

Australian Government Department of Defence Science and Technology

On the Use of Vectors, Reference Frames, and Coordinate Systems in Aerospace Analysis

Don Koks

Cyber and Electronic Warfare Division Defence Science and Technology Group

DST-Group-TR-3309

ABSTRACT

This report describes the core foundational concepts of aerospace modelling, an understanding of which is necessary for the analysis of complex environments that hold many entities, each of which might employ a separate reference frame and coordinate system. I begin by defining vectors, frames, and coordinate systems, and then discuss the quantities that allow Newton's laws to be applied to a complex scenario. In particular, I explain the crucial distinction between the "coordinates of the time-derivative of a vector" and the "time-derivative of the coordinates of a vector". I finish by drawing a parallel between this aerospace language and the notation found in seemingly unrelated areas such as relativity theory and fluid dynamics, and make some comments on various supposedly different derivatives as found in the literature.

RELEASE LIMITATION

Approved for Public Release

Published by Cyber and Electronic Warfare Division Defence Science and Technology Group PO Box 1500 Edinburgh, SA 5111, Australia Telephone: 1300 333 362 Facsimile: (08) 7389 6567 © Commonwealth of Australia 2017 AR-016-726 January 2017

APPROVED FOR PUBLIC RELEASE

On the Use of Vectors, Reference Frames, and Coordinate Systems in Aerospace Analysis

Executive Summary

Classical concepts of kinematics are by now well established. But the many differences of opinion on the subject easily found on various internet physics and maths discussion sites indicate that despite being well established, these concepts are not necessarily well understood. I believe that this confusion stems from standard textbook presentations on the subject, whose applicability is limited to only very simple scenarios. In the complex environments encountered in aerospace, a far more advanced understanding of the relevant concepts and their notation is pivotal to the success of any analysis.

This report discusses the more advanced (but still standard) definitions of vectors, reference frames, and coordinate systems that allow their use to model the most complex aerospace scenarios. I begin by defining vectors, highlighting the crucial distinction between a *proper vector* (an arrow) and a *coordinate vector* (an array of numbers). Next I define a reference frame as a quasi-physical scaffold relative to which all motion is defined. I then define a coordinate system essentially as a set of rulers attached to a frame, but with the proviso that one's choice of frame need not be tied to one's choice of coordinates; that is, we are free to quantify the events in our chosen frame by using the coordinates natural to another frame.

Having defined the basic concepts, I discuss the quantities that allow Newton's laws to be applied to a complex system. In particular, I explain the important difference between the "coordinates of the time-derivative of a vector" and the "time-derivative of the coordinates of a vector", which is central to aerospace calculations. This difference is not widely appreciated in the field, nor indeed in physics more generally, where it is generally taught only in advanced courses in relativity—but where its meaning is easily lost in a forest of notation. And yet this difference is a basic part of vector analysis that could easily be taught at a first-year university level.

To unify the discussions of this report with the bigger picture, I finish by drawing a parallel between this aerospace language and the tensor notation common in relativity theory. One theme of this report is that some apparently different types of derivative found in the literature are, in essence, identical: although often viewed as "new" or special, they are in fact nothing more than standard derivatives written in a way that is meant to aid practitioners in the various fields that use them.

This page is intentionally blank.

Author



Don Koks

Cyber & Electronic Warfare Division

Don Koks completed a doctorate in mathematical physics at Adelaide University in 1996, with a thesis describing the use of quantum statistical methods to analyse decoherence, entropy and thermal radiance in both the early universe and black hole theory. He holds a Bachelor of Science from the University of Auckland in pure and applied mathematics and physics, and a Master of Science in physics from the same university with a thesis in applied accelerator physics (proton-induced X ray and γ ray emission for trace element analysis). He has worked on the accelerator mass spectrometry programme at the Australian National University in Canberra, and in commercial Internet development.

Currently he is a Research Scientist with the Position Navigation and Timing Technologies and Systems Group in the Cyber and Electronic Warfare Division at the Defence Science and Technology Group, specialising in the relativity aspects of precision timing, satellite orbital prediction, 6-degree-of-freedom orientation concepts, and radar signal processing. He is the author of the book *Explorations in Mathematical Physics: the Concepts Behind an Elegant Language* (Springer, 2006).

This page is intentionally blank.

Contents

1	Introduction								
2	The First Concept: Proper Vectors								
3	The Second Concept: Reference Frames								
4	The Third Concept: Coordinate Systems								
5	The Fourth Concept: Coordinate Vectors								
6	Basis Vectors								
	6.1 Example: Basis Vectors for Polar Coordinates	10							
7	Relating Different Coordinate Systems								
	7.1 Relating Vector Components	12							
	7.2 Relating Rank-2 Tensor Components	13							
	7.3 Interpretations of "Position"	15							
8	The Velocity Vector	16							
9	The Angular Velocity Vector	18							
	9.1 Angular Velocity of a Point Relative to Another Point in Some Frame	18							
	9.2 Angular Velocity of a Frame Relative to Another Frame	25							
10	Coordinates and the Time Derivative	26							
	10.1Three Forms of the Time Derivative of a Vector	27 28 29 29							
11	The Acceleration Vector	33							
12	Inertial Frames								
13	Torque and Angular Momentum								
14	4 Final Comments								
15	References	36							
Appendix A: Relating Vector Components for Non-Orthonormal Coordinates 39									
	A.1 Example: Calculating ∇ in Spherical Polar Coordinates	41							
	A.2 Comment on the Vector Components v_{α}	42							
UNCLASSIFIED									

Appendi	x B: Time Derivative of a Rank-2 Tensor	45
Appendi	x C: Two Derivatives found in Tensor Analysis and Fluid Flow	47
C.1	The Covariant Derivative	47
C.2	A Derivative Used in Fluid Flow	48

1 Introduction

Six-degree-of-freedom modelling is a classic example of a calculation in aerospace theory that can involve multiple reference frames and coordinate systems. One might think that the concepts of how to relate such frames and coordinate systems are well established, taught to all physics/maths/engineering undergraduates, and with nothing new left to say. But a study of non-specialist discussions and differences of opinion on the subject on various internet physics and maths discussion sites [1] reveals a large variance in their contributors' declared understanding of the theory of vectors, reference frames, and coordinate systems, the basic tools that describe position and motion. I believe that this general confusion begins early on, with standard first-year university textbooks that were necessarily written for the simplest of systems, and which understandably use abbreviated notation that is fully appropriate to the tasks they address. These first-principles ideas and notation are simply inadequate to handle more advanced systems such as the complex dynamics of aerospace. There, some concepts and notation must be redesigned to handle environments involving multiple entities each with their own reference frame and coordinate system. In particular, when calculating velocity, the crucial distinction between the "coordinates of the time-derivative of a vector" and the "time-derivative of the coordinates of a vector" is generally not given the attention it deserves, and tends to be absent from many discussions. These two quantities are unequal for all but the simplest scenarios.

The objects described in kinematic scenarios range from a pendulum or a train confined to a rail, to an aircraft flying over a curved Earth while being tracked by a spinning satellite. Simplistic definitions of vectors, reference frames, and coordinate systems are usually quite adequate for non-demanding scenarios, but they tend to fail when used to describe the complex environments that are of real interest in aerospace.

This report aims to highlight and address such difficulties by drawing attention to more sophisticated definitions and notation for the subject. I begin by defining vectors, frames, and coordinate systems, and then discuss the ideas that allow Newton's laws to be applied to a complex system. Newton's laws are of course fundamental to the field, but I include them in this report to highlight the full attention which must be given to the various frames and reference points that are necessary to make full sense of these laws in complex environments.

The notation used in the following pages is similar to that used by Zipfel in his book *Modeling and Simulation of Aerospace Vehicle Dynamics* [2]. What might at first seem like notational clutter, both in this report and in Zipfel's book, quickly becomes useful when a plethora of vectors, frames, and coordinate systems are all playing concurrent roles in a scenario being analysed. The simple notation of an undergraduate physics textbook is very useful for simple scenarios, but it fails to cope with anything more advanced.

A good starting point is the notion of a continuous space populated with objects whose behaviour we wish to analyse. Discussions of reference frames can easily become esoteric such as with Mach's ideas of inertia—and so we might not hope to define everything in an absolute sense; but neither should our definitions become an infinite chain of turtles standing on the backs of turtles: we must stop somewhere and rely on usage to dictate the meaning of the core ideas. The concepts described in the next few sections are introduced in an ordered way, but this does not rule out their being re-ordered with a different description, depending on which is treated as constructible from the others.

DST-Group-TR-3309

2 The First Concept: Proper Vectors

Two quite distinct objects are routinely called a "vector" [3]. The first is a *proper vector*, which is a geometrical quantity, drawn very usefully as an arrow, and envisaged without any possibility of confusion as a real physical object, such as a straight piece of wood with an arrowhead at one end. This concrete view makes it clear that, like a point, a proper vector can be thought of as existing in an absolute sense: it can be "seen" by all. It has no numbers explicitly associated with it. The second type of object commonly called a vector is a *coordinate vector*, an ordered set of numbers discussed in Section 5. We will write a given proper vector in boldface as \boldsymbol{v} , with length \boldsymbol{v} . Note that \boldsymbol{v} is an arrow, a geometrical quantity with a kind of absolute existence, with no numerals associated with it.

The proper vector is useful because it can be treated as a building block of the world that we wish to analyse. Given fundamental ideas of distance and angle, we can use vectors to construct a complex object and hence describe it.

We must be careful to distinguish a proper vector from a point. A proper vector requires two points to define it: a head and a tail, and it follows that a single point is *not* a proper vector. The position of a point A can only be quantified by describing its displacement relative to a given point B. This *displacement vector* is the fundamental proper vector: it can be modelled as an arrow, with tail at point B and head at A. We will write this displacement vector of A relative to B as \mathbf{r}_{AB} . The usual "triangle law of addition" for vectors (arrows) says that

$$\boldsymbol{r}_{AB} + \boldsymbol{r}_{BC} = \boldsymbol{r}_{AC} \,. \tag{2.1}$$

Multiplying a proper vector by a positive number simply increases its length by that factor. Multiplying it by -1 just reverses its direction, so

$$\boldsymbol{r}_{AB} = -\boldsymbol{r}_{BA} \,. \tag{2.2}$$

When a point's position is loosely described as a (proper) vector, what is really meant is that the *displacement* of the point from some given origin is a proper vector. Although one can add vectors (both proper and coordinate) and multiply them by numbers, points cannot be added or multiplied in the same way.¹

Describing objects by the arrows that are the primordial proper vectors has given rise to the field of linear algebra, in which proper vectors become codified into elements of a *vector space*. All scientists are familiar with "adding" arrows top-to-tail, and with this operation, the axioms that define a vector space—found in every book on linear algebra—are easily understood when we use arrows as examples of the elements of the vector space. For example, scaling a proper vector, and adding two of them, produces another proper vector.

¹In relativity an event becomes a point in spacetime, and proper vectors are called *four-vectors* in the four dimensions of spacetime. Just as a point is not a vector in ordinary spatial analysis, in relativity an event is not a four-vector. But you will often see the spacetime coordinates (t, x, y, z) of an event mistakenly called a four-vector. We might well construct a four-vector whose tail is at some event designated as the origin of the spacetime coordinates, and whose head is at the event (t, x, y, z); but that is a very specific state of affairs, because there is generally nothing special about the coordinate origin. This is discussed further (for space only) in Section 7.3.

3 The Second Concept: Reference Frames

A reference frame, or simply "frame", can be pictured as a rigid lattice attached to an observer, relative to which that observer quantifies the motion of objects. In the classical mechanics of aerospace, this single observer makes measurements in some way that is not (and need not be) always specified. I use the words "frame" and "observer" interchangeably in this report. Although not discussed further here, in the subject of special relativity it becomes necessary to imagine the lattice as populated with a continuum of agents, each occupying a fixed point on the lattice and holding their own clock, who each record the positions and times of events only in their "very close" vicinity, and who send their data to a master station that periodically collates this information to form a global picture of the scenario of interest. But even this complex setup is called both a frame and a single observer.

The idea of reference frames is rooted in notions of statics, kinematics (motion), and dynamics (forces and masses). Begin with an idea of the points at the corners of the lattice of the frame being used. Consider the *displacement* of a given point from some reference point. Being an arrow, displacement is a proper vector, and in fact is the primordial proper vector. For example, the displacement of Adelaide relative to (i.e. from) Perth is the arrow embedded in Earth whose tail is at Perth and whose head is at Adelaide. All observers agree on the nature and identity of this displacement vector, because it can be constructed as a giant wooden arrow connecting Perth and Adelaide. All observers, regardless of their relative orientation or motion, will construct the same arrow—neglecting considerations of special relativity here and throughout this report. The displacement of Adelaide from Perth is a well-defined, completely unique proper vector, and requires no further information attached to it, such as the identity of any observer. So the notation \mathbf{r}_{AP} for the displacement (vector) of Adelaide relative to Perth does not need any choice of frame specified.

Next, consider kinematics: the subject of velocity, or how displacement vectors vary with time:

velocity vector
$$\equiv \frac{\mathrm{d}}{\mathrm{d}t}$$
 (displacement vector). (3.1)

Time-differentiating a changing proper vector entails subtracting that vector at time t from that vector at time $t + \Delta t$, dividing the result by Δt , and taking the limit $\Delta t \rightarrow 0$. In what amounts to the same thing, we can subtract the vector at time t from the vector at time t + dtand divide the result by dt (this process invisibly incorporates taking the limit), as shown in Figure 1. This differentiation can in principle be carried out purely geometrically, without using numbers.

Observers who rotate relative to each other will obtain different results when they measure how a displacement vector is changing with time. For example, everyone at rest on Earth agrees that Adelaide has no velocity relative to Perth, because they each measure the displacement vector of Adelaide from Perth not to change with time. But an observer hovering outside Earth, at rest relative to the distant stars and not rotating relative to them, will see the displacement vector of Adelaide from Perth rotate through 360° once every "sidereal day" (about 23 hours 56 minutes), and will conclude that the velocity of Adelaide relative to Perth is non-zero (and in fact the velocity vector itself also rotates through 360° once per sidereal day, as do all of its time derivatives). I signal this difference in the calculation of velocity by saying that whereas all observers fixed to Earth's surface inhabit the same *frame*—usually called the Earth-Centred Earth-Fixed frame, or ECEF²—the observer hovering at rest rela-

²"Earth-Centred" actually has nothing to do with the frame, and presumably refers to one choice of coordinates for it. See the footnote on page 5 for further discussion.



Figure 1: In frame F, the velocity of Adelaide A relative to Perth P is found from the (black) displacement vectors of Adelaide relative to Perth at times t and t + dt: we translate one displacement vector so that both share a common tail, which facilitates subtracting the earlier from the later vector and dividing the (red) resulting vector by dt. The result is the velocity \mathbf{v}_{AP}^F of Adelaide A relative to Perth P in frame F.

tive to the stars inhabits a different frame, the Earth-Centred Inertial frame (ECI). So, two observers are defined to inhabit the same frame if they construct the same velocity (proper vector, an arrow) of any given point relative to some other given point.

It follows that when describing or analysing a kinematical situation, one must *always* specify the frame being used. Any analysis that does not make this choice clear is set to fail at some point.

Finally, consider dynamics, the study of how forces affect masses. This is the realm of Newton's laws, which apply in their simplest form only to *inertial frames*. (Using other frames requires introducing "pseudo forces", which are really just book-keeping devices calculated by applying Newton's laws in an inertial frame.) A frame is defined to be inertial if a free mass maintains a constant velocity in that frame. No frame is truly inertial. We can imagine hovering high above Earth, at rest relative to the distant stars, watching Earth rotate. If we throw a ball, the ball will hold an almost-constant velocity for a short while before the acceleration it receives from Earth's gravity begins to change the ball's velocity to a measurable extent. The higher we hover, the weaker Earth's gravity will be at our location, and the more inertial our frame will be, and over a longer time. The ECI is the non-rotating frame constructed by one who is far enough from Earth to be treated as permanently inertial; but this frame does extend far enough into space from that observer to include Earth within it.

4 The Third Concept: Coordinate Systems

In any scenario, kinematic quantities such as position, velocity, acceleration, and angular momentum must be definable as proper vectors by examining the scenario of interest in a chosen frame. Only when we wish to perform numerical calculations might we decide to introduce a coordinate system.

A frame and a coordinate system are very often confused and thought to be the same thing, but they are quite different entities and ideas. A frame is a construct that enables us to recognise motion (if not quantify it), whereas a coordinate system is a collection of sets of

numbers that specifies the locations of *points*. The world can be treated as formed of points that exist in an absolute sense, and which have nothing to do with any frame. Two observers (i.e. different frames) will agree completely on the identity of a particular point or event. While it might sound obvious or trivial, it is worth emphasising that if each observer places a fingertip at the chosen point, then their fingertips will touch. The observers may well have some relative rotation, but a point in space is a primal concept of location: the point cannot rotate because it simply has no structure that is rotatable.

Given a frame, a coordinate system can be created by labelling with numbers the lattice points that make up the frame, so that each point has associated with it a set of coordinates. This is one way of constructing a coordinate system, but it does not imply that given a frame, we are obliged to use coordinates that were constructed in this way. Instead we are free to use a different coordinate system. The fact that this other coordinate system was constructed by attaching numbers to a different frame is of no importance. A given coordinate system need not have any relationship to a given frame. For example, coordinates that are fixed in the ECEF can certainly be used to quantify a velocity measured in the ECI. Or we might choose cartesian coordinates originating at an aircraft's centre and with axes embedded in the aircraft's body, to describe the orbit of a satellite around Earth. But although frames and coordinates are not related, a choice of frame may well suggest some natural choices of coordinates. Consider the ECEF, which describes the view of an observer for whom no part of Earth's body moves. A particular choice of cartesian coordinates (defined in Section 6) that is often used with the ECEF has its axes originating at Earth's centre and embedded in the solid body of the planet. The x axis emerges at $\frac{1}{\sqrt{0}} \frac{0^{\circ}}{0^{\circ}}$, the y axis emerges at lat/long 0°/90°, and the z axis completes a right-handed set.³

A second coordinate system used for the ECEF is latitude/longitude/height. A third is, say, the north–east–down set of axes originating at any *fixed* place of interest, such as Adelaide; like the previous two coordinate systems, these axes define a set of xyz coordinates for *any* point in the universe (and not just "near" Earth's surface).

A natural coordinate system for the ECI is a cartesian set of axes originating at Earth's centre, that are fixed relative to the distant stars. The x axis points to the *First Point of Aries*, a point in the sky defined by Earth's spin axis (the z axis) and Earth's orbit around the Sun, and which changes only slowly over the course of centuries. The y axis completes the right-handed set.

In relativity each reference frame is traditionally given its own dedicated coordinate system, so that discussion of frames tends to be inseparable from discussion of the relevant coordinates. Also, the arrows that all observers agree to have some kind of "objective" reality become four dimensional, joining *events* in space and time rather than *points* in space. The interesting topic of whether one observer might gain from using the coordinate system of another observer is not conventionally discussed in the subject.

³Does "ECEF" denote a frame or a coordinate set? While "Earth-Fixed" describes the view of Earth as fixed (so relates to a frame), "Earth-Centred" relates to the cartesian coordinates just described, that originate at Earth's centre. Perhaps this very acronym highlights the general confusion that exists between frames and coordinates.

DST-Group-TR-3309

5 The Fourth Concept: Coordinate Vectors

Proper vectors are usually simply called vectors. But another type of object is also commonly called a vector: a *coordinate vector*. This is a row or column of numbers that describes a proper vector using some choice of coordinates: we will normally write it as a column, so in n dimensions it's an $n \times 1$ matrix of numbers; but we make an exception when it occurs within a line of text, where for space reasons it will be written as a row, and with parentheses instead of brackets to indicate this exception is being made. And although points are *not* vectors, we will also write the coordinates of points as a parenthesised row of numbers; there should be no confusion.

The *Correspondence Principle*, a basic statement of linear algebra, says that any proper vector (an arrow) can always be represented numerically by a coordinate vector (an array of numbers). Proper vectors are the arrows that we draw to develop an intuitive feel for a scenario; coordinate vectors are the arrays of numbers that we use for calculation. A given proper vector can be associated with an infinite number of different coordinate vectors, corresponding to an infinite number of different choices of coordinate system. For the same reason, a given coordinate vector can be associated with an infinite number of different proper vectors. In the absence of a coordinate system, a given proper vector has no connection with a given coordinate vector. They are two entirely different objects.

For any proper vector \boldsymbol{v} , we write the coordinate vector that quantifies \boldsymbol{v} in a coordinate system S as $[\boldsymbol{v}]_S$ (as in [3]), so that $[\boldsymbol{v}]_S$ is a (say, 3×1) matrix of numbers. Given \boldsymbol{v} and S, the proper and coordinate vectors \boldsymbol{v} and $[\boldsymbol{v}]_S$ are two ways of describing the same entity. In this report I will usually write an arbitrary proper vector as \boldsymbol{v} , where the " \boldsymbol{v} " stands for "vector". (Note in particular that although velocity is a proper vector, \boldsymbol{v} need not specifically denote velocity in the generic discussions of this report.)

The "[]_S" notation is a linear operation, so is very easy to work with. The every-day adding of proper vectors by adding their components is written as $[\boldsymbol{a} + \boldsymbol{b}]_S = [\boldsymbol{a}]_S + [\boldsymbol{b}]_S$: "the components of the sum are the sums of the components". Scaling a vector by a real number r is written $[r\boldsymbol{a}]_S = r[\boldsymbol{a}]_S$: "the components of a multiple are the multiples of the components".

Both the dot product and cross product are defined for both proper vectors and coordinate vectors in such a way that the following hold for any vectors $\boldsymbol{a}, \boldsymbol{b}$ and choice of coordinates S:

$$\boldsymbol{a} \cdot \boldsymbol{b} = [\boldsymbol{a}]_S \cdot [\boldsymbol{b}]_S, \quad [\boldsymbol{a} \times \boldsymbol{b}]_S = [\boldsymbol{a}]_S \times [\boldsymbol{b}]_S.$$
 (5.1)

Normally of course, we don't apply such rules explicitly; they simply describe what we are doing when we "dot" or "cross" two vectors in the usual way by working with their components. But it is worth noting that, for example, the dot is being used in two separate ways here: it combines two proper vectors via the usual expression " $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos \theta$ ", and it combines two coordinate vectors via a rule that is written for cartesian coordinates as $(a_x, a_y, a_z) \cdot (b_x, b_y, b_z) = a_x b_x + a_y b_y + a_z b_z$. (For more general coordinates this rule is written similarly but with a weight factor attached to each of the six possible products of the components. These weights are the relevant "metric coefficients".) A similar comment applies to the cross-product operation.

A final point is that whereas a matrix can multiply a coordinate vector (a column of numbers), it cannot multiply a proper vector, since "matrix times arrow" is not defined. Thus in the context of this report where nothing is omitted in the notation, " $M[v]_S$ " is meaningful

when M is a matrix, but "Mv" is not. In the sections to follow, we will see similarly that when F denotes a frame, " $d^F v/dt$ " and " $d[v]_S/dt$ " are meaningful, but "dv/dt" is not.

The relationship of a proper vector (the actual object) to a coordinate vector (that object's representation is some chosen coordinate system) is similar to the relationship of a number to its representation in some base. A number is an object with a concrete existence (such as the number of fingers on each of my hands), and it has a numeric-string representation in some chosen basis: this number of fingers is written "5" in base 10 (or indeed in any base higher than 5), but "12" in base 3. A given number can be written as various strings of numerