1 Motivation

The goal is to develop "proofs" or, short of that, sketches—informal demonstrations—of the truth of important results which are more "instructive" than those that have made their way into textbooks. I believe there is a certain loss of information that happens through the ages and, in microcosm, between the inception and the proof of a result. I am not happy with the presentation of, say, so-called Linear Algebra or Multivariable Calculus in the textbooks today. Perhaps these textbooks have a particular rhetorical end that is served by their organization and, perhaps, obscurity of their presentation. Additionally, particular proofs may sometimes be chosen because they are short. (Perhaps there is a conflation by mathematicians between brevity and elegance! While "elegance" is trotted out as one of the motivating aesthetical considerations of mathematics, there is a danger that it be understood simply as a synonym for short, mean and lean, or even "economical" as opposed to something truly transcendent and of high value.) But perhaps their would be a greater cognitive (etc.) benefit to another presesntation.

2 Gauss' Theorem

Let S be an arbitrary tetrahedron. Actually, we will soon imagaine that this tetrahedron is very small. As you may know, a tetrahedron has 4 triangular faces. Since every edge is shared by two faces, this makes for

$$\frac{[\text{number of faces}][\text{number of edges per face}]}{2} = \frac{4 \cdot 3}{2} = 6$$

edges.

S is totally determined by as few as three of these edges—at least by three edges that share a common vertex. Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be vectors parallel to three edges that share a common vertex. Identify that vertex as $\mathbf{x}_0$. Note that the other three edges can be identified with (by being parallel to)

$$\mathbf{w} - \mathbf{u}, \mathbf{v} - \mathbf{u}, \mathbf{w} - \mathbf{v}$$

With an eye towards generalization of this result, you can remark that these remaining three vectors are what you get by choosing the first three in pairs; after all, $\binom{3}{2} = 3$.

Now, we describe the four faces. First, let's fix a system of orienting ourselves as we imagine looking at the tetrahedron S. Imagine that you, the observer, are located at the point $\mathbf{x}_0$. Label any one of the vectors extending from this point as $\mathbf{u}$. Then, proceeding counterclockwise—at least from the perspective of this point of observation—we choose to label the next vector v, followed by $\mathbf{w}$. Continuing to proceed counterclockwise should return you to vector $\mathbf{u}$. Next, imagine an arbitrary vector pointing outward from a given face (that is to say, so that if it were anchored at an appropriate point on the interior of S, it would pierce through the face on its way to the outside). For instance, imagine the face, which we will call S_1, subtended by $\mathbf{u}$ and $\mathbf{v}$ [Note sure about that yet...]

$$S_1 = \left\{ \mathbf{x}_0 + s\mathbf{v} + t\mathbf{u} \;\middle|\; \begin{matrix} 0 \leq s, 0 \leq t \\ s + t \leq 1 \end{matrix} \right\}$$

$$S_2 = \left\{ \mathbf{x}_0 + s\mathbf{w} + t\mathbf{v} \;\middle|\; \begin{matrix} 0 \leq s, 0 \leq t \\ s + t \leq 1 \end{matrix} \right\}$$

$$S_3 = \left\{ \mathbf{x}_0 + s\mathbf{u} + t\mathbf{w} \;\middle|\; \begin{matrix} 0 \leq s, 0 \leq t \\ s + t \leq 1 \end{matrix} \right\}$$

$$S_4 = \left\{ \mathbf{x}_0 + s(\mathbf{v} - \mathbf{u}) + t(\mathbf{w} - \mathbf{u}) \;\middle|\; \begin{matrix} 0 \leq s, 0 \leq t \\ s + t \leq 1 \end{matrix} \right\}$$

