A Tutorial on Tensor Calculus

Ammar H. Hakim

Contents

1 Tangent Vector Space on Smooth Manifolds **1**

2 Euclidean Space **4**

3 Tensor Products and Tensors **5**

4 The Fundamental Derivative Operator **8**

5 Cross Product and Curl in Three Dimensions **10**

6 Select Equations of Mathematical Physics **12**
- 6.1 The Advection-Diffusion Equation 12
- 6.2 Maxwell Equations 13
- 6.3 The Ideal MHD Equations 14
  - 6.3.1 Evolution of Kinetic, Internal, Magnetic and Total Energy 15
- 6.4 The Navier-Stokes Equations 17

7 Coordinate Representations, Choice of Basis **17**
- 7.1 Divergence of Vector Field 18
- 7.2 Laplacian of a Scalar Field 19
- 7.3 Divergence of Tensor Field 20
- 7.4 Tensor Products with Derivative Operator 20
- 7.5 Curl in Three Dimensions 21
- 7.6 Coordinate Forms of Various Terms 21

8 Bibliographical Notes **23**

1 Tangent Vector Space on Smooth Manifolds

A n-dimensional smooth manifold M is a mathematical object that locally looks like n-dimension Euclidean space. Intuitively, on each "patch" of the manifold, say Ω, we can define a smooth one-to-one mapping between points $p \in \Omega$ in that region to $\mathbb{R}^n$. By a "smooth mapping" we mean one for which continuous partial derivatives of all orders exist. Such a mapping is called a *coordinate system*. A collection of such patches can be constructed that completely cover the manifold. An example of a manifold is the surface of a sphere. The 3D space of classical physics is also a manifold, however, a very simple one in that is *globally* Euclidean. This tutorial shows how to formulate physics on such manifolds without explicit reference to the coordinate systems on any particular patch. Such formulations are called *geometrical* formulations. Our 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.

Our fundamental building block will be *vectors*, i.e directed line segments. Consider a point $p \in M$. At this point we will assume we can construct a linear vector space that consist of all vectors anchored at that point and that are "tangential" to the manifold. That is, given two tangent vectors $\mathbf{a}$ and $\mathbf{b}$ anchored at p, we can add them and produce another vector in the usual way geometrical way of constructing parallelograms. We will also assume that we can compute the *dot-product* $\mathbf{a} \cdot \mathbf{b}$ of the tangent vectors. Note that one can't add vectors anchored at two *different points*: just because I pull my car in Princeton and you pull your car in Seattle, does not mean we can add the force vectors representing our individual forces! To emphasize the concept of tangent vector spaces, we will denote the space of tangent vectors anchored at point p by T_pM. This is called the *tangent-space* of the manifold at point p. We will assume that our manifold allows us to construct a tangent-space at each point $p \in M$. Further, we will also assume that we can compute the dot-product of any pair of vectors in T_pM. 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* of the manifold. We will discuss *tensors* later in this tutorial, however, for now it suffices to say that tensors are multi-linear functions of vectors.

To precisely define what we mean by a tangent-space consider a patch Ω of M and a smooth, invertible mapping, φ, from $p \in \Omega$ to $\mathbb{R}^n$. That is, $\varphi : \Omega \to \mathbb{R}^n$. The *inverse* mapping φ^{-1} hence maps a point $z^1, \ldots, z^n$ in $\mathbb{R}^n$ to a point $\mathbf{x} \in \Omega$ on the manifold:

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

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

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

As φ^{-1} is smooth, the basis vectors are also smooth in the patch $\Omega \subseteq M$. Such a set of vectors generated from the coordinate mappings are called *coordinate basis*. Note that the existence of such a basis in a finite region around each point $p \in M$ is guaranteed by the definition of a smooth manifold. Sometimes, coordinate basis are denoted by $\mathbf{e}_i(\mathbf{x})$ to indicate that they are a *vector fields* defined over a patch of the manifold.

We can represent any vector in T_pM 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 $\mathbf{a} \in T_pM$ 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* T_pM as the basis-vectors themselves. Hence, for any $\mathbf{a} \in T_pM$ 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.

Given a coordinate basis we can compute it's reciprocal basis-set in the same way as defined above. For coordinate basis, however, there is another way to express the reciprocal basis. Consider the transform φ defined above. A point $\mathbf{x} \in \Omega$ is mapped to the coordinates $z^i(\mathbf{x})$, $i = 1, \ldots, n$ in $\mathbb{R}^n$. Clearly, the reciprocal basis are merely gradients of the 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.

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

2 Euclidean Space

The arena of classical physics is a simple 3D Euclidean manifold. In this special case we can cover the entire manifold with a single Cartesian coordinate system (x, y, z). We will denote unit vectors in this coordinate system as $\boldsymbol{\sigma}_\mu$, where μ and other Greek letters will range over x, y, z. At each point p in Euclidean space the tangent vector space T_pM can also use the same basis-vectors $\boldsymbol{\sigma}_\mu$, which also serve as the reciprocals basis. Note than in general, when the manifold is not Euclidean, T_pM and the manifold itself are not the same spaces. For example, on the surface of a sphere, at a point p the tangent-vector space T_pM can be visualized as the 2D plane tangent to the sphere at that point.

Even in Euclidean space we often want to use coordinate systems that are not Cartesian. In this case again we can introduce mappings from the non-Cartesian coordinate system to the global Cartesian system in the form Eq. (1). Consider the following example.

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

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

From this the tangent vectors (coordinate basis) are

$$\mathbf{e}_r = \frac{\partial \mathbf{x}}{\partial r} = \cos\theta\,\boldsymbol{\sigma}_x + \sin\theta\,\boldsymbol{\sigma}_y \tag{9}$$

$$\mathbf{e}_\theta = \frac{\partial \mathbf{x}}{\partial \theta} = -r\sin\theta\,\boldsymbol{\sigma}_x + r\cos\theta\,\boldsymbol{\sigma}_y. \tag{10}$$

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

$$\mathbf{e}^r = \cos\theta\,\boldsymbol{\sigma}_x + \sin\theta\,\boldsymbol{\sigma}_y \tag{11}$$

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

Example 2. With polar coordinates defined in the above example, let $\mathbf{u} = u_x\mathbf{e}_x + u_y\mathbf{e}_y$ 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

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

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

$$\mathbf{x} = r\sin\theta\cos\phi\,\boldsymbol{\sigma}_x + r\sin\theta\sin\phi\,\boldsymbol{\sigma}_y + r\cos\theta\boldsymbol{\sigma}_z. \tag{22}$$

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

For the basis $\mathbf{e}_i$, $i = 1, 2, 3$ in 3D space, let $\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3) = \mathcal{J}$. As $\mathbf{e}_1 \cdot \mathbf{e}^1 = 1$, this expression (and permutations of it) allows us to compute explicit expressions for the reciprocal basis as

$$\mathbf{e}^1 = \frac{1}{\mathcal{J}}\mathbf{e}_2 \times \mathbf{e}_3 \tag{23}$$
$$\mathbf{e}^2 = \frac{1}{\mathcal{J}}\mathbf{e}_3 \times \mathbf{e}_1 \tag{24}$$
$$\mathbf{e}^3 = \frac{1}{\mathcal{J}}\mathbf{e}_1 \times \mathbf{e}_2. \tag{25}$$

A simple calculation also shows that $\mathbf{e}^1 \cdot (\mathbf{e}^2 \times \mathbf{e}^3) = \mathcal{J}^{-1}$.

3 Tensor Products and Tensors

Given two or more vectors in T_pM, we will denote their *tensor product* with the $\otimes$ symbol. For example, given $\mathbf{u}, \mathbf{v} \in T_pM$ we can write their tensor-product as $\mathbf{u} \otimes \mathbf{v}$[2]. The tensor-product creates a *multilinear mapping* from n vectors, where n is the number of vectors in the product, to

[2] In the literature, often $\mathbf{u} \otimes \mathbf{v}$ is written as $\mathbf{uv}$. We will use the more explicit notation for the tensor-product, leaving the "invisible" operator in $\mathbf{uv}$ between $\mathbf{u}$ and $\mathbf{v}$ to mean the *geometric product* defined below.

a real number. Given $\mathbf{u} \otimes \mathbf{v}$, the mapping is $\mathbf{u} \otimes \mathbf{v} : T_pM \times T_pM \to \mathbb{R}$, and can evaluate it for the vectors $\mathbf{a}$ and $\mathbf{b}$ as follows

$$(\mathbf{u} \otimes \mathbf{v})(\mathbf{a}, \mathbf{b}) = (\mathbf{u} \cdot \mathbf{a})(\mathbf{v} \cdot \mathbf{b}). \tag{26}$$

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

$$\mathbf{T}(\alpha\mathbf{a} + \delta\mathbf{d}, \mathbf{b}) = \alpha\mathbf{T}(\mathbf{a}, \mathbf{b}) + \delta\mathbf{T}(\mathbf{d}, \mathbf{b}). \tag{27}$$

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

$$\mathbf{u}(\mathbf{a}) = \mathbf{u} \cdot \mathbf{a}. \tag{28}$$

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

$$\mathbf{u}(\mathbf{e}_i) = \mathbf{u} \cdot \mathbf{e}_i = u_i \tag{29}$$

and

$$\mathbf{u}(\mathbf{e}^i) = \mathbf{u} \cdot \mathbf{e}^i = u^i. \tag{30}$$

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

$$T_{ij} \equiv \mathbf{T}(\mathbf{e}_i, \mathbf{e}_j) \tag{31}$$

and

$$T^{ij} \equiv \mathbf{T}(\mathbf{e}^i, \mathbf{e}^j). \tag{32}$$

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,

$$\mathbf{T} = T^{ij}\mathbf{e}_i \otimes \mathbf{e}_j = T_{ij}\mathbf{e}^i \otimes \mathbf{e}^j \tag{33}$$

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

$$\mathbf{T}(\mathbf{a}, \mathbf{b}) = T^{ij}(\mathbf{a} \cdot \mathbf{e}_i)(\mathbf{b} \cdot \mathbf{e}_j) = T_{ij}(\mathbf{a} \cdot \mathbf{e}^i)(\mathbf{b} \cdot \mathbf{e}^j). \tag{34}$$

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

$$\mathbf{T}(\mathbf{a}, \_) = T^{ij}\mathbf{a} \cdot (\breve{\mathbf{e}}_i \otimes \mathbf{e}_j) = T^{ij}(\mathbf{a} \cdot \mathbf{e}_i)\mathbf{e}_j \tag{35}$$

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,

$$\mathbf{T}(\_, \mathbf{a}) = T^{ij}\mathbf{a} \cdot (\mathbf{e}_i \otimes \breve{\mathbf{e}}_j) = T^{ij}\mathbf{e}_i(\mathbf{a} \cdot \mathbf{e}_j). \tag{36}$$

