\partial}{\partial r}\left(r\hat{u}^r\right) + \frac{1}{r}\frac{\partial \hat{u}^\theta}{\partial \theta}. \tag{139}$$

9.2 Laplacian of a Scalar Field

An immediate and important application of Eq. (135) is to compute the Laplacian of a scalar field f. The gradient can be written in the tangent vectors as

$$\nabla f = g^{ij}\frac{\partial f}{\partial z^i}\mathbf{e}_j. \tag{140}$$

Using this in Eq. (135) to take the divergence, we get the *Laplacian*

$$\nabla^2 f = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}g^{ij}\frac{\partial f}{\partial z^i}\right). \tag{141}$$

Exercise 22. Show that in polar coordinates, the Laplacian is

$$\nabla^2 f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}. \tag{142}$$

9.3 Divergence of Tensor Field

We can also take the divergence of a second or higher order tensor. For higher-order tensors we need to indicate the slot with which the divergence is to be taken. To compute the divergence of a second-order tensor with the first slot we can use

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = \mathbf{e}^i \cdot \frac{\partial}{\partial z^i} \left(T^{mn} \breve{\mathbf{e}}_m \otimes \mathbf{e}_n \right) = \frac{\partial T^{in}}{\partial z^i} \mathbf{e}_n + T^{mn} \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_m}{\partial z^i} + T^{in} \frac{\partial \mathbf{e}_n}{\partial z^i}. \tag{143}$$

Using the definition of the Christoffel symbols we get

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = \left(\frac{\partial T^{in}}{\partial z^i} + T^{mn} \Gamma^i{}_{mi} + T^{im} \Gamma^n{}_{im} \right) \mathbf{e}_n. \tag{144}$$

Using the identity Eq. (135) we can write this instead in a more compact form as

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i} \left(\mathcal{J} T^{in} \mathbf{e}_n \right) = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i} \left(\mathcal{J} \mathbf{T}(\mathbf{e}^i, \_) \right) \tag{145}$$

The last form is suggestive as it involves the geometric tensor $\mathbf{T}$ and not its specific representation, hence allowing use of any representation one chooses.

Exercise 23. Compute the divergence $\nabla \cdot \mathbf{T}(\breve{\_}, \_)$ of a second-order tensor represented as $\mathbf{T} = T_{mn} \mathbf{e}^m \otimes \mathbf{e}^n$.

An important special case is when the tensor $\mathbf{T}$ is *anti-symmetric*. In this case, we can swap i and m in the last term in Eq. (144) and then use the symmetry of the Christoffel symbol and anti-symmetry of $\mathbf{T}$ to find that

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = \left(\frac{\partial T^{in}}{\partial z^i} + T^{mn} \Gamma^i{}_{mi} \right) \mathbf{e}_n. \qquad \text{(Antisymmetric } \mathbf{T}) \tag{146}$$

Again, using the identity Eq. (135) we can write this instead

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i} \left(\mathcal{J} T^{in} \right) \mathbf{e}_n. \qquad \text{(Antisymmetric } \mathbf{T}) \tag{147}$$

Of course, we can also write

$$\nabla \cdot \mathbf{T}(\_, \breve{\_}) = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i} \left(\mathcal{J} T^{ni} \right) \mathbf{e}_n. \qquad \text{(Antisymmetric } \mathbf{T}) \tag{148}$$

9.4 Tensor Products with Derivative Operator

Another operation that is often needed is the ability to compute tensor products like $\nabla \otimes \mathbf{u}$, where $\mathbf{u}$ is a vector. Representing $\mathbf{u} = u^m \mathbf{e}_m$ we can compute this as

$$\nabla \otimes \mathbf{u} = \mathbf{e}^i \otimes \frac{\partial}{\partial z^i} (u^m \mathbf{e}_m) = \underbrace{\left(\frac{\partial u^m}{\partial z^i} + u^k \Gamma^m{}_{ki} \right)}_{\equiv u^m{}_{;i}} \mathbf{e}^i \otimes \mathbf{e}_m. \tag{149}$$

The expression in the bracket is known as the *covariant derivative* and is denoted by $u^m{}_{;i}$. Hence, we can write this expression as

$$\nabla \otimes \mathbf{u} = u^m{}_{;i} \mathbf{e}^i \otimes \mathbf{e}_m. \tag{150}$$