2.1 Flux

What we aim to do is to provide a (somewhat informal) demonstration of the truth of Gauss' Theorem—one that promises (in the author's mind) to link up with a new, more vibrant way of organizing, presenting, and narrating the results of so-called Linear Algebra and Multivariable Calculus. Gauss' Theorem, also called the Divergence Theorem, concerns what is called *flux* across a surface. Just what that is from an intutive point of view is something I hope to better understand through this particular exercise and of the project, in general, or reorganizing the subject matter. But, for now, let's proceed on the hunch that the study of flux ultimately concerns looking at the effects on a surrounding surface of a source of a force—with a particular geometry—centered on a particular point. The geometry of this force field is the interesting thing, with it having different effects depending on the location of the effect relative to the center or source. Authors discussing this subject often invite their readers to imagine a fluid passing through a particular element of surface area. The flux, across that particular element, if the amount of fluid that passes through the surface in a unit of time. As such, one can imagine the volume of a parallelepiped with its base the element of surface area and extending in the direction of the flow $\mathbf{F}$ of the fluid (so, in this case, $\mathbf{F}$ would actually be a velocity vector). Thus, the element of flux $d\Phi$ of $\mathbf{F}$ across element of surface area dS can be obtained by computing

$$d\Phi = \det\left(\mathbf{F}, \mathbf{a}\, ds, \mathbf{b}\, dt\right) = \det\left(\mathbf{F}, \mathbf{a}, \mathbf{b}\right) ds dt$$

where $\mathbf{a}$ and $\mathbf{b}$ are the vectors subtending this element of surface area (that is, they are parallel to the sides of the parallelogram representing dS) and ds and dt are scalars (which are meant to be small or even "infinitesimal" increments, of course). Also, you may appreciate seeing what I mean by the above notation:

$$\det(\mathbf{a},\mathbf{b},\mathbf{c}) := \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

The total flux over S is just the sum of the fluxes over each face. On the other hand the vector field field $\mathbf{F}$ is expected to vary across a given face. We will repeatedly use the approximation

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + D\mathbf{F}(\mathbf{x}_0)\Delta\mathbf{x},$$

where

$$D\mathbf{F}(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial y} & \frac{\partial F_3}{\partial z} \end{bmatrix}$$

Note that each of the partial derivatives are evaluated at $\mathbf{x} = \mathbf{x}_0$. Actually, for even shorter, let's label

$$A = D\mathbf{F}(\mathbf{x}_0)$$

Applying this approximation and using the continuity and then the multilinearity of det, we get for $\mathbf{x}$ near $\mathbf{x}_0$,

$$\begin{aligned} \det(\mathbf{F}(\mathbf{x}),\mathbf{a},\mathbf{b}) &\approx \det\big(\mathbf{F}(\mathbf{x}_0) + A\Delta\mathbf{x},\mathbf{a},\mathbf{b}\big) \\ &= \det\big(\mathbf{F}(\mathbf{x}_0),\mathbf{a},\mathbf{b}\big) + \det\big(A\Delta\mathbf{x},\mathbf{a},\mathbf{b}\big) \end{aligned}$$

Now, note that for faces, $S_i, i = 1,\ldots,3$, we have that any $\mathbf{x}$ on S_i is of the form

$$\mathbf{x} = \mathbf{x}_0 + s\mathbf{a} + t\mathbf{b}$$

In the case of S_4, $\mathbf{x}$ is just a little different in form:

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{u} + s\mathbf{a} + t\mathbf{b}$$

For each $i = 1,\ldots,3$, we also have that $\Delta\mathbf{x}$ is of the form

$$\Delta\mathbf{x} = \Delta\mathbf{x}(s,t) = \mathbf{x}(s,t) - \mathbf{x}_0 = s\mathbf{a} + t\mathbf{b}$$

For S_4, we have

$$\Delta\mathbf{x} = \Delta\mathbf{x}(s,t) = \mathbf{x}(s,t) - \mathbf{x}_0 = \mathbf{u} + s\mathbf{a} + t\mathbf{b}$$