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,

$$\mathbf{T}(\hat{\mathbf{e}}_i, \hat{\mathbf{e}}_j) \equiv \hat{T}_{ij} = T_{mn}(\hat{\mathbf{e}}_i \cdot \mathbf{e}^m)(\hat{\mathbf{e}}_j \cdot \mathbf{e}^n). \tag{37}$$

Exercise 2. 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

$$\mathrm{Tr}(\mathbf{u} \otimes \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}. \tag{38}$$

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

$$\mathrm{Tr}(\breve{\mathbf{u}} \otimes \mathbf{v} \otimes \breve{\mathbf{w}}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v}. \tag{39}$$

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

$$\mathbf{g}(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}. \tag{40}$$

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

$$\mathbf{a} \cdot \mathbf{g}(\breve{\_}, \_) = \mathbf{g}(\mathbf{a}, \_) = \mathbf{a}. \tag{41}$$

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

$$\mathbf{g}(\mathbf{e}_i, \mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j \equiv g_{ij}. \tag{42}$$

$$\mathbf{g}(\mathbf{e}^i, \mathbf{e}^j) = \mathbf{e}^i \cdot \mathbf{e}^j \equiv g^{ij}. \tag{43}$$

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

$$u^i = \mathbf{u} \cdot \mathbf{e}^i = (\mathbf{u} \cdot \mathbf{e}_j)(\mathbf{e}^j \cdot \mathbf{e}^i) = g^{ij} u_j. \tag{44}$$

Similarly, we have

$$u_i = \mathbf{u} \cdot \mathbf{e}_i = (\mathbf{u} \cdot \mathbf{e}^j)(\mathbf{e}_j \cdot \mathbf{e}_i) = g_{ij} u^j. \tag{45}$$

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

$$\mathbf{e}^i = g^{ij} \mathbf{e}_j \tag{46}$$

$$\mathbf{e}_i = g_{ij} \mathbf{e}^j. \tag{47}$$

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

Exercise 3. 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 $f : T_p M \to \mathbb{R}$ 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*.

4 The Fundamental Derivative Operator

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

$$\nabla = \mathbf{e}^i \frac{\partial}{\partial z^i}. \tag{48}$$

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 in the patch of the manifold, ensures that we can compute their partial derivatives with respect to the coordinates.

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

$$\nabla f = \mathbf{e}^r \frac{\partial f}{\partial r} + \mathbf{e}^\theta \frac{\partial f}{\partial \theta} = \mathbf{e}_r \frac{\partial f}{\partial r} + \frac{\mathbf{e}_\theta}{r^2} \frac{\partial f}{\partial \theta} = \hat{\mathbf{e}}_r \frac{\partial f}{\partial r} + \frac{\hat{\mathbf{e}}_\theta}{r} \frac{\partial f}{\partial \theta}. \tag{49}$$

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

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

$$\nabla = g^{ij} \mathbf{e}_j \frac{\partial}{\partial z^i} = g^{ij} \|\mathbf{e}_{\underline{j}}\| \hat{\mathbf{e}}_{\underline{j}} \frac{\partial}{\partial z^i} \tag{50}$$

As with vectors, the last representation in normalized tangent vectors allows interpreting the gradient with the 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.

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. (48) 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:

$$\mathbf{u} \otimes \overleftarrow{\nabla} \equiv \frac{\partial \mathbf{u}}{\partial z^i} \otimes \mathbf{e}^i. \tag{51}$$

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

$$\nabla \cdot \mathbf{u} = \mathrm{Tr}(\nabla \otimes \mathbf{u}) = \mathrm{Tr}\Big(\mathbf{u} \otimes \overleftarrow{\nabla}\Big). \tag{52}$$

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 on indicating 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

$$\nabla \cdot (\nabla f) = (\nabla \cdot \nabla) f \equiv \nabla^2 f. \tag{53}$$

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

$$\nabla \cdot (\breve{\nabla} \otimes \mathbf{u}) = (\nabla \cdot \nabla)\mathbf{u} \equiv \nabla^2 \mathbf{u}. \tag{54}$$

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

$$\nabla \cdot \mathbf{g}(\breve{\_}, \_) = \mathbf{e}^k \cdot \frac{\partial}{\partial z^k}(\breve{\mathbf{e}}_i \otimes \mathbf{e}^i) = (\mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial z^k})\mathbf{e}^i + \underbrace{\mathbf{e}^k \cdot \mathbf{e}_i}_{\delta^k{}_i} \frac{\partial \mathbf{e}_i}{\partial z^k}. \tag{55}$$

Now using the chain-rule

$$\mathbf{e}^k \cdot \frac{\partial \mathbf{e}_i}{\partial z^k} = \underbrace{\frac{\partial}{\partial z^k}(\mathbf{e}^k \cdot \mathbf{e}_i)}_{0} - \mathbf{e}_i \cdot \frac{\partial \mathbf{e}_k}{\partial z^k}. \tag{56}$$

Hence,

$$\nabla \cdot \mathbf{g} = -(\mathbf{e}_i \cdot \frac{\partial \mathbf{e}_k}{\partial z^k})\mathbf{e}^i + \frac{\partial \mathbf{e}_k}{\partial z^k} = 0. \tag{57}$$

