

Tensor Calculus: Unlearning Vector Calculus

Wha-Suck Lee¹, Johann Engelbrecht², Rita Moller¹.

(whasuck.lee@up.ac.za, johann.engelbrecht@up.ac.za;
rita.moller@up.ac.za.)

¹Department of Mathematics and Applied Mathematics,
University of Pretoria, Private Bag X20, Hatfield 0028,
Pretoria, South Africa;

²Department of Science, Mathematics and Technology Education,
University of Pretoria, Technika building 1-4, Groenkloof campus,
Pretoria, South Africa.

(MS received Received date; Accepted date)

Tensor calculus is critical in the study of the vector calculus of the surface of a body. Indeed, tensor calculus is a natural step-up for vector calculus. This paper presents some pitfalls of a traditional course in vector calculus in transitioning to tensor calculus. We show how a deeper emphasis on traditional topics such as the Jacobian can serve as a bridge for vector calculus into tensor calculus.

Keywords. Tensor calculus, vector calculus,

Subject. 97140, 97160.

1. Introduction

A typical course in vector calculus sometimes leaves some students with concepts that appear to be scattered and unrelated. Much of this confusion results from an inability to move away from using Cartesian coordinates as a reference point to judge all other coordinate systems. Since high school all the way into the first year of Mathematics, students are exposed to Cartesian coordinates. Furthermore, the distinction between the actual physical point in space and its Cartesian coordinates has not been emphasized. This addiction to the Cartesian coordinate system when applied to the ubiquitous geometric concept of the gradient manifests in the form of the algebraic formulation of the gradient (2.1). Many students perceive this as the geometric definition of the gradient. In Section 2, we show the pitfall of this misconception in order to break away from the addiction to Cartesian coordinates. More precisely, we show (i) that the algebraic undergraduate definition of the gradient works in exactly one coordinate system, the Cartesian coordinate system and (ii) a case at where the algebraic formulation of the gradient (2.1) gives the wrong gradient in a slightly modified Cartesian coordinate system. Tensor calculus is an upgraded vector calculus with improved algebraic formulations which are not tied down to any single coordinate system. We use the tensor calculus notation of [1].

In Section 3, we introduce the idea that that a coordinate transformation is a dualism in the sense of seeing the same object through the different lenses of a different coordinate systems. We present the notion of the Jacobian as a coordinate

transformation that has been linearized by differentiation. The Jacobian is a local linear substitute to the original coordinate transformation which needs not be linear. This coordinate transformation approach then adds a geometric flavour to the well drilled concept of the determinant of a Jacobian as a fudge factor which is the limit of the ratio of the area of the image set (under the coordinate transformation) to the area of the pre-image set. Furthermore, the Jacobian has the critical property of preserving derivatives from one coordinate system to another. Thus, a Jacobian is a linear dualism map between the two dual realities of coordinate systems with the peculiar property of preserving tangents. This property of preserving of tangents explains the fundamental role of the Jacobian in tensor calculus. In tensor calculus, objects are treated as derivatives. In Section 4, we show that even the fundamental objects of linear algebra, coordinate basis vectors, are derivatives. The change of basis matrix in linear algebra is also a derivative (Proposition 3.1). So, in this sense, calculus (differentiation) is more fundamental than linear algebra.

The power of algebra in the form of coordinate systems is undeniable. Tensor calculus is a framework where one constructs an algebraic formulation that is independent of the coordinate system: it works in all coordinate systems, unlike in (2.1) of Section 2. The Jacobian is the correction factor when the same algebraic formulation (which is dependent on the coordinate system) is executed mechanically in two different coordinate systems. Tensor calculus calls objects created by such (coordinate dependent) algebraic formulations with a Jacobian correction factor, *tensors* (equation (6.1), [1]): the inadequate algebraic formulation of the gradient (2.1) is a tensor (see (4.8) of Section 4). Thus the Jacobian is a positive step forward to the final goal of a coordinate-free universal algebraic formulation of the gradient.