In particular, for $i = 1,\ldots,3$,

$$\det\big(A\Delta\mathbf{x},\mathbf{a},\mathbf{b}\big) = s\det\big(A\mathbf{a},\mathbf{a},\mathbf{b}\big) + t\det\big(A\mathbf{b},\mathbf{a},\mathbf{b}\big)$$

while, for S_4,

$$\det\big(A\Delta\mathbf{x},\mathbf{a},\mathbf{b}\big) = \det\big(A\mathbf{u},\mathbf{a},\mathbf{b}\big) + s\det\big(A\mathbf{a},\mathbf{a},\mathbf{b}\big) + t\det\big(A\mathbf{b},\mathbf{a},\mathbf{b}\big)$$

Thus, we get the following information:

S_1: $\Delta\mathbf{x}(s,t) = s\mathbf{v} + t\mathbf{u}$

$$\det\left(\mathbf{F}(\mathbf{x}),\mathbf{a}_1,\mathbf{b}_1\right) \approx \det\big(\mathbf{F}(\mathbf{x}_0),\mathbf{v},\mathbf{u}\big) + s\det\big(A\mathbf{v},\mathbf{v},\mathbf{u}\big) + t\det\big(A\mathbf{u},\mathbf{v},\mathbf{u}\big)$$

S_2: $\Delta\mathbf{x}(s,t) = s\mathbf{w} + t\mathbf{v}$

$$\det\left(\mathbf{F}(\mathbf{x}),\mathbf{a}_2,\mathbf{b}_2\right) \approx \det\big(\mathbf{F}(\mathbf{x}_0),\mathbf{w},\mathbf{v}\big) + s\det\big(A\mathbf{w},\mathbf{w},\mathbf{v}\big) + t\det\big(A\mathbf{v},\mathbf{w},\mathbf{v}\big)$$

S_3: $\Delta\mathbf{x}(s,t) = s\mathbf{u} + t\mathbf{w}$

$$\det\left(\mathbf{F}(\mathbf{x}),\mathbf{a}_3,\mathbf{b}_3\right) \approx \det\big(\mathbf{F}(\mathbf{x}_0),\mathbf{u},\mathbf{w}\big) + s\det\big(A\mathbf{u},\mathbf{u},\mathbf{w}\big) + t\det\big(A\mathbf{w},\mathbf{u},\mathbf{w}\big)$$

S_4: $\Delta\mathbf{x}(s,t) = \mathbf{u} + s(\mathbf{v}-\mathbf{u}) + t(\mathbf{w}-\mathbf{u})$

$$\begin{aligned}\det\left(\mathbf{F}(\mathbf{x}),\mathbf{a}_4,\mathbf{b}_4\right) \approx \det\big(\mathbf{F}(\mathbf{x}_0),\mathbf{v}-\mathbf{u},\mathbf{w}-\mathbf{u}\big) &+ \det(A\mathbf{u},\mathbf{v}-\mathbf{u},\mathbf{w}-\mathbf{u}) \\ &+ s\det\big(A(\mathbf{v}-\mathbf{u}),\mathbf{v}-\mathbf{u},\mathbf{w}-\mathbf{u}\big) \\ &+ t\det\big(A(\mathbf{w}-\mathbf{u}),\mathbf{v}-\mathbf{u},\mathbf{w}-\mathbf{u}\big)\end{aligned}$$

In preparation of adding the fluxes across each face, we note that the sum of the "constant" terms, excluding the term $\det(A\mathbf{u},\mathbf{v}-\mathbf{u},\mathbf{w}-\mathbf{u})$: is 0:

$$\begin{aligned}\text{Sum of constant terms} &= \det(\mathbf{y}_0,\mathbf{v},\mathbf{u}) \\ &+ \det(\mathbf{y}_0,\mathbf{w},\mathbf{v}) \\ &+ \det(\mathbf{y}_0,\mathbf{u},\mathbf{w}) \\ &+ \det(\mathbf{y}_0,\mathbf{v},\mathbf{w}) \\ &- \det(\mathbf{y}_0,\mathbf{v},\mathbf{u}) \\ &- \det(\mathbf{y}_0,\mathbf{u},\mathbf{w}) \\ &= 0\end{aligned}$$

Note that I have abbreviated $\mathbf{y}_0 := \mathbf{F}(\mathbf{x}_0)$. Also, in case you are wondering, I obtained the last three terms using the multilinearity of det. Along the way, I also neglected the term $\det(\mathbf{y}_0,\mathbf{u},\mathbf{u})$ as it equals 0, using the fact that a determinant with a repeated column (or row) is 0. And, in order to accomplish the sum, one needs also the following property of det:

$$\det(\mathbf{a},\mathbf{b},\mathbf{c}) = -\det(\mathbf{b},\mathbf{a},\mathbf{c}) = -\det(\mathbf{a},\mathbf{c},\mathbf{b})$$

Now, we take the sum of the fluxes across each face:

$$
\begin{aligned}
\text{Total Flux} &= \int\int_S d\Phi \\
&= \sum_{i=1}^{4} \int\int_{S_i} d\Phi \\
&= \sum_{i=1}^{4} \int\int_{S_i} \det(\mathbf{F}((x)), \mathbf{a}_i, \mathbf{b}_i) dS \\
&= \sum_{i=1}^{4} \int_0^1 \int_0^{1-s} \det(\mathbf{F}((x(s,t))), \mathbf{a}_i, \mathbf{b}_i) dt ds \\
&= \int_0^1 \int_0^{1-s} \sum_{i=1}^{4} \det(\mathbf{F}((x(s,t))), \mathbf{a}_i, \mathbf{b}_i) dt ds \\
&\approx \int_0^1 \int_0^{1-s} [\text{sum of } s\text{-terms}]s + [\text{sum of } t\text{-terms}]t + \det(A\mathbf{u}, \mathbf{v}-\mathbf{u}, \mathbf{w}-\mathbf{u})\, dt ds
\end{aligned}
$$

Now, noting that no terms in these sums depend on s or t, let's write $P = [\text{sum of } s\text{-terms}]$ and $Q = [\text{sum of } t\text{-terms}]$. While we are at it, let's write $R = \det(A\mathbf{u}, \mathbf{v}-\mathbf{u}, \mathbf{w}-\mathbf{u})$. Thus, we compute:

$$
\begin{aligned}
\int_0^1 \int_0^{1-s} Ps + Qt + R\, dt ds &= \int_0^1 \left[Pst + \frac{1}{2}Qt^2 + Rt \Big|_0^{1-s} \right] ds \\
&= \int_0^1 P(1-s)s + \frac{1}{2}Q(1-s)^2 + R(1-s)\, ds \\
&\vdots \\
&= \frac{1}{6}(P+Q) + \frac{1}{2}R
\end{aligned}
$$

Now, let's examine these terms:

$$
\begin{array}{ccc}
P & Q & R \\
\det(A\mathbf{v},\mathbf{v},\mathbf{u}) & \det(A\mathbf{u},\mathbf{v},\mathbf{u}) & \det(A\mathbf{u},\mathbf{v},\mathbf{w}) \\
+\det(A\mathbf{w},\mathbf{w},\mathbf{v}) & +\det(A\mathbf{v},\mathbf{w},\mathbf{v}) & -\det(A\mathbf{u},\mathbf{v},\mathbf{u}) \\
+\det(A\mathbf{u},\mathbf{u},\mathbf{w}) & +\det(A\mathbf{w},\mathbf{u},\mathbf{w}) & -\det(A\mathbf{u},\mathbf{u},\mathbf{w}) \\
+\det(A\mathbf{v},\mathbf{v},\mathbf{w}) & +\det(A\mathbf{w},\mathbf{v},\mathbf{w}) & +\det(A\mathbf{u},\mathbf{u},\mathbf{u}) \\
-\det(A\mathbf{v},\mathbf{v},\mathbf{u}) & -\det(A\mathbf{w},\mathbf{v},\mathbf{u}) & \\
-\det(A\mathbf{v},\mathbf{u},\mathbf{w}) & -\det(A\mathbf{w},\mathbf{u},\mathbf{w}) & \\
+\det(A\mathbf{v},\mathbf{u},\mathbf{u}) & +\det(A\mathbf{w},\mathbf{u},\mathbf{u}) & \\
-\det(A\mathbf{u},\mathbf{v},\mathbf{w}) & -\det(A\mathbf{u},\mathbf{v},\mathbf{w}) & \\
+\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & +\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & \\
+\det(A\mathbf{u},\mathbf{u},\mathbf{w}) & +\det(A\mathbf{u},\mathbf{u},\mathbf{w}) & \\
-\det(A\mathbf{u},\mathbf{u},\mathbf{u}) & -\det(A\mathbf{u},\mathbf{u},\mathbf{u}) &
\end{array}
$$

Note that determinants with repeated columns are 0. There are also two terms that cancel immediately in the first column. Applying these, we get the slightly simplified terms:

$$
\begin{array}{ccc}
P & Q & R \\
\det(A\mathbf{w},\mathbf{w},\mathbf{v}) & \det(A\mathbf{u},\mathbf{v},\mathbf{u}) & \det(A\mathbf{u},\mathbf{v},\mathbf{w}) \\
+\det(A\mathbf{u},\mathbf{u},\mathbf{w}) & +\det(A\mathbf{v},\mathbf{w},\mathbf{v}) & -\det(A\mathbf{u},\mathbf{v},\mathbf{u}) \\
+\det(A\mathbf{v},\mathbf{v},\mathbf{w}) & +\det(A\mathbf{w},\mathbf{u},\mathbf{w}) & -\det(A\mathbf{u},\mathbf{u},\mathbf{w}) \\
-\det(A\mathbf{v},\mathbf{u},\mathbf{w}) & +\det(A\mathbf{w},\mathbf{v},\mathbf{w}) & +\det(A\mathbf{u},\mathbf{u},\mathbf{u}) \\
-\det(A\mathbf{u},\mathbf{v},\mathbf{w}) & -\det(A\mathbf{w},\mathbf{v},\mathbf{u}) & \\
+\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & -\det(A\mathbf{w},\mathbf{u},\mathbf{w}) & \\
+\det(A\mathbf{u},\mathbf{u},\mathbf{w}) & -\det(A\mathbf{u},\mathbf{v},\mathbf{w}) & \\
 & +\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & \\
 & +\det(A\mathbf{u},\mathbf{u},\mathbf{w}) &
\end{array}
$$

Transposing and applying the coefficients, the terms in the final sum $\frac{1}{6}(P+Q)+\frac{1}{2}R$ are as follows:

$$
\begin{array}{rccc}
 & Au & Av & Aw \\
 & \frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & \frac{1}{6}\det(A\mathbf{v},\mathbf{v},\mathbf{w}) & \frac{1}{6}\det(A\mathbf{w},\mathbf{w},\mathbf{v}) \\
 & +\frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & -\frac{1}{6}\det(A\mathbf{v},\mathbf{u},\mathbf{w}) & +\frac{1}{6}\det(A\mathbf{w},\mathbf{u},\mathbf{w}) \\
 & -\frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & +\frac{1}{6}\det(A\mathbf{v},\mathbf{w},\mathbf{v}) & +\frac{1}{6}\det(A\mathbf{w},\mathbf{v},\mathbf{w}) \\
 & +\frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & & -\frac{1}{6}\det(A\mathbf{w},\mathbf{v},\mathbf{u}) \\
 & -\frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & & -\frac{1}{6}\det(A\mathbf{w},\mathbf{u},\mathbf{w}) \\
 & +\frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & & \\
 & +\frac{1}{2}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & & \\
 & -\frac{1}{2}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & & \\
 & -\frac{1}{2}\det(A\mathbf{u},\mathbf{v},\mathbf{u}) & & \\
\hline
\text{Sum:} & \frac{1}{6}\det(A\mathbf{u},\mathbf{v},\mathbf{w}) & -\frac{1}{6}\det(A\mathbf{v},\mathbf{u},\mathbf{w}) & -\frac{1}{6}\det(A\mathbf{w},\mathbf{v},\mathbf{u})
\end{array}
$$