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

Exercise 4. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors, and f be a scalar. Use the definition Eq. (48) to prove the following

$$\nabla \otimes (\mathbf{a} \otimes \mathbf{b}) = (\nabla \otimes \mathbf{a}) \otimes \mathbf{b} + \mathbf{a} \otimes (\nabla \otimes \mathbf{b}) \tag{58}$$

$$\nabla \otimes (f\mathbf{a}) = \nabla \otimes (\mathbf{a}f) = \nabla f \otimes \mathbf{a} + f \nabla \otimes \mathbf{a} \tag{59}$$

$$\nabla \cdot (\breve{\mathbf{a}} \otimes \mathbf{b}) = \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{a} \cdot \nabla)\mathbf{b} \tag{60}$$

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (\nabla \otimes \breve{\mathbf{b}}) + \mathbf{b} \cdot (\nabla \otimes \breve{\mathbf{a}}). \tag{61}$$

5 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 cure to this was discovered in the 19th century by Grassman and built upon by Clifford, in the form of Geometric Algebra (GA). GA allows a complete description of geometry in all dimensions. We will discuss this at a later part of this tutorial. 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 $\varepsilon(\mathbf{a}, \mathbf{b}, \mathbf{c})$ as

$$\varepsilon(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}). \tag{62}$$

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

$$\varepsilon(\_, \mathbf{b}, \mathbf{c}) = \mathbf{b} \times \mathbf{c}, \tag{63}$$

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, $\varepsilon(\_, \_, \mathbf{c})$, has the property that

$$\mathbf{b} \cdot \varepsilon(\_, \breve{\_}, \mathbf{c}) = \varepsilon(\_, \mathbf{b}, \mathbf{c}) = \mathbf{b} \times \mathbf{c}. \tag{64}$$

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 5. Show that ε is anti-symmetric in each pair of slots. Such tensors are called *totally anti-symmetric* tensors.

Exercise 6. Consider the components of ε along a basis, for example, $\varepsilon(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k)$. Show that all non-zero components are identical, except for a sign.

Exercise 7. Show that

$$\varepsilon(\_, \breve{\_}, \mathbf{B}) \cdot \nabla = -\mathbf{B} \times \nabla. \tag{65}$$

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

$$\mathbf{u} \times \mathbf{P}(\breve{\_}, \_) = P^{mn}(\mathbf{u} \times \mathbf{e}_m) \otimes \mathbf{e}_n. \tag{66}$$

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

Exercise 8. 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 ε tensor above, we can write the curl instead as a *divergence* as

$$\nabla \times \mathbf{B} = \nabla \cdot [\varepsilon(\_, \breve{\_}, \mathbf{B})]. \tag{67}$$

Exercise 9. Show that this is indeed true.

6 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.

6.1 The Advection-Diffusion Equation

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

$$\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{u} f) = \nu \nabla^2 f \tag{68}$$

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 $\nabla \cdot [\mathbf{D}(\_, \nabla f)]$.

$$\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{u} f) = \nabla \cdot [\mathbf{D}(\nabla f, \_)] \tag{69}$$

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

Exercise 10. 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 11. Consider the diffusion tensor for heat-flux in a magnetized plasma. This is given by

$$\mathbf{D} = \kappa_\perp (\mathbf{g} - \mathbf{b} \otimes \mathbf{b}) + \kappa_\parallel \mathbf{b} \otimes \mathbf{b} - \kappa_\wedge \boldsymbol{\varepsilon}(\_, \mathbf{b}, \_) \tag{70}$$

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

$$\nabla \cdot [\mathbf{D}(\nabla f, \_)] = \kappa_\perp \nabla^2 f + (\kappa_\parallel - \kappa_\perp) \nabla \cdot [\mathbf{b}\,(\mathbf{b} \cdot \nabla f)] + \kappa_\wedge \mathbf{b} \times \nabla f. \tag{71}$$

6.2 Maxwell Equations

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

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0 \tag{72}$$

$$\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} = -\mu_0 \mathbf{J} \tag{73}$$

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

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{E})] = 0 \tag{74}$$

$$\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} - \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B})] = -\mu_0 \mathbf{J} \tag{75}$$

Taking the dot product of Eq. (72) by $\mathbf{B}/\mu_0$ and Eq. (73) by $\mathbf{E}/\mu_0$ and adding the two resulting equations gives the evolution equation for electromagnetic energy

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{B}^2}{2\mu_0} + \frac{\epsilon_0 \mathbf{E}^2}{2}\right) + \frac{1}{\mu_0}\nabla \cdot (\mathbf{E} \times \mathbf{B}) = -\mathbf{E} \cdot \mathbf{J}. \tag{76}$$

In deriving this expression we have used the identity

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{E} - \mathbf{E} \cdot \nabla \times \mathbf{B}. \tag{77}$$

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. (72) by $\epsilon_0 \mathbf{E}$ and Eq. (73) by $\mathbf{B}/\mu_0$ and subtract the two resulting equations to get

$$\epsilon_0 \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) + \epsilon_0 \mathbf{E} \times (\nabla \times \mathbf{E}) + \frac{1}{\mu_0}\mathbf{B} \times (\nabla \times \mathbf{B}) = -\mathbf{J} \times \mathbf{B}. \tag{78}$$

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

$$\mathbf{A} \times (\nabla \times \mathbf{A}) = \nabla \cdot \left[\frac{1}{2}\mathbf{A}^2 \mathbf{g}(\_, \breve{\_}) - \mathbf{A} \otimes \breve{\mathbf{A}}\right] + \mathbf{A}(\nabla \cdot \mathbf{A}). \tag{79}$$

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

$$\epsilon_0 \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) + \nabla \cdot \mathbf{T}(\_, \breve{\_}) = -\varrho_c \mathbf{E} - \mathbf{J} \times \mathbf{B}, \tag{80}$$