Representing instead $\mathbf{u} = u_m \mathbf{e}^m$ we can compute $\nabla \otimes \mathbf{u}$ as

$$\nabla \otimes \mathbf{u} = \mathbf{e}^i \otimes \frac{\partial}{\partial z^i}(u_m \mathbf{e}^m) = \left(\frac{\partial u_m}{\partial z^i} - u_k \Gamma^k{}_{mi} \right) \mathbf{e}^i \otimes \mathbf{e}^m, \tag{151}$$

or using the notation of covariant derivatives as $\nabla \otimes \mathbf{u} = u_{m;i} \mathbf{e}^i \otimes \mathbf{e}^m$.

Exercise 24. Show that $\nabla \otimes \mathbf{x} = \mathbf{g}$, where $\mathbf{x}$ is the position vector.

Exercise 25. The *divergence* of a vector $\mathbf{u}$ can be computed as $\nabla \cdot \mathbf{u} = \mathrm{Tr}(\nabla \otimes \mathbf{u})$. Show the expressions obtained from taking the trace with either representation of the vector yields the same expressions we derived before.

We have defined the covariant derivative in component form as

$$u^m{}_{;i} = \frac{\partial u^m}{\partial z^i} + u^k \Gamma^m{}_{ki} \tag{152}$$

$$u_{m;i} = \frac{\partial u_m}{\partial z^i} - u_k \Gamma^k{}_{mi}. \tag{153}$$

Using these expressions as prototypes, we can compute the covariant dervatives of a tensor in a coordinate system. For example, we have

$$T^{mn}{}_{;i} = \frac{\partial T^{mn}}{\partial z^i} + T^{kn} \Gamma^m{}_{ki} + T^{mk} \Gamma^n{}_{ki} \tag{154}$$

$$T_{mn;i} = \frac{\partial T_{mn}}{\partial z^i} - T_{kn} \Gamma^k{}_{mi} - T_{mk} \Gamma^k{}_{ni} \tag{155}$$

$$T^m{}_{n;i} = \frac{\partial T^m{}_n}{\partial z^i} + T^k{}_n \Gamma^m{}_{ki} - T^m{}_k \Gamma^k{}_{ni} \tag{156}$$

(Note that the components $T^m{}_k$ and $T_m{}^k$ are, in general, different). With these, we can now compute the tensor products of the vector dervative with a second-order tensor $\mathbf{T}$ as

$$\nabla \otimes \mathbf{T} = T^{mn}{}_{;i} \mathbf{e}^i \otimes \mathbf{e}_m \otimes \mathbf{e}_n = T_{mn;i} \mathbf{e}^i \otimes \mathbf{e}^m \otimes \mathbf{e}^n = T^m{}_{n;i} \mathbf{e}^i \otimes \mathbf{e}_m \otimes \mathbf{e}^n. \tag{157}$$

Note the order of the basis in the tensor product: the dual basis $\mathbf{e}^i$ appears first as the derivative operator appears on the left. The other basis (or duals) appear in the same order as they do in the tensor components.

As for the vector case, we can compute the divergence of the second-order tensor by taking the trace of the above expression with the appropriate slots. For example, we can compute

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = T^{mn}{}_{;i} \, \mathrm{Tr}\big(\breve{\mathbf{e}}^i \otimes \breve{\mathbf{e}}_m \otimes \mathbf{e}_n\big) = T^{in}{}_{;i} \mathbf{e}_n. \tag{158}$$

This is of course just the expression we obtained before (see Eq. (144)). We can also write this expresion in terms of the the covariant components as

$$\nabla \cdot \mathbf{T}(\breve{\_}, \_) = T_{mn;i} \operatorname{Tr}\left(\breve{\mathbf{e}}^i \otimes \breve{\mathbf{e}}^m \otimes \mathbf{e}^n\right) = g^{im} T_{mn;i} \mathbf{e}^n. \tag{159}$$

For some applications, we will also need the tensor product $\mathbf{u} \otimes \overleftarrow{\nabla}$. The left-pointing arrow on the derivative indicates that it acts on the object on the left. Again, letting $\mathbf{u} = u^m \mathbf{e}_m$ we can compute this as

$$\mathbf{u} \otimes \overleftarrow{\nabla} = \frac{\partial}{\partial z^i}(u^m \mathbf{e}_m) \otimes \mathbf{e}^i = u^m{}_{;i} \mathbf{e}_m \otimes \mathbf{e}^i. \tag{160}$$

9.5 Curl in Three Dimensions

Given a vector field $\mathbf{B} = B^m \mathbf{e}_m$ we can compute its curl as

$$\nabla \times \mathbf{B} = \mathbf{e}^i \times \frac{\partial}{\partial z^i}(B^m \mathbf{e}_m) = \frac{\partial B^m}{\partial z^i} \mathbf{e}^i \times \mathbf{e}_m + B^m \mathbf{e}^i \times \frac{\partial \mathbf{e}_m}{\partial z^i} = B^m{}_{;i} \mathbf{e}^i \times \mathbf{e}_m \tag{161}$$

Previously, we showed that we can write the curl instead as a *divergence* as follows.

$$\nabla \times \mathbf{B} = \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B})]. \tag{162}$$

Using the expression Eq. (145) we can write this as

$$\nabla \times \mathbf{B} = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i}[\mathcal{J} \boldsymbol{\varepsilon}(\_, \mathbf{e}^i, \mathbf{B})] = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i}[\mathcal{J}(\mathbf{e}^i \times \mathbf{B})]. \tag{163}$$

Exercise 26. Show that Eq. (161) and Eq. (163) result in the same expression for the the curl.

Exercise 27. We can write $\nabla \times \mathbf{B} = \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B})]$. Also, the $\boldsymbol{\varepsilon}$ tensor is anti-symmetric. Use the special form for the divergence of an anti-symmetric tensor, Eq. (148), to derive yet another coordinate form for the curl.

Exercise 28. Let $\mathbf{u} = \boldsymbol{\omega} \times \mathbf{x}$, where $\boldsymbol{\omega}$ is a constant vector. Compute $\nabla \otimes \mathbf{u}$.

As an application, consider the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0. \tag{164}$$

We rewrite this as a divergence as

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{E})] = 0. \tag{165}$$

In coordinates, using Eq. (148), we can write this as

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left[\mathcal{J}\mathbf{e}^n \cdot (\mathbf{e}^i \times \mathbf{E})\right]\mathbf{e}_n = 0. \tag{166}$$

This can be written as evolution equations for $B = \mathbf{e}_n B^n$ by taking the dot-product of the above expression with $\mathbf{e}^n$ to get

$$\frac{\partial B^n}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left[\mathcal{J}\mathbf{e}^n \cdot (\mathbf{e}^i \times \mathbf{E})\right] = 0. \tag{167}$$

We now need to choose a representation for $\mathbf{E}$. Using $\mathbf{E} = \mathbf{e}^m E_m$ and the relations between basis vectors and their duals in three dimensions (see Eq. (28)), we can compute the independent non-zero components

$$\mathbf{e}^1 \cdot (\mathbf{e}^2 \times \mathbf{E}) = \frac{E_3}{\mathcal{J}} \tag{168a}$$

$$\mathbf{e}^3 \cdot (\mathbf{e}^1 \times \mathbf{E}) = \frac{E_2}{\mathcal{J}} \tag{168b}$$

$$\mathbf{e}^2 \cdot (\mathbf{e}^3 \times \mathbf{E}) = \frac{E_1}{\mathcal{J}}. \tag{168c}$$

Using these we get the evolution equations

$$\frac{\partial B^1}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial E_3}{\partial z^2} - \frac{1}{\mathcal{J}}\frac{\partial E_2}{\partial z^3} = 0 \tag{169a}$$

$$\frac{\partial B^2}{\partial t} - \frac{1}{\mathcal{J}}\frac{\partial E_3}{\partial z^1} + \frac{1}{\mathcal{J}}\frac{\partial E_1}{\partial z^3} = 0 \tag{169b}$$

$$\frac{\partial B^3}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial E_2}{\partial z^1} - \frac{1}{\mathcal{J}}\frac{\partial E_1}{\partial z^2} = 0. \tag{169c}$$

Exercise 29. Using instead $\mathbf{E} = \mathbf{e}_m E^m$ derive the evolution equations for the components B^k that result from the induction equation.

10 Solving Vector Equations

In this section I show how to solve commonly encountered vector equations. In general, one should avoid the temptation to decompose the equations into components and just use various vector/tensor identities.

10.1 Find $\mathbf{A}$ if $\mathbf{A} = \mathbf{R} + \mathbf{A} \times \mathbf{B}$

Often we want to find $\mathbf{A}$ such that it satisfies the vector equation

$$\mathbf{A} = \mathbf{R} + \mathbf{A} \times \mathbf{B}. \tag{170}$$

Take the cross-product with $\mathbf{B}$:

$$\mathbf{A} \times \mathbf{B} = \mathbf{R} \times \mathbf{B} + (\mathbf{A} \times \mathbf{B}) \times \mathbf{B} \tag{171}$$

Now rewrite the last term as[3]

$$(\mathbf{A} \times \mathbf{B}) \times \mathbf{B} = (\mathbf{B} \cdot \mathbf{A})\mathbf{B} - B^2\mathbf{A} \tag{172}$$

to get

$$\mathbf{A} \times \mathbf{B} = \mathbf{R} \times \mathbf{B} + (\mathbf{B} \cdot \mathbf{A})\mathbf{B} - B^2\mathbf{A}. \tag{173}$$

From the original equation we have $\mathbf{A} \times \mathbf{B} = \mathbf{A} - \mathbf{R}$ and $\mathbf{B} \cdot \mathbf{A} = \mathbf{B} \cdot \mathbf{R}$. Plugging these into the last expression we finally get

$$\mathbf{A} = \frac{\mathbf{R} + \mathbf{R} \times \mathbf{B} + (\mathbf{B} \cdot \mathbf{R})\mathbf{B}}{1 + B^2}. \tag{174}$$

We can write this in an alternate way as follows. Let $\mathbf{b} = \mathbf{B}/B$ be the unit vector along $\mathbf{B}$. Then, every vector can be decomposed into parallel and perpendicular components, for example $\mathbf{A} = A_\parallel \mathbf{b} + \mathbf{A}_\perp$, where $A_\parallel = \mathbf{A} \cdot \mathbf{b}$. From Eq. (170), we see that $A_\parallel = R_\parallel$. Decomposing the right-hand of Eq. (174) into parallel and perpendicular components, we get

$$\mathbf{A}_\perp = \frac{\mathbf{R}_\perp + \mathbf{R}_\perp \times \mathbf{B}}{1 + B^2}, \tag{175}$$

which is the solution to the equation $\mathbf{A}_\perp = \mathbf{R}_\perp + \mathbf{A}_\perp \times \mathbf{B}$. This form is often more useful than Eq. (174).

Example 4. As an example of the application of this formula, consider we wish to discretize the ODE describing the motion of a charged particle in a given EM field.

$$\frac{d\mathbf{v}}{dt} = \frac{q}{m} \left(\mathbf{E} + \mathbf{v} \times \mathbf{B} \right). \tag{176}$$

We will use a time-centered implicit scheme

$$\frac{\mathbf{v}^{n+1} - \mathbf{v}^n}{\Delta t} = \frac{q}{m} \left(\mathbf{E} + \frac{\mathbf{v}^{n+1} + \mathbf{v}^n}{2} \times \mathbf{B} \right). \tag{177}$$

Defining $\mathbf{v}^{n+1/2} \equiv (\mathbf{v}^{n+1} + \mathbf{v}^n)/2$ we can write this as

$$\mathbf{v}^{n+1/2} = \mathbf{v}^n + \frac{q\Delta t}{2m}\mathbf{E} + \mathbf{v}^{n+1/2} \times \frac{q\Delta t \mathbf{B}}{2m}. \tag{178}$$

[3] Follows from $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{C} \times \mathbf{B}) \times \mathbf{A} = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$.

This is in the same form as Eq. (170), and using Eq. (174) we can write the explicit update formula

$$\mathbf{v}^{n+1/2} = \lambda\left(\mathbf{v}^n + \frac{q\Delta t}{2m}\mathbf{v}^n\times\mathbf{B} + \frac{q^2\Delta t^2}{4m^2}(\mathbf{v}^n\cdot\mathbf{B})\mathbf{B}\right) + \frac{\lambda q\Delta t}{2m}\left(\mathbf{E} + \frac{q\Delta t}{2m}\mathbf{E}\times\mathbf{B} + \frac{q^2\Delta t^2}{4m^2}(\mathbf{E}\cdot\mathbf{B})\mathbf{B}\right) \tag{179}$$

where $\lambda = 1/(1+q^2\Delta t^2B^2/4m^2)$. Once $\mathbf{v}^{n+1/2}$ is determined, we can calculate $\mathbf{v}^{n+1} = 2\mathbf{v}^{n+1/2} - \mathbf{v}^n$.

Exercise 30. Solve the equation $\mathbf{q} + \omega\mathbf{b}\times\mathbf{q} = -k\nabla T$ for $\mathbf{q}$. Here, $\mathbf{q}$ is the heat-flux vector, T is the temperature, $\mathbf{b}$ is the unit-vector along the magnetic field and ω is the product of the collision time and cyclotron frequency. Write the answer in the form

$$\mathbf{q} = -\kappa_\perp\nabla T - (\kappa_\| - \kappa_\perp)\mathbf{b}\otimes\mathbf{b} : \nabla T + \kappa_\wedge\mathbf{b}\times\nabla T, \tag{180}$$

expressing $\kappa_\|, \kappa_\perp, \kappa_\wedge$ in terms of ω and k.

10.2 Find A if $\mathbf{A} = \mathbf{R} + \mathbf{A}\times\mathbf{B} + \xi\mathbf{B}(\mathbf{B}\cdot\mathbf{A})$

This is a small modification to the previous problem. Left as an exercise to the reader, one can write the solution as

$$\mathbf{A} = \frac{\mathbf{R} + \mathbf{R}\times\mathbf{B} + (\mathbf{B}\cdot\mathbf{R})\mathbf{B}}{1+B^2} + \frac{\xi}{1-B^2\xi}\mathbf{B}(\mathbf{B}\cdot\mathbf{R}). \tag{181}$$

This shows that there is no solution to this vector equation if $\xi B^2 = 1$.

A Note on the Appendices

Most of the stuff in the appendix are merely place-holders for material I am yet to work into the main text.

B 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 reciprocals, consider the vector field

$$\mathbf{B} = C(\mathbf{x})\mathbf{e}^1\times\mathbf{e}^2 \tag{182}$$

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:

$$\nabla\cdot\mathbf{B} = \nabla\cdot\big(C(\mathbf{x})\mathbf{e}^1\times\mathbf{e}^2\big) = C\big(\mathbf{e}^2\cdot\underbrace{\nabla\times\mathbf{e}^1}_{\nabla\times\nabla z^1=0} - \mathbf{e}^1\cdot\underbrace{\nabla\times\mathbf{e}^2}_{\nabla\times\nabla z^2=0}\big) + \mathbf{e}^1\times\mathbf{e}^2\cdot\nabla C. \tag{183}$$

We are now left with

$$\nabla \cdot \mathbf{B} = \frac{1}{C} \mathbf{B} \cdot \nabla C \tag{184}$$

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*[4]. 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. (182). 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. (182) we have ensured that $\mathbf{B} \cdot \mathbf{e}^1 = \mathbf{B} \cdot \mathbf{e}^2 = 0$. Hence, we can also write

$$\mathbf{B} = \big(\mathbf{B} \cdot \mathbf{e}^3\big) \mathbf{e}_3. \tag{185}$$

Using Eq. (182) in this we get[5]

$$\mathbf{B} = C \underbrace{(\mathbf{e}^1 \times \mathbf{e}^2) \cdot \mathbf{e}^3}_{J^{-1}} \mathbf{e}_3. \tag{186}$$

Hence, we must have

$$B^2 = \mathbf{B} \cdot \mathbf{B} = \frac{C^2}{J^2} \underbrace{\mathbf{e}_3 \cdot \mathbf{e}_3}_{g_{33}} \tag{187}$$

and hence

$$C = \frac{BJ}{\sqrt{g_{33}}}. \tag{188}$$

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

$$\mathbf{B} \cdot \nabla (BJ) = 0 \tag{189}$$

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.

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

[5]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.