Permuting the arguments of the determinant (note that permuting arguments of a determinant comes at the consequence of reversing the sign), we have that

$$[\text{Total flux}] \approx \frac{1}{6}\Big(\det(A\mathbf{u},\mathbf{v},\mathbf{w}) + \det(\mathbf{u},A\mathbf{v},\mathbf{w}) + \det(\mathbf{u},\mathbf{v},A\mathbf{w})\Big) \tag{1}$$

Of course, the accuracy of this approximation depends upon the tetrahedron S being rather small; on the other hand, this approximation gets even more accurate the smaller S is.

Now, do you see Gauss' Theorem yet? No, perhaps not, but let's take a look at the terms $A\mathbf{u}, A\mathbf{v}, A\mathbf{w}$. Note that $\mathbf{u},\mathbf{v},\mathbf{w}$ together form a tetrahedron. In particular, they are not coplanar and thus span $\mathbb{R}^3$. This in turn implies that we can write each of $A\mathbf{u}, A\mathbf{v}, A\mathbf{w}$ in terms of $\mathbf{u},\mathbf{v},\mathbf{w}$. Knowing that a determinant with repeated columns is 0 (and using also its multilinearity), this provides an avenue for simplifying the expression for total flux above.

Write

$$
\begin{aligned}
A\mathbf{u} &= c_{11}\mathbf{u} + c_{21}\mathbf{v} + c_{31}\mathbf{w} \\
A\mathbf{v} &= c_{12}\mathbf{u} + c_{22}\mathbf{v} + c_{32}\mathbf{w} \\
A\mathbf{w} &= c_{13}\mathbf{u} + c_{23}\mathbf{v} + c_{33}\mathbf{w}
\end{aligned}
$$

Thus, our approximation for total flux over S can be rewritten:

$$[\text{Total flux}] \approx (c_{11} + c_{22} + c_{33})\,\frac{1}{6}\det(\mathbf{u},\mathbf{v},\mathbf{w}) \tag{2}$$

This equation has a nice interpretation! Note that $\frac{1}{6}\det(\mathbf{u},\mathbf{v},\mathbf{w})$ is the volume of tetrahedron S. Since

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

is the matrix representation of (the linear transformation represented by) A using the basis $\mathcal{U} = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$, one can regard expression $c_{11} + c_{22} + c_{33}$ as the divergence of $\mathbf{F}$ relative to the basis $\mathcal{U}$.

Moving in another direction, we can also use the expression (1) to prove Gauss' Theorem (modulo some informality). Consider the case of

$$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \{\Delta x \mathbf{e}_1, \Delta y \mathbf{e}_2, \Delta z \mathbf{e}_3\}$$

where $\Delta x, \Delta y, \Delta z$ are some small increments—which can be negative!—and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the usual basis vectors. Notice that I am leaving it flexible as to which vector on the left side gets mapped to which vector on the right. Why am I doing that? Well, what I have actually been interested in from the start with this arbitrary tetrahedron is because I am interested in the "base case", if you like, of the divergence theorem in $\mathcal{R}^3$. That is, my aim is now to note that the types of surfaces that we probably want to apply Gauss' Theorem to are those than can be approximated by polygons. Imagine some sort of nice surface like that. Call it S, if you can stand it. Fill the space $\mathcal{R}^3$ with a fine, uniform lattice of points. Send out "feelers" in the 6 "cardinal" directions in $\mathcal{R}^3$ form each lattice point. You will find some lattice points whose feelers intersect S at a distance of less than the uniform distance between adjacent (rectilinearly) lattice points. These feelers will be your $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$; however, I have to leave it up how the labels go so that $\det(\mathbf{u}, \mathbf{v}, \mathbf{w})$ will be a (positive) volume rather than a negative one. Accepting that we have settled on the correct labeling (or "orientation" if you prefer), we can rewrite:

$$\begin{aligned} \frac{1}{6}\det(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \frac{1}{6}\det(\Delta x_i \mathbf{e}_i, \Delta x_j, \mathbf{e}_j, \Delta x_k, \mathbf{e}_k) \\ &= \frac{1}{6}\Delta x_i \Delta x_j \Delta x_k \det(\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k) \\ &= \frac{1}{6}\Delta x \Delta y \Delta z, \end{aligned}$$

where the indices are just a way of anonymizing the vectors. (In other words, the indices i, j, k are *not* meant to correspond to the basis vectors $\mathbf{i} = \mathbf{e}_1, \mathbf{j} = \mathbf{e}_2, \mathbf{k} = \mathbf{e}_3$ but are just supposed to be neutral and arbitrary.) Note that the sign of the Δ-increments are allowed to be negative and that the sign of this expression is positive by the choice of the ordering $\langle \mathbf{u}, \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k \rangle$. In other words, in case it still isn't clear, the mapping from $\{i, j, k\}$ to $\{1, 2, 3\}$ is a permutation chosen to make this calculation come out positive.

Finally, note that, for $i = 1, 2, 3,$

$$\begin{aligned}
A\boldsymbol{\Delta} x_i e_i &= \Delta x_i A\mathbf{e}_i \\
&= \Delta x_i \left(a_{1i}\mathbf{e}_1 + a_{2i}\mathbf{e}_2 + a_{3i}\mathbf{e}_3\right) \\
&= a_{1i}\Delta x_i \mathbf{e}_1 + a_{2i}\Delta x_i \mathbf{e}_2 + a_{3i}\Delta x_i \mathbf{e}_3
\end{aligned}$$

from which we may conclude that $c_{ii} = a_{ii}$. Recalling the definition of A we may conclude that the divergence theorem (at least approximately) holds over this tetrahedron. That is, calling this tetrahedron T_i, we have that equation (2), we have

$$\begin{aligned}
[\text{Total flux over } T_i] &\approx (c_{11} + c_{22} + c_{33})\,\frac{1}{6}\det(\mathbf{u}, \mathbf{v}, \mathbf{w}) \\
&= (a_{11} + a_{22} + a_{33})\,\frac{1}{6}\Delta x \Delta y \Delta z \\
&= \left(\frac{\partial \mathbf{F}_1}{\partial x} + \frac{\partial \mathbf{F}_2}{\partial y} + \frac{\partial \mathbf{F}_3}{\partial z}\right)\Delta V_i
\end{aligned}$$

where $\Delta V_i = \frac{1}{6}\Delta x \Delta y \Delta z$, the volume of tetrahedron T_i. To extend the divergence theorem to surface $S = \partial D$, we note that both the volume of D and the surface S are well-approximated by our scheme of choosing tetrahedrons. (We have paid special attention to the surface, but one can, of course, continue into the interior of D filling it completely with these tetrahedrons.) Finally, in the "usual trick," we should note that the flux, $\det(\mathbf{F}, \mathbf{a}, \mathbf{b})$ over a given face of a tetrahedron will instead be negative when computed from the perspective of the adjoining tetrahedron (sharing that face). In point of fact, the flux over this face will be computed in the adjacent tetrahedron as $\det(\mathbf{F}, \mathbf{b}, \mathbf{a})$, as this reverse choice must be made to register "inflow" as a negative area and "outflow" as positive.