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

$$\mathbf{T} = \left(\frac{\mathbf{B}^2}{2\mu_0} + \frac{\epsilon_0 \mathbf{E}^2}{2}\right)\mathbf{g} - \left(\epsilon_0 \mathbf{E} \otimes \mathbf{E} + \frac{1}{\mu_0}\mathbf{B} \otimes \mathbf{B}\right). \tag{81}$$

6.3 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

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{82}$$

and the *momentum equation*

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \frac{\nabla p}{\rho} = \frac{1}{\mu_0 \rho}(\nabla \times \mathbf{B}) \times \mathbf{B} \tag{83}$$

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

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0 \tag{84}$$

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

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \nabla p = -\gamma p \nabla \cdot \mathbf{u}. \tag{85}$$

The momentum equation can be written in conservative form as follows. We use the identity Eq. (79) and the fact that $\nabla \cdot \mathbf{B} = 0$ to write

$$(\nabla \times \mathbf{B}) \times \mathbf{B} = \nabla \cdot \left[-\frac{1}{2} \mathbf{B}^2 \mathbf{g}(\_, \breve{\_}) + \mathbf{B} \otimes \breve{\mathbf{B}} \right] \tag{86}$$

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

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \nabla \cdot \mathbf{T}(\_, \breve{\_}) = 0 \tag{87}$$

where

$$\mathbf{T} = \rho \mathbf{u} \otimes \mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \mathbf{g} - \frac{1}{\mu_0} \mathbf{B} \otimes \mathbf{B}. \tag{88}$$

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

The induction equation can be written in divergence form as

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{u} \times \mathbf{B})] = 0. \tag{89}$$

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

$$\boldsymbol{\varepsilon}(\_, \mathbf{a}, \mathbf{u} \times \mathbf{B}) = \mathbf{g}(\mathbf{a}, \mathbf{B})\mathbf{u} - \mathbf{g}(\mathbf{a}, \mathbf{u})\mathbf{B} \tag{90}$$

and hence

$$\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{u} \times \mathbf{B}) = \mathbf{u} \otimes \mathbf{g}(\breve{\_}, \mathbf{B}) - \mathbf{B} \otimes \mathbf{g}(\breve{\_}, \mathbf{u}) = \mathbf{u} \otimes \mathbf{B} - \mathbf{B} \otimes \mathbf{u} \tag{91}$$

(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

$$\frac{\partial \mathbf{B}}{\partial t} - \nabla \cdot \left[\mathbf{u} \otimes \breve{\mathbf{B}} - \mathbf{B} \otimes \breve{\mathbf{u}} \right] = 0. \tag{92}$$

6.3.1 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. (87) with $\mathbf{u}$ to get

$$\underbrace{\mathbf{u} \cdot \frac{\partial}{\partial t}(\rho \mathbf{u})}_{\frac{\partial}{\partial t}(\frac{1}{2}\rho \mathbf{u}^2) + \frac{1}{2}\mathbf{u}^2 \frac{\partial \rho}{\partial t}} \quad + \mathbf{u} \cdot [\nabla \cdot \mathbf{T}(\_, \breve{\_})] = 0. \tag{93}$$

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

$$\nabla \cdot \mathbf{T}(\mathbf{u}, \_) = \mathbf{u} \cdot [\nabla \cdot \mathbf{T}(\_, \breve{\_})] + [\mathbf{T}(\_, \breve{\_}) \cdot \nabla] \cdot \mathbf{u}. \tag{94}$$

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

$$\mathbf{T}(\mathbf{u}, \_) = \rho \mathbf{u}^2 \mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \mathbf{u} - \frac{1}{\mu_0} (\mathbf{B} \cdot \mathbf{u}) \mathbf{B} \tag{95}$$

and

$$[\mathbf{T}(\_, \breve{\_}) \cdot \nabla] \cdot \mathbf{u} = \underbrace{\rho \mathbf{u} \cdot [(\mathbf{u} \cdot \nabla)\mathbf{u}]}_{\rho \mathbf{u} \cdot \nabla \frac{1}{2} \mathbf{u}^2} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \nabla \cdot \mathbf{u} - \frac{1}{\mu_0} \mathbf{B} \cdot [(\mathbf{B} \cdot \nabla)\mathbf{u}]. \tag{96}$$

Now, we can show that

$$\mathbf{B} \cdot [(\mathbf{B} \cdot \nabla)\mathbf{u}] = \nabla \cdot [\mathbf{B}(\mathbf{u} \cdot \mathbf{B})] - \mathbf{u} \cdot [(\mathbf{B} \cdot \nabla)\mathbf{B}]. \tag{97}$$

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

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho \mathbf{u}^2 \right) + \nabla \cdot \left[\frac{1}{2} \rho \mathbf{u}^2 \mathbf{u} + \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \mathbf{u} \right] = \left(p + \frac{\mathbf{B}^2}{2\mu_0} \right) \nabla \cdot \mathbf{u} + \frac{1}{\mu_0} \mathbf{u} \cdot [(\mathbf{B} \cdot \nabla)\mathbf{B}]. \tag{98}$$

The pressure evolution equation Eq. (85) can be rearranged as

$$\frac{\partial}{\partial t}\left(\frac{p}{\gamma-1}\right)+\nabla\cdot\left(\mathbf{u}\frac{p}{\gamma-1}\right)=-p\nabla\cdot\mathbf{u}. \tag{99}$$

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. (92), by $\mathbf{B}/\mu_0$. To simplify the second term we again use the identity Eq. (94) to write

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{B}^2}{2\mu_0}\right) \underbrace{-\frac{1}{\mu_0}\nabla\cdot\left[(\mathbf{u}\cdot\mathbf{B})\mathbf{B}-\mathbf{B}^2\mathbf{u}\right]}_{-\nabla\cdot\left(\frac{1}{\mu_0}(\mathbf{u}\cdot\mathbf{B})\mathbf{B}-\frac{\mathbf{B}^2}{2\mu_0}\mathbf{u}\right)+\mathbf{u}\cdot\nabla\left(\frac{\mathbf{B}^2}{2\mu_0}\right)+\frac{\mathbf{B}^2}{2\mu_0}\nabla\cdot\mathbf{u}} +\frac{1}{\mu_0}\mathbf{u}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{B}] \underbrace{-\frac{1}{\mu_0}\mathbf{B}\cdot[(\mathbf{u}\cdot\nabla)\mathbf{B}]}_{-\mathbf{u}\cdot\nabla\left(\frac{\mathbf{B}^2}{2\mu_0}\right)} \tag{100}$$

or

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{B}^2}{2\mu_0}\right)-\nabla\cdot\left(\frac{1}{\mu_0}(\mathbf{u}\cdot\mathbf{B})\mathbf{B}-\frac{\mathbf{B}^2}{2\mu_0}\mathbf{u}\right)=-\frac{\mathbf{B}^2}{2\mu_0}\nabla\cdot\mathbf{u}-\frac{1}{\mu_0}\mathbf{u}\cdot[(\mathbf{B}\cdot\nabla)\mathbf{B}]. \tag{101}$$

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

$$\mathcal{E}\equiv\frac{1}{2}\rho\mathbf{u}^2+\frac{p}{\gamma-1}+\frac{\mathbf{B}^2}{2\mu_0} \tag{102}$$

evolves as

$$\frac{\partial\mathcal{E}}{\partial t}+\nabla\cdot\left[(\mathcal{E}+p^*)\mathbf{u}-\frac{1}{\mu_0}(\mathbf{u}\cdot\mathbf{B})\mathbf{B}\right]=0 \tag{103}$$

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

$$p^*\equiv p+\frac{\mathbf{B}^2}{2\mu_0}. \tag{104}$$

Exercise 12. Prove the identity Eq. (94).

Exercise 13. Prove that Eq. (97) is indeed true.

6.4 The Navier-Stokes Equations

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

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{105}$$

the momentum equation

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \nabla \cdot \mathbf{T}(\_, \breve{\_}) = 0, \tag{106}$$

and the energy equation

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot [\mathbf{u}(\mathcal{E} + p) - \boldsymbol{\tau}(\mathbf{u}, \_) + \mathbf{q}] = 0. \tag{107}$$

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

$$\mathcal{E} = \frac{1}{2}\rho \mathbf{u} \cdot \mathbf{u} + \frac{p}{\gamma - 1} \tag{108}$$

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

$$\mathbf{T} = \rho \mathbf{u} \otimes \mathbf{u} + p\mathbf{g} - \boldsymbol{\tau} \tag{109}$$

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

$$\boldsymbol{\tau} = \mu \left(\nabla \otimes \mathbf{u} + \mathbf{u} \otimes \overleftarrow{\nabla} - \frac{2}{3}(\nabla \cdot \mathbf{u})\, \mathbf{g} \right) + \lambda (\nabla \cdot \mathbf{u})\, \mathbf{g} \tag{110}$$

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

$$\mathbf{q} = -\kappa \nabla \varepsilon \tag{111}$$

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

7 Coordinate Representations, Choice of Basis

In this section we work out the various operators in generic coordinates on a patch of the manifold. 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.

7.1 Divergence of Vector Field

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

$$\nabla \cdot \mathbf{u} = \mathbf{e}^i \cdot \frac{\partial \mathbf{u}}{\partial z^i}. \tag{112}$$

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

$$\nabla \cdot \mathbf{u} = \mathbf{e}^i \cdot \left(\frac{\partial u^j}{\partial z^i} \mathbf{e}_j + u^j \frac{\partial \mathbf{e}_j}{\partial z^i} \right) = \frac{\partial u^i}{\partial z^i} + \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_j}{\partial z^i} u^j. \tag{113}$$

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

$$\frac{\partial \mathbf{e}_j}{\partial z^i} = \Gamma^m{}_{ji} \mathbf{e}_m \tag{114}$$

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

$$\nabla \cdot \mathbf{u} = \frac{\partial u^i}{\partial z^i} + \Gamma^i{}_{ji} u^j. \tag{115}$$

The Christoffel symbols can be computed by taking the dot product of Eq. (114) by $\mathbf{e}^m$ to get

$$\Gamma^m{}_{ij} = \mathbf{e}^m \cdot \frac{\partial \mathbf{e}_i}{\partial z^j}. \tag{116}$$

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

$$\frac{\partial \mathbf{e}_i}{\partial z^j} = \frac{\partial}{\partial z^j} \left(\frac{\partial \mathbf{x}}{\partial z^i} \right) = \frac{\partial}{\partial z^i} \left(\frac{\partial \mathbf{x}}{\partial z^j} \right) = \frac{\partial \mathbf{e}_j}{\partial z^i} \tag{117}$$

showing that

$$\Gamma^m{}_{ij} = \Gamma^m{}_{ji}. \tag{118}$$

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

$$\frac{1}{\mathcal{J}} \frac{\partial \mathcal{J}}{\partial z^j} = \mathbf{e}^i \cdot \frac{\partial \mathbf{e}_j}{\partial z^i}. \tag{119}$$