In Section 5 we introduce another dualism map called the *metric tensor*. Functional analysts would not accept the Jacobian as a true dualism map. However the metric tensor is what functional analysts would call a dualism map since it produces a dual (contravariant) basis and a dual notion to the (covariant) tensor introduced in Section 4. The covariant tensor and its dual, the contravariant tensor, together provide a method of generating coordinate-free universal algebraic formulations or calculus (Proposition 6.2, Section 6)

Finally in Section 7, we bring together the key vector calculus concepts such as the *Jacobian, coordinate systems, coordinate basis and gradient* into an organized platform where geometric and algebraic notion of the gradient is truly married in the sense that the new algebraic formulation of the gradient (7.1) is independent of the coordinate system. By nature, the geometric object (like the gradient) is independent of any coordinate system and the algebraic formulation is independent of the coordinate system in the sense that the algebraic formulation is universally applicable to any coordinate system (without recourse to a correction factor like a Jacobian).

2. Breaking the addiction to Cartesian coordinates

The *geometric* notion of the gradient $\vec{\nabla} F(P_0)$ is captured by the *arrow/vector* in the domain $\text{Dom}(F) \subset \mathbb{R}^2$ of a scalar field F with the following properties: (i) the *direction* points where F increases quickest from a point $P_0 \in \text{Dom}(F)$ on one level curve to another nearby level curve and (ii) the *magnitude* gives the actual size of

increase. So, the geometric notion of the gradient is independent of the Cartesian coordinate system or any coordinate system. This geometric notion of the gradient is a mathematical formulation of a law in nature that states that natural phenomena (like heat) flow in the direction of least resistance or quickest increase from one temperature level curve to a nearby temperature level curve. So the classical heat equation which requires the geometric notion of the gradient, $\vec{\nabla}F$, is formulated independent of coordinate systems.

In our lectures, we commit the mistake of making no distinction between the geometric notion of the gradient, $\vec{\nabla}F$, and its algebraic formulation by immediately presenting the scalar field F in its Cartesian coordinatized format $F(x, y)$ and the gradient as the pair of numbers:

$$\vec{\nabla}F(x, y) := \begin{pmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{pmatrix}. \quad (2.1)$$

All unsubscripted column vectors will be assumed to be in the familiar (\mathbf{i}, \mathbf{j}) -basis form: $\begin{pmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{pmatrix}_{(\mathbf{i}, \mathbf{j})}$. We denote the induced standard Cartesian coordinate system by Z_i ; all unsubscripted ordered pairs (x, y) will assume coordinatization by the familiar standard Cartesian coordinate system Z_i .

We now show how easy it is to expose the cardinal error of thinking of a vector algebraically as a pair of numbers (which implicitly assumes a coordinate system) instead of as a geometric arrow.

Example 2.1. Consider a simple scalar field F in Cartesian coordinatized format $F(x, y) = x + y$. Then

$$\vec{\nabla}F(x, y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1\mathbf{i} + 1\mathbf{j}$$

by (2.1) where (\mathbf{i}, \mathbf{j}) denotes the Cartesian coordinate basis vectors. Now suppose we present F in terms of a 'stretched' Cartesian coordinate system $Z_{i'}$ induced by the basis vectors $(\mathbf{i}', \mathbf{j}')$ where $\mathbf{i}' = 2\mathbf{i}$ and $\mathbf{j}' = 2\mathbf{j}$ ¹. Then,

$$x := 2x' \quad y := 2y'$$

and it follows that

$$F(x', y') = 2x' + 2y'$$

so

$$\vec{\nabla}F(x', y') = \begin{pmatrix} \frac{\partial F}{\partial x'} \\ \frac{\partial F}{\partial y'} \end{pmatrix}_{(\mathbf{i}', \mathbf{j}')} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}_{(\mathbf{i}', \mathbf{j}')} = 2\mathbf{i}' + 2\mathbf{j}' = 4\mathbf{i} + 4\mathbf{j}.$$

Therefore even with a slightly modified Cartesian coordinate system $(\mathbf{i}', \mathbf{j}')$, the algebraic formulation (2.1) gives the wrong geometric gradient:

$$\vec{\nabla}F(x', y') = 4\vec{\nabla}F(x, y). \quad (2.2)$$

¹The dimensions of the unit square of the $(\mathbf{i}', \mathbf{j}')$ -coordinate system $Z_{i'}$ are twice as large as the dimensions of the unit square of the (\mathbf{i}, \mathbf{j}) -coordinate system, Z_i .

The algebraic formulation (2.1) gives the correct gradient vector only in the Cartesian coordinate system Z_i . This inadequacy should be a warning to maintain addition to Cartesian coordinate systems despite the ease of computation afforded by (2.1).

What went wrong with the $(\mathbf{i}', \mathbf{j}')$ -formulation of the gradient in Example 2.1? We blurred the distinction between the geometric nature of the gradient and the algebraic formulation of the gradient. We did not recognize the differences in the two natures and blindly married the two notions (without respect of their individual natures), hoping the marriage will hold.

Coordinate systems give an algebraic formulation of geometric notions such as the gradient. Algebraic formulations are powerful in solving problems by providing computational power, so we need coordinate systems. For a successful marriage between the geometric formulation and the algebraic formulation, we need *an algebraic formulation in one coordinate system to hold equally well in any other coordinate system*. We call such formulations *coordinate-free universal algebraic formulations*. Specifically to the problem raised in Example 2.1, we shall show coordinate-free universal algebraic formulation of the gradient vector. For this to happen, tools need to be developed to track how one algebraic formulation in one coordinate system gets transformed into another coordinate system.

The Jacobian is one such tool; the Jacobian will be formulated in such a way to lay the foundation of a fundamental principle of tensor calculus that

all coordinate systems are essentially the same.

Indeed, tensor calculus always assumes some coordinate system in the background by virtue of a universal algebraic formulation thus maintaining the power of computation. In the final analysis, tensor calculus chooses a particular coordinate system or particular algebraic formulation most compatible with the problem at hand. So tensor calculus delays the decision to commit to a coordinate system but at the same time has the computational power of the coordinate system hovering over the problem available at any time.

3. The Jacobian as a localized linear coordinate transformation

Consider the scalar field F of Example 2.1. Through the lens of the Cartesian (\mathbf{i}, \mathbf{j}) -coordinate system, Z_i , the object F assumed the simple linear algebraic form $F(x, y) = x + y$. Through the lens of the stretched-Cartesian $(\mathbf{i}', \mathbf{j}')$ -coordinate system, $Z_{i'}$, the same object F assumed an equally simple linear algebraic form $F(x', y') = 2x' + 2y'$. Therefore we regard $F(x', y')$ as the dual of $F(x, y)$.

In traditional vector calculus courses, the concept of a Jacobian is introduced by first considering the *pullback* coordinate transformation $T_{i'}^i : (x', y') \in Z_{i'} \rightarrow (x, y) \in Z_i$ (p. 184-187, [2]),

$$x = f(x', y') = 2x'; y = g(x', y') = 2y'. \quad (3.1)$$

which has the matrix format $T_{i'}^i = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}. \quad (3.2)$$

The term *pullback* emphasizes the point that the Cartesian Z_i -coordinate system takes centre stage and the new $Z_{i'}$ -coordinate system the back stage.

We shall call the pullback coordinate transformation $T_{i'}^i$ a *global* dualism map since $T_{i'}^i$ describes how for each and every point in \mathbb{R}^2 , the $Z_{i'}$ -lens needs to be adjusted to the preferred Z_i -lens. For example, consider the physical point P_0 which when seen through the awkward $Z_{i'}$ -lens has coordinates $(1, 1)_{Z_{i'}}$; the global dualism map $T_{i'}^i$ then gives an algebraic algorithm to re-view that same physical point P_0 as $(2, 2)$ which is now easy to visualize due to the past addiction to the standard Cartesian coordinate system.

More often than not, the global dualism map $T_{i'}^i$ is highly non-linear like the pullback polar coordinate transformation $T_{i'}^i : (x', y') = (r, \theta) \in Z_{i'} \mapsto (x, y) \in Z_i$,

$$x = r \cos \theta; y = r \sin \theta. \quad (3.3)$$

which has the matrix format $T_{i'}^i = \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 0 \end{pmatrix}$:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & 0 \\ \sin \theta & 0 \end{pmatrix} \begin{pmatrix} r \\ \theta \end{pmatrix}. \quad (3.4)$$

Differentiation comes to our rescue: differentiation linearizes non linear objects as linear objects while preserving the type of the object. The differentiated object is the same type of creature as the original object itself. The price to pay is localization. So, the Jacobian $J_{i'}^i$ of the pullback coordinate system, $T_{i'}^i$, is nothing but the differentiated pullback transformation: $J_{i'}^i = DT_{i'}^i$ which will act as a (linear) substitute for the original non-linear global dualism map $T_{i'}^i$ at least locally. Furthermore, $J_{i'}^i$ is also a $Z_{i'} \rightarrow Z_i$ coordinate transformation which we define locally at $P_0 = (x'_0, y'_0)$ as follows:

$$J_{i'}^i(P_0) = DT_{i'}^i(P_0) := \lim_{P \rightarrow P_0} \frac{\|T_{i'}^i(P) - T_{i'}^i(P_0) - J_{i'}^i(P_0)(P - P_0)\|}{\|P - P_0\|} \quad (3.5)$$