C Symplectic Geometry and Transforming Hamiltonian Systems

Hamiltonian systems can be written in the generic form

$$\frac{\partial f}{\partial t} + \{f, H\} = 0 \tag{190}$$

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

$$\{f, g\} = \nabla_{\mathbf{x}} f \cdot \nabla_{\mathbf{v}} g - \nabla_{\mathbf{v}} f \cdot \nabla_{\mathbf{x}} g \tag{191}$$

Transforming these equations to general coordinate system is not trivial. See Chapter II of Cary and Brizard[2] for details on how to do this[6]. Here I only state the final result. The mapping to a general coordinate system z^i, $i = 1, \ldots, 2n$ is written as

$$x^i = x^i(z^1, \ldots, z^{2n}) \tag{192}$$

$$v^i = v^i(z^1, \ldots, z^{2n}) \tag{193}$$

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

$$\{f, g\} = \frac{\partial f}{\partial z^i} \Pi^{ij} \frac{\partial g}{\partial z^j} \tag{194}$$

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 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

$$\mathbf{\Pi} = \mathbf{M} \boldsymbol{\sigma} \mathbf{M}^T \tag{195}$$

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 reciprocal 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):

$$\boldsymbol{\sigma} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \tag{196}$$

[6] The notation in Cary and Brizard[2] 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 reciprocal vectors (rows of $\mathbf{M}$) play an important role, rather than Riemannian geometry, in which the metric plays an important role.

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. (195) is

$$\Pi^{ij} = \nabla_{\mathbf{x}} z^i \cdot \nabla_{\mathbf{v}} z^j - \nabla_{\mathbf{v}} z^i \cdot \nabla_{\mathbf{x}} z^j. \tag{197}$$

This follows from the fact that the reciprocal 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 reciprocal vectors

$$\mathbf{e}^i = \nabla_{\mathbf{x}} z^i \tag{198}$$

$$\tilde{\mathbf{e}}^i = \nabla_{\mathbf{v}} z^i. \tag{199}$$

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

$$\Pi^{ij} = \mathbf{e}^i \cdot \tilde{\mathbf{e}}^j - \mathbf{e}^j \cdot \tilde{\mathbf{e}}^i. \tag{200}$$

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

$$\partial \equiv \mathbf{e}^i \frac{\partial}{\partial z^i} \tag{201}$$

$$\tilde{\partial} \equiv \tilde{\mathbf{e}}^i \frac{\partial}{\partial z^i} \tag{202}$$

then the Poisson bracket can be represented in the compact form

$$\{f, g\} = \partial f \cdot \tilde{\partial} g - \partial g \cdot \tilde{\partial} f. \tag{203}$$

Notice that this has the same structure as the canonical Poisson bracket defined in Eq. (191), 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. (190) can be written in the elegant form

$$\frac{\partial f}{\partial t} + \tilde{\partial} H \cdot \partial f - \partial H \cdot \tilde{\partial} f = 0. \tag{204}$$

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. (190) in conservative form as follows:

$$\frac{\partial f}{\partial t} + \frac{1}{\mathcal{J}} \frac{\partial}{\partial z_j} \left(\mathcal{J} \alpha^j f \right) = 0 \tag{205}$$

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

$$\alpha^i \equiv \dot{z}^i = \{z^i, H\} = \Pi^{ij} \frac{\partial H}{\partial z^j}. \tag{206}$$

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

$$\alpha^i = \mathbf{e}^i \cdot \tilde{\partial} H - \tilde{\mathbf{e}}^i \cdot \partial H. \tag{207}$$

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

$$\frac{1}{\mathcal{J}} \frac{\partial}{\partial z_j} \left(\mathcal{J} \alpha^j \right) = 0. \tag{208}$$

Using Eq. (194) we can also write Eq. (190) in the non-conservative form:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial z_j} \Pi^{jk} \frac{\partial H}{\partial z_k} = \frac{\partial f}{\partial t} + \alpha^j \frac{\partial f}{\partial z_j} = 0. \tag{209}$$

For example, in Binney and Termaine[1] 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.

C.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

$$\{f, g\} = \frac{\partial f}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial g}{\partial x} \tag{210}$$

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:

$$x = r \cos\theta, \qquad y = r \sin\theta. \tag{211}$$

The Jacobian matrix of this transform and its inverse are:

$$\mathbf{D} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix} \tag{212}$$

The metric tensor can be read off from $ds^2 = dr^2 + r^2 d\theta^2$ and, $\mathcal{J} = r$. From Eq. (195) the Poisson tensor in these coordinates is

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 1/r \\ -1/r & 0 \end{pmatrix}. \tag{213}$$

From this, the characteristic velocity can be computed as

$$\dot{r} = \{r, H\} = \frac{1}{r} \frac{\partial \phi}{\partial \theta} \tag{214}$$

$$\dot{\theta} = \{\theta, H\} = -\frac{1}{r} \frac{\partial \phi}{\partial r}. \tag{215}$$

C.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. (191), and the Hamiltonian:

$$H = \frac{1}{2}\mathbf{v}^2 + \phi(\mathbf{x}) \tag{216}$$

(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 $\mathbf{e}_c^i$ be their reciprocals. Then there are two "natural" choices of velocity coordinate transforms. The first is to choose w^i ($i = 1, 2, 3$) such that

$$\mathbf{v} = \mathbf{e}_{c,i} w^i. \tag{217}$$

With this choice, the kinetic energy is

$$\frac{1}{2}\mathbf{v}^2 = \frac{1}{2}\mathbf{e}_{c,i} w^i \cdot \mathbf{e}_{c,j} w^j = \frac{1}{2} \underbrace{\mathbf{e}_{c,i} \cdot \mathbf{e}_{c,j}}_{h_{ij}} w^i w^j = \frac{1}{2} h_{ij} w^i w^j \tag{218}$$

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

$$\mathbf{v} = \mathbf{e}_c^i w_i. \tag{219}$$

With this choice, the kinetic energy is

$$\frac{1}{2}\mathbf{v}^2 = \frac{1}{2}\mathbf{e}_c^i w_i \cdot \mathbf{e}_c^j w_j = \frac{1}{2} \underbrace{\mathbf{e}_c^i \cdot \mathbf{e}_c^j}_{h^{ij}} w_i w_j = \frac{1}{2} h^{ij} w_i w_j \tag{220}$$

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 reciprocals are

$$\mathbf{e}_r = \cos\theta\, \mathbf{e}_x + \sin\theta\, \mathbf{e}_y \tag{221}$$

$$\mathbf{e}_\theta = -r\sin\theta\, \mathbf{e}_x + r\cos\theta\, \mathbf{e}_y \tag{222}$$

and

$$\mathbf{e}^r = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{223}$$

$$\mathbf{e}^\theta = -\frac{1}{r}\sin\theta\,\mathbf{e}_x + \frac{1}{r}\cos\theta\,\mathbf{e}_y \tag{224}$$

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. (**??**), 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. (211) but keep the velocity space coordinates unchanged. The Jacobian matrix of the transform and its inverse are

$$\mathbf{D} = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 & 0 \\ \sin\theta & r\cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{M} = \begin{pmatrix} \cos\theta & \sin\theta & 0 & 0 \\ -\sin\theta/r & \cos\theta/r & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{225}$$

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

$$\mathbf{e}^1 = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{226}$$

$$\mathbf{e}^2 = -\frac{1}{r}\sin\theta\,\mathbf{e}_x + \frac{1}{r}\cos\theta\,\mathbf{e}_y \tag{227}$$

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

$$\partial = \mathbf{e}_x\left(\cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}\right) + \mathbf{e}_y\left(\sin\theta\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\frac{\partial}{\partial\theta}\right) \tag{228}$$

and

$$\tilde{\partial} = \mathbf{e}_x\frac{\partial}{\partial v_x} + \mathbf{e}_y\frac{\partial}{\partial v_y}. \tag{229}$$

Having obtained the gradient operators, we can write the Vlasov equation directly from Eq. (204). We can also, instead, compute first compute the Poisson tensor:

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta/r & \cos\theta/r \\ -\cos\theta & \sin\theta/r & 0 & 0 \\ -\sin\theta & -\cos\theta/r & 0 & 0 \end{pmatrix}. \tag{230}$$

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

$$\dot{r} = v_x\cos\theta + v_y\sin\theta \tag{231}$$

$$\dot{\theta} = -\frac{v_x}{r}\sin\theta + \frac{v_y}{r}\cos\theta. \tag{232}$$

and in velocity space as

$$\dot{v}_x = -\frac{\partial \phi}{\partial r}\cos\theta + \frac{1}{r}\frac{\partial \phi}{\partial \theta}\sin\theta \tag{233}$$

$$\dot{v}_y = -\frac{\partial \phi}{\partial r}\sin\theta - \frac{1}{r}\frac{\partial \phi}{\partial \theta}\cos\theta. \tag{234}$$

It is left as a (simple) exercise for the reader to show that using Eqns. 228 and 229 in Eq. (204) gives the same expressions as above.

Transforming phase-space, choice NC. In this example, we will use the velocity transform given by Eq. (217), which, for polar coordinates becomes:

$$v_x = w_r \cos\theta - w_\theta r \sin\theta \tag{235}$$

$$v_y = w_r \sin\theta + w_\theta r \cos\theta. \tag{236}$$

With this choice, the Hamiltonian becomes

$$H = \frac{1}{2}(w_r^2 + r^2 w_\theta^2) + \phi(r,\theta). \tag{237}$$

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

$$\mathbf{e}^1 = \cos\theta\,\mathbf{e}_x + \sin\theta\,\mathbf{e}_y \tag{238}$$

$$\mathbf{e}^2 = -\frac{1}{r}\sin\theta\,\mathbf{e}_x + \frac{1}{r}\cos\theta\,\mathbf{e}_y \tag{239}$$

$$\mathbf{e}^3 = -\sin\theta w_\theta\,\mathbf{e}_x + \cos\theta w_\theta\,\mathbf{e}_y \tag{240}$$

$$\mathbf{e}^4 = \left(\sin\theta\frac{w_r}{r^2} - \cos\theta\frac{w_\theta}{r}\right)\,\mathbf{e}_x - \left(\sin\theta\frac{w_\theta}{r} + \cos\theta\frac{w_r}{r^2}\right)\,\mathbf{e}_y. \tag{241}$$

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. (200) we can compute

$$\mathbf{\Pi} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1/r^2 \\ -1 & 0 & 0 & 2w_\theta/r \\ 0 & -1/r^2 & -2w_\theta/r & 0 \end{pmatrix}. \tag{242}$$

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

$$\dot{w}_r = r w_\theta^2 - \frac{\partial \phi}{\partial r} \tag{243}$$

$$\dot{w}_\theta = -\frac{2 w_r w_\theta}{r} - \frac{1}{r^2}\frac{\partial \phi}{\partial \theta}. \tag{244}$$

With these definitions, the Vlasov equation in these coordinates becomes

$$\frac{\partial f}{\partial t} + w_r \frac{\partial f}{\partial r} + w_\theta \frac{\partial f}{\partial \theta} + \left(r w_\theta^2 - \frac{\partial \phi}{\partial r} \right) \frac{\partial f}{\partial w_r} - \left(\frac{2 w_r w_\theta}{r} + \frac{1}{r^2} \frac{\partial \phi}{\partial \theta} \right) \frac{\partial f}{\partial w_\theta} = 0. \tag{245}$$

Transforming phase-space, choice C. Now consider the velocity transform given by Eq. (219), which, for polar coordinates becomes:

$$v_x = \cos\theta w_r - \frac{w_\theta}{r} \sin\theta \tag{246}$$

$$v_y = \sin\theta w_r + \frac{w_\theta}{r} \cos\theta. \tag{247}$$

With this choice, the Hamiltonian becomes

$$H = \frac{1}{2r^2} (r^2 w_r^2 + w_\theta^2) + \phi(r, \theta). \tag{248}$$

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

$$\dot{w}_r = \frac{w_\theta^2}{r^3} - \frac{\partial \phi}{\partial r} \tag{249}$$

$$\dot{w}_\theta = -\frac{\partial \phi}{\partial \theta}. \tag{250}$$

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

$$\frac{\partial f}{\partial t} + w_r \frac{\partial f}{\partial r} + \frac{w_\theta}{r^2} \frac{\partial f}{\partial \theta} + \left(\frac{w_\theta^2}{r^3} - \frac{\partial \phi}{\partial r} \right) \frac{\partial f}{\partial w_r} - \frac{\partial \phi}{\partial \theta} \frac{\partial f}{\partial w_\theta} = 0. \tag{251}$$

As verification, this is identical to Eq. (4.12) in[1]. Note that Binney and Tremaine do not give the general transformation formalism presented here, but only deal with canonical coordinates.

References

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

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