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

Now, in terms of the Christoffel symbols we can write the identity Eq. (119) as

$$\frac{1}{\mathcal{J}}\frac{\partial \mathcal{J}}{\partial z^j} = \Gamma^i{}_{ij} \tag{120}$$

Using this, we can write the divergence, Eq. (115), as

$$\nabla \cdot \mathbf{u} = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J} u^i\right). \tag{121}$$

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

Exercise 14. Use $\mathbf{e}_j \cdot \mathbf{e}^k = \delta_j{}^k$ to show that

$$\mathbf{e}_j \cdot \frac{\partial \mathbf{e}^k}{\partial z^i} = -\frac{\partial \mathbf{e}_j}{\partial z^i} \cdot \mathbf{e}^k = -\Gamma^k{}_{ji}. \tag{122}$$

Exercise 15. Take the derivative of $g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$, cyclically permute indices and use the symmetry property Eq. (117) to express the Christoffel symbols in terms of the metric tensor as

$$\Gamma^n{}_{kj} = \frac{1}{2} g^{nm}\left(\frac{\partial g_{mk}}{\partial z^j} + \frac{\partial g_{jm}}{\partial z^k} - \frac{\partial g_{kj}}{\partial z^m}\right). \tag{123}$$

Exercise 16. 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

$$\nabla \cdot \mathbf{u} = g^{ij}\left(\frac{\partial u_j}{\partial z_i} - u_k \Gamma^k{}_{ji}\right). \tag{124}$$

Exercise 17. 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

$$\nabla \cdot \mathbf{u} = \frac{1}{r}\frac{\partial}{\partial r}\left(r \hat{u}^r\right) + \frac{1}{r}\frac{\partial \hat{u}^\theta}{\partial \theta}. \tag{125}$$

7.2 Laplacian of a Scalar Field

An immediate and important application of Eq. (121) is to compute the Laplacian of a scalar field f. The gradient can be written in the tangent vectors as

$$\nabla f = g^{ij}\frac{\partial f}{\partial z^i}\mathbf{e}_j. \tag{126}$$

Using this in Eq. (121) to take the divergence, we get the *Laplacian*

$$\nabla^2 f = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J} g^{ij}\frac{\partial f}{\partial z^i}\right). \tag{127}$$

Exercise 18. Show that in polar coordinates, the Laplacian is

$$\nabla^2 f = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2}. \tag{128}$$

7.3 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

$$\nabla\cdot\mathbf{T}(\breve{\_},\_) = \mathbf{e}^i\cdot\frac{\partial}{\partial z^i}\left(T^{mn}\breve{\mathbf{e}}_m\otimes\mathbf{e}_n\right) = \frac{\partial T^{in}}{\partial z^i}\mathbf{e}_n + T^{mn}\mathbf{e}^i\cdot\frac{\partial\mathbf{e}_m}{\partial z^i} + T^{in}\frac{\partial\mathbf{e}_n}{\partial z^i}. \tag{129}$$

Using the definition of the Christoffel symbols we get

$$\nabla\cdot\mathbf{T}(\breve{\_},\_) = \left(\frac{\partial T^{in}}{\partial z^i} + T^{mn}\Gamma^i{}_{mi} + T^{im}\Gamma^n{}_{im}\right)\mathbf{e}_n. \tag{130}$$

Using the identity Eq. (121) we can write this instead in a more compact form as

$$\nabla\cdot\mathbf{T}(\breve{\_},\_) = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}T^{in}\mathbf{e}_n\right) = \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}\mathbf{T}(\mathbf{e}^i,\_)\right) \tag{131}$$

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

7.4 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

$$\nabla\otimes\mathbf{u} = \mathbf{e}^i\otimes\frac{\partial}{\partial z^i}(u^m\mathbf{e}_m) = \underbrace{\left(\frac{\partial u^m}{\partial z^i} + u^k\Gamma^m{}_{ki}\right)}_{\equiv u^m{}_{;i}}\mathbf{e}^i\otimes\mathbf{e}_m. \tag{132}$$

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

$$\nabla\otimes\mathbf{u} = u^m{}_{;i}\mathbf{e}^i\otimes\mathbf{e}_m. \tag{133}$$

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

$$\nabla\otimes\mathbf{u} = \mathbf{e}^i\otimes\frac{\partial}{\partial z^i}(u_m\mathbf{e}^m) = \left(\frac{\partial u_m}{\partial z^i} - u_k\Gamma^k{}_{mi}\right)\mathbf{e}^i\otimes\mathbf{e}^m, \tag{134}$$

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

Exercise 20. Show that $\nabla \otimes \mathbf{x} = \mathbf{g}$, where $\mathbf{x}$ is the position vector.

Exercise 21. 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.

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

$$\mathbf{u} \otimes \overleftarrow{\nabla} = \frac{\partial}{\partial z^i}(u^m \mathbf{e}_m) \otimes \mathbf{e}^i = u^m{}_{;i} \mathbf{e}_m \otimes \mathbf{e}^i. \tag{135}$$

7.5 Curl in Three Dimensions

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

$$\nabla \times \mathbf{B} = \mathbf{e}^i \times \frac{\partial}{\partial z^i}\left(B^m \mathbf{e}_m\right) = \frac{\partial B^m}{\partial z^i} \mathbf{e}^i \times \mathbf{e}_m + B^m \mathbf{e}^i \times \frac{\partial \mathbf{e}_m}{\partial z^i} = B^m{}_{;i} \mathbf{e}^i \times \mathbf{e}_m \tag{136}$$

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

$$\nabla \times \mathbf{B} = \nabla \cdot [\boldsymbol{\varepsilon}(\_, \breve{\_}, \mathbf{B})]. \tag{137}$$

Using the expression Eq. (131) we can write this as

$$\nabla \times \mathbf{B} = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i}[\mathcal{J} \boldsymbol{\varepsilon}(\_, \mathbf{e}^i, \mathbf{B})] = \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i}[\mathcal{J}(\mathbf{e}^i \times \mathbf{B})]. \tag{138}$$

Exercise 22. Show that Eq. (136) and Eq. (138) result in the same expression for the the curl.

7.6 Coordinate Forms of Various Terms

The equations listed in Section 6 are completely coordinate independent, i.e. they are formulated using only geometrical quantities. To solve them for specific problems, however, we need to introduce some coordinate system. Besides this freedom to choose the coordinate system, we also have the freedom to choose the *representation* of the various vectors and tensors that appear. These choices are *independent* of each other, and often once choice of the representation is better that another, depending on the problem at hand. The advantage of the geometric forms above is that they remain valid independent of the representation chosen, and specific forms of the equations can be derived by choosing a representation.

Denote the computation coordinates as z^1, z^2, z^3. Most of the terms in the above equations are straightforward to transcribe using the formulas derived in the previous sections. For example, the continuity equation becomes, using the expression of divergence Eq. (121),

$$\frac{\partial \rho}{\partial t} + \frac{1}{\mathcal{J}} \frac{\partial}{\partial z^i}\left(\mathcal{J} \rho u^i\right) = 0. \tag{139}$$

Let us look at the Navier-Stokes momentum equation. This can be written as, on using Eq. (131),

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}T^{ji}\mathbf{e}_j\right) = 0. \tag{140}$$

We now need to make a choice on the representation of the velocity. Say we choose $\mathbf{u} = u^m \mathbf{e}_m$. Then, expanding the right-hand side and taking the dot-product with $\mathbf{e}^m$ we get that the momentum component ρu^m evolves as

$$\frac{\partial}{\partial t}(\rho u^m) + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}T^{mi}\right) + T^{ji}\mathbf{e}^m \cdot \frac{\partial \mathbf{e}_j}{\partial z^i} = 0. \tag{141}$$

Using the definition of the Christoffel symbols we get

$$\frac{\partial}{\partial t}(\rho u^m) + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}T^{mi}\right) + T^{ji}\Gamma^m{}_{ji} = 0. \tag{142}$$

We now see a potential problem: even though the momentum is conserved, the use of the representation $\mathbf{u} = u^m \mathbf{e}_m$ has led to a term that can't be written as a divergence. Hence, a numerical scheme will have difficulties ensuring conservation of momentum. Note that such a non-conservative term does not arise in the continuity equation. The difference in the momentum equation is that involves the divergence of a *tensor*, hence inevitably leading to the non-divergence term. Physically, this term represents non-inertial forces as the coordinate system is not Cartesian.

To get around this, we can represent the velocity in the *global Cartesian* basis $\boldsymbol{\sigma}_\alpha$, where Greek indices like α range over $\{x, y, z\}$. In terms of these the momentum equation can be written as

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}\mathbf{T}(\boldsymbol{\sigma}_\alpha, \mathbf{e}^i)\boldsymbol{\sigma}_\alpha\right) = 0 \tag{143}$$

Noting that $\boldsymbol{\sigma}_\alpha$ are fixed vectors and can be pulled out of the divergence, we get that the *Cartesian* component of momentum ρu_α, evolves as

$$\frac{\partial}{\partial t}(\rho u_\alpha) + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}\left(\mathcal{J}\mathbf{T}(\boldsymbol{\sigma}_\alpha, \mathbf{e}^i)\right) = 0. \tag{144}$$

Here something very curious has occurred: despite using a potentially non-Cartesian coordinate system, the momentum equation is in conservation-law form as the representation of the velocity was chosen along the Cartesian basis.

Maxwell equations can also be written in a similar way. For example, we can write

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}[\mathcal{J}\boldsymbol{\varepsilon}(\_, \mathbf{e}^i, \mathbf{E})] = 0. \tag{145}$$

If we were to express the electromagnetic field using the tangent-vectors as basis, we would have a non-conservative term involving the Christoffel symbols. However, if we instead represent the fields using *Cartesian* basis we get

$$\frac{\partial B_\alpha}{\partial t} + \frac{1}{\mathcal{J}}\frac{\partial}{\partial z^i}[\mathcal{J}\boldsymbol{\varepsilon}(\boldsymbol{\sigma}_\alpha, \mathbf{e}^i, \mathbf{E})] = 0 \tag{146}$$

Again, as for the momentum equation there are no non-conservative terms from the geometry in this formulation.

Exercise 23. Derive the form of the induction equation for the representation $\mathbf{B} = B^m \mathbf{e}_m$.

8 Bibliographical Notes

There are a lot of books on tensor calculus and Riemannian geometry. The book by Lovelock and Rund[1] is very good and available in a cheap Dover edition. Chapter 1 of Thorne and Blanford[2] is a good introduction to tensors without use of coordinates. Incidentally, this is an excellent textbook, 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] David Lovelock and Hanno Rund. *Tensors, Differential Forms, and Variational Principles*. Dover, 1989.

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