Differentiable objects are nothing more than objects that look like ‘straight lines’ locally at a fixed point (x_0, y_0) (p. vii, [3]); that is, a linear transformation that is the same type of creature as the original object: in our case of a derivative of a $(Z_{i'} \rightarrow Z_i)$ -transformation $T_{i'}^i$, the ‘straight line’ is nothing but a linear $(Z_{i'} \rightarrow Z_i)$ -transformation, that is, a 2×2 -matrix that can be shown to be

$$J_{i'}^i(P_0 = (x'_0, y'_0)) = \begin{pmatrix} x_{x'} & x_{y'} \\ y_{x'} & y_{y'} \end{pmatrix}_{(x'_0, y'_0)}. \quad (3.6)$$

So for the nonlinear polar pullback coordinate transformation $T_{i'}^i$ of (3.4), at each fixed point $P_0 = (r_0, \theta_0) \in Z_{i'}$, the familiar polar Jacobian matrix,

$$J_{i'}^i(P_0) = \begin{pmatrix} \cos \theta_0 & -r_0 \sin \theta_0 \\ \sin \theta_0 & r_0 \cos \theta_0 \end{pmatrix}, \quad (3.7)$$

is a linear $(Z_{i'} \rightarrow Z_i)$ substitute of the non-linear $T_{i'}^i : Z_{i'} \rightarrow Z_i$ of (3.4) at P_0 .

For the linear coordinate transformation $T_{i'}^i$ of (3.1), the Jacobian $J_{i'}^i$ is always the change of basis matrix $P_{(i',j') \rightarrow (i,j)} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ for every point P_0 . It is immediate from (3.2), $J_{i'}^i$ coincides with $T_{i'}^i$. For the non-linear coordinate transformation $T_{i'}^i$ of (3.3), the Jacobian J will vary for different P_0 but the differentiability of $T_{i'}^i$ always ensures $J_{i'}^i$ is a well behaved matrix at every fixed point P_0 . Thus,

The Jacobian $J_{i'}^i$ is a linear coordinate transformation that replaces the actual coordinate transformation $T_{i'}^i$ at least locally.

The special case of $J_{i'}^i$ coinciding with $T_{i'}^i$ of (3.1) holds for all pullback coordinate transformation $T_{i'}^i : Z_{i'} \rightarrow Z_i$ associated with the coordinate systems induced by the two different sets of basis vectors.

Proposition 3.1. *Every change-of-basis matrix is a derivative.*

Proof. Let $Z_i, Z_{i'}$ denote the coordinate systems induced by the basis vectors $(\mathbf{i}, \mathbf{j}), (\mathbf{i}', \mathbf{j}')$ respectively. The change in basis matrix $P_{(i',j') \rightarrow (i,j)}$ is nothing but the Jacobian or derivative of the pullback coordinate transformation $T_{i'}^i : Z_{i'} \rightarrow Z_i$ associated with the coordinate system induced by the basis vectors. \square

Unfortunately in many traditional vector calculus courses, the emphasis has always been on the *determinant* of the Jacobian $|J_{i'}^i|$ with little regard to the actual Jacobian $J_{i'}^i$ as a linear coordinate transformation which replaces the actual coordinate transformation $T_{i'}^i$ at least locally (Chapter 4, [2]). The first immediate drawback with this totally algebraic approach is that it prevents the student from seeing the need of the determinant of the Jacobian $|J_{i'}^i|$ when one does a double integration in the pullback space $Z_{i'}$. In the pull back space of $T_{i'}^i$ of (3.1), the area of a unit square in $Z_{i'}$ is 4 times the unit square in the original space Z_i (see footnote 1). So the determinant of the Jacobian $|J_{i'}^i| = 4$, which is the fudge factor when one does a double integration in $Z_{i'}$. The second drawback is that it prevents the student from seeing that the Jacobian locally preserves tangent vectors *exactly*:

Example 3.1. *Consider a curve $\mathbf{R}'(t) := \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$ in the pullback space $Z_{i'}$. The dual curve*

$$\mathbf{R}(t) := \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = T_{i'}^i \mathbf{R}'(t), \quad (3.8)$$

lies in the Z_i -space. By the chain rule, the Jacobian $J_{i'}^i(P_0)$ exactly preserves tangent vectors to any (position) curve at the point $P_0 = (x'(t_0), y'(t_0))_{Z_{i'}}$ on the curve in the $Z_{i'}$ -space to the corresponding (dual) tangent vector at the same physical point $P_0 = (x(t_0), y(t_0))_{Z_i}$ to the dual (position) curve in the Z_i -space.

$$\mathbf{V}(t_0) := \begin{pmatrix} \dot{x}(t_0) \\ \dot{y}(t_0) \end{pmatrix} = J_{i'}^i(P_0) \mathbf{V}'(t_0), \quad (3.9)$$

where $\mathbf{V}'(t) := \begin{pmatrix} \dot{x}'(t) \\ \dot{y}'(t) \end{pmatrix}$.

From (3.5), the Jacobian $J_{\mathbf{i}}^i(P_0)$ is a good *approximation* of a differentiable coordinate transformation $T_{\mathbf{i}}^i$, when it comes to preserving dual positional vectors (see (3.8)). But when it comes to locally preserving dual tangent vectors or derivatives, the (local) preservation is exact by (3.9) thus replacing $T_{\mathbf{i}}^i$ perfectly.

Proposition 3.2. *The Jacobian locally preserves (tangent) derivatives.*

Change-of-basis matrices are derivatives (Proposition 3.1). Given that the Jacobian preserves dual tangent vectors *exactly* (Proposition 3.2), it comes to no surprise that tensor calculus expresses basis vectors (like the familiar Cartesian basis vectors \mathbf{i}, \mathbf{j}) as derivatives, more precisely, tangent vectors. Therefore, in this sense *calculus is more fundamental than linear algebra*.

4. Coordinate basis vectors as derivatives

Let Z denote the Cartesian coordinates $(Z^1, Z^2) := (x, y)$ of a physical point P_0 . Then the vector field

$$\mathbf{R} : Z \mapsto \mathbf{R}(Z) := \overrightarrow{OP_0}$$

can be differentiated with respect to both $Z^1 = x$ and $Z^2 = y$. The vector field \mathbf{R} will be called a position vector field and it is constructed by making a clear distinction between the coordinates Z and the physical arrow $\mathbf{R}(Z)$. Note that the position vector \mathbf{R} is a geometric object independent of the coordinate system and so we use the notation $\mathbf{R}(Z')$ should the position vector be coordinatized by another coordinate system Z' (we leave the symbol \mathbf{R} unprimed).

Fix the coordinates $Z := (x, y)$. Then the vector $\Delta\mathbf{R}(x + \Delta x, y) - \Delta\mathbf{R}(x, y)$ is $\Delta x \mathbf{i}$. Hence the derivative

$$\mathbf{Z}_1 := \frac{\partial \mathbf{R}(Z)}{\partial Z^1} = \frac{\partial \mathbf{R}(Z)}{\partial x} = \mathbf{i}, \quad (4.1)$$

at all points $Z = (x, y)$. Similarly, for the other basis vector \mathbf{j} ,

$$\mathbf{Z}_2 := \frac{\partial \mathbf{R}(Z)}{\partial Z^2} = \mathbf{j}, \quad (4.2)$$

at all points $Z = (x, y)$. So the rate of change of the position vector with respect to the Cartesian coordinate is constant.

Remark 4.1. *Differentiation preserves the type of object and this \mathbf{Z}_1 is a vector field which returns the basis vector \mathbf{i} at each point $Z = (x, y)$. This is precisely the notion of a vector as an equivalence class of arrows!*

Fix the polar coordinates $Z' := (x', y') = (r, \theta)$. Then the vector

$$\Delta\mathbf{R}(r + \Delta r, \theta) - \Delta\mathbf{R}(r, \theta),$$

shall be viewed through the lens of the Cartesian coordinate system for immediate visualization. Here $\Delta\mathbf{R}(r + \Delta r, \theta) - \Delta\mathbf{R}(r, \theta)$ is $\Delta r \mathbf{r}$ where $\mathbf{r}(r, \theta)$ is the unit radial vector at the point with coordinates $Z' = (r, \theta)$. Hence the derivative

$$\mathbf{Z}_{1'} := \frac{\partial \mathbf{R}(Z')}{\partial Z^{1'}} = \frac{\partial \mathbf{R}(Z')}{\partial r} = \mathbf{r}(r, \theta) \quad (4.3)$$

and thus varies for each point $Z' = (r, \theta)$. We therefore have a *non-constant* vector field. Similarly,

$$|\Delta \mathbf{R}(r + \Delta r, \theta) - \Delta \mathbf{R}(r, \theta)| = r2 \sin \frac{\Delta \theta}{2}.$$

So $\|\mathbf{Z}_{2'}\| := \left\| \frac{\partial \mathbf{R}(Z)}{\partial Z^{2'}} \right\| = r$ at the point with coordinates $Z' = (r, \theta)$. Therefore,

$$\mathbf{Z}_{2'} = r\boldsymbol{\theta}(r, \theta) \quad (4.4)$$

where $\boldsymbol{\theta}(r, \theta)$ is $\mathbf{r}(r, \theta)$ rotated clockwise by $\frac{\pi}{2}$. So $(\mathbf{Z}_{1'}, \mathbf{Z}_{2'})$ is an orthogonal basis that varies/moves from point to point.

The Jacobian preserves time derivatives of position curves (Proposition 3.2). The Jacobian similarly relates the ‘derivative’ basis $(\mathbf{Z}_{1'}, \mathbf{Z}_{2'})$ in the $Z_{i'}$ -coordinate system to the ‘derivative’ basis $(\mathbf{Z}_1, \mathbf{Z}_2)$ in the Z_i -coordinate system once again by virtue of the chain rule:

$$\begin{pmatrix} \mathbf{Z}_{1'} \\ \mathbf{Z}_{2'} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{pmatrix}^T J_{i'}^i, \quad (4.5)$$

or equivalently, for each fixed i'

$$\mathbf{Z}_{i'} = \Sigma_i \mathbf{Z}_i J_{i'}^i, \quad (4.6)$$

which we contract in Einstein summation as

$$\mathbf{Z}_{i'} = \mathbf{Z}_i J_{i'}^i. \quad (4.7)$$

Remark 4.2. *The right hand side of (4.7), $J_{i'}^i$ is a $Z_{i'} \rightarrow Z_i$ transformation and the derivative basis vector \mathbf{Z}_i which lives in the Z_i -coordinate system space. This is in contrast to the right hand side of (3.9), Example 3.1, where $J_{i'}^i$ is a $Z_{i'} \rightarrow Z_i$ transformation acting on the derivative $\mathbf{V}'(t_0)$ which lives in the $Z_{i'}$ -coordinate system space. For this reason we shall refer to the Jacobian $J_{i'}^i$ in (4.7) as locally reverse-preserving coordinate basis vectors. Only in the special case of an orthonormal Jacobian will the operation of taking transposes ‘convert’ (3.9) to (4.7).*

Proposition 4.1. *The Jacobian locally reverse-preserves coordinate basis vectors.*

The common theme running behind Proposition 4.1 and Proposition 3.2 is that the Jacobian is the *correction factor* when the same algebraic formulation is executed blindly in two different coordinate systems. This is what the Einstein notation of equation (4.7) best expresses.

Objects in one coordinate system constructed by differentiation (algebraic formulation) differ from the corresponding object in another coordinate system created when that algebraic formulation is blindly followed exactly by a Jacobian which locally reverse-preserves.

These objects created algebraically are given the formal name *tensors*. So,

Tensors are (algebraically created) objects that can be tracked across any coordinate system by the Jacobian which locally reverse-preserves.

Proposition 4.2. *The Cartesian coordinate basis vectors \mathbf{i}, \mathbf{j} are tensors. So also are the polar coordinate basis vectors.*

Proof. By Remark 4.1, \mathbf{i}, \mathbf{j} can be treated as (constant) vector fields $\mathbf{Z}_1, \mathbf{Z}_2$ constructed algebraically by equations (4.1)-(4.2), respectively. Similarly, the polar coordinate basis vectors can be identified with (non-constant) vector fields $(\mathbf{Z}_{1'}, \mathbf{Z}_{2'})$ of (4.3)-(4.4) where $J_{i'}^i$ of (3.7) locally reverse preserves as in (4.5). \square

Remark 4.3. *From this point onwards we shall not make a distinction between a vector space basis such as (\mathbf{i}, \mathbf{j}) and the derivative basis such as $(\mathbf{Z}_1, \mathbf{Z}_2)$.*

So the algebraic formulation (2.1) of the gradient is a tensor at best: applying (4.5) to the coordinate dependent formulation (2.1),

$$\vec{\nabla} F(x', y') = \vec{\nabla} F(x, y)^T J_{i'}^i, \quad (4.8)$$

is precisely equation (2.2) since

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix}_{Z_{i'}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{Z_i}^T \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

The notion of a *tensor* is a formalization of equation (4.8). Formally,

Definition 4.1. *A tensor is an algebraically constructed object T_i in some coordinate system Z_i that is locally reverse preserved by the Jacobian $J_{i'}^i$ when that algorithm is implemented in the coordinate system $Z_{i'}$ to create the (dual) object $T_{i'}$:*

$$T_{i'} = T_i J_{i'}^i \quad (4.9)$$

In light of Remark 4.3, the algebraic formulation (2.1) of the gradient in tensor form becomes

$$\vec{\nabla} F(Z_i) := \begin{pmatrix} \frac{\partial F}{\partial Z_1} \\ \frac{\partial F}{\partial Z_2} \end{pmatrix}_{Z_i} \quad (4.10)$$

in the coordinate system Z_i .

Proposition 4.3. *The gradient defined as in (4.10) is a tensor.*

Proof. Set $T_i := \vec{\nabla} F(Z_i)$ and $T_{i'} := \vec{\nabla} F(Z_{i'})$. \square

The Jacobian is a positive step forward to the final goal of a coordinate-free universal algebraic formulation which would bring about a true marriage of geometry and algebra; geometric objects are also coordinate-free!

5. Another dualism map

Most functional analysts would not call the Jacobian a dualism map. The term *duality* has a specific meaning: given a basis $(\mathbf{e}_1, \mathbf{e}_2)$ of $V := \mathbb{R}^2$, where $\mathbf{e}_1 := \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $\mathbf{e}_2 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, the dual basis $(\mathbf{e}_1^*, \mathbf{e}_2^*)$ consists of linear functionals on V where for each fixed j ,

$$\mathbf{e}_i^*(\mathbf{e}_j) = \delta_j^i \quad (5.1)$$

for $j = 1, 2$. Each fixed $\mathbf{e}_j^* \in V^*$ can be represented by a fixed vector \mathbf{v}_j in V in the sense that

$$\mathbf{v}_i \bullet \mathbf{e}_j = \delta_j^i \quad (5.2)$$

for $j = 1, 2$. So the representation (5.2), allows us to identify V^* with V . For the purpose of distinguishing the basis from the dual basis, we use subscript and superscript notations: the dual basis of the basis $(\mathbf{e}_1^*, \mathbf{e}_2^*)$ is denoted by $(\mathbf{e}^1, \mathbf{e}^2)$ and the map

$$E : \mathbf{e}_i \mapsto \mathbf{e}^i \quad (5.3)$$

is called the *dualism* map. Functional analysts would not be up in arms to call E a dualism map. The Gram matrix formed by dotting the basis vectors

$$[E_{ij}] := \begin{pmatrix} \mathbf{e}_1 \bullet \mathbf{e}_1 & \mathbf{e}_1 \bullet \mathbf{e}_2 \\ \mathbf{e}_2 \bullet \mathbf{e}_1 & \mathbf{e}_2 \bullet \mathbf{e}_2 \end{pmatrix},$$

plays a critical role in computing the representation vectors \mathbf{v}_i of (5.2). We can identify its inverse E^{ij} with the dualism map E :

$$\mathbf{e}^i := \mathbf{v}_i = \Sigma_j E^{ij} \mathbf{e}_j \quad (5.4)$$

The Einstein notation once again contracts (5.4) in a more suggestive form:

$$\mathbf{e}^i = E^{ij} \mathbf{e}_j \quad (5.5)$$

The constructs (5.1)-(5.5), carry over to the derivative basis $(\mathbf{Z}_1, \mathbf{Z}_2)$ of tensor calculus (Section 4). Tensor calculus calls the derivative basis $(\mathbf{Z}_1, \mathbf{Z}_2)$ the *covariant* basis and the dual basis $(\mathbf{Z}^1, \mathbf{Z}^2)$ the *contravariant* basis. The matrix inverse Z^{ij} of the Gram matrix for the *covariant* basis plays the role of a dualism map $Z : \mathbf{Z}_i \mapsto \mathbf{Z}^i$:

$$\mathbf{Z}^i = Z^{ij} \mathbf{Z}_j \quad (5.6)$$

We shall use the notation Z and Z^{ij} interchangeably. Note that the dualism map Z^{ij} is local to coordinate system of choice:

Example 5.1. Consider the Cartesian and polar derivative (covariant) basis $(\mathbf{Z}_1, \mathbf{Z}_2)$ and $(\mathbf{Z}_{1'}, \mathbf{Z}_{2'})$ defined by equations (4.1)-(4.2) and (4.3)-(4.4), respectively. Then

1. The dual Cartesian derivative (contravariant) basis $(\mathbf{Z}^1, \mathbf{Z}^2)$ coincides with the covariant basis.
2. The dual polar (contravariant) basis $(\mathbf{Z}^{1'}, \mathbf{Z}^{2'})$ is $(\mathbf{Z}_{1'}, \frac{1}{r^2} \mathbf{Z}_{2'})$.

Proof. The Cartesian dualism map Z^{ij} is:

$$Z^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5.7)$$

The polar dualism map $Z^{i'j'}$ is:

$$Z^{i'j'} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{pmatrix}. \quad (5.8)$$

□

For the the final goal of a coordinate-free universal algebraic formulation of the gradient, the contravariant basis will prove essential. For this we introduce the notion of covariant and contravariant tensors of tensor calculus that we have already met unknowingly.

6. Coordinate-free algebraic formulations

The notion of a *covariant* tensor is precisely Definition 4.1. The term covariant is suitable since we use the covariant basis in (4.10). The notion of a *contravariant* tensor means transformation happens in the same direction as the Jacobian in the sense of Remark 4.2:

Definition 6.1. *A contravariant tensor is an algebraically constructed object T^i in some coordinate system Z_i that is locally preserved by the Jacobian $J_{i'}^i$ when that algorithm is implemented in the coordinate system $Z_{i'}$ to create the (dual) object $T^{i'}$:*

$$T^{i'} = T^i J_i^{i'} \quad (6.1)$$

Proposition 6.1. *The dual basis $(\mathbf{Z}^1, \mathbf{Z}^2)$ and $(\mathbf{Z}^{1'}, \mathbf{Z}^{2'})$ of Example 5.1 are contravariant tensors.*

The dualism map Z of Section 5 created a dual to the covariant tensor in the form of a contravariant tensor. Together they hold the key to a coordinate-free algebraic formulation:

Proposition 6.2. *The object created by the Einstein summation of a covariant and contravariant tensor*

$$U := S_i T^i$$

evaluates to the same value in all coordinate systems.

7. Fixed-up gradient

The problem with the algebraic formulation of the gradient $\vec{\nabla} F(Z_i)$ of (4.10) is that $\vec{\nabla} F(Z_i) = \left(\begin{array}{c} \frac{\partial F}{\partial Z_1} \\ \frac{\partial F}{\partial Z_2} \end{array} \right)_{Z_i}$ is only a covariant tensor (Proposition 4.3). So taking the Einstein summation (dot product) with the contravariant tensor $(\mathbf{Z}^1, \mathbf{Z}^2)$ (Proposition 6.1) we end up with the coordinate-free algebraic formulation of the gradient

$$\vec{\nabla} F(Z_i) = \sum_i \frac{\partial F}{\partial Z_i} \mathbf{Z}^i. \quad (7.1)$$

by Proposition 6.2.

Remark 7.1. *The flawed algebraic formulation (2.1) is a special case of the universal formulation (7.1) since the contravariant basis and the covariant basis coincides with Cartesian coordinates (Example 5.1, item 1).*

We end off by recommending Chapters 5, 6 of [4] as good follow up reading.

Acknowledgement. This work was gratefully funded by the Claude Leon Foundation.

References

- 1 GRINFELD, P, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*, Springer Science+Business Media New York, 2013.
- 2 FADELL, A.G, *Vector Calculus and Differential Equations*, D. Van Nostrand, United States, 1968.
- 3 DINEEN, S, *Multivariate Calculus and Geometry*, Third Edition, Springer-Verlag London, 2014.
- 4 GARRITY, T.A, *All the Mathematics You Missed But Need to Know for Graduate School*, Cambridge University Press, 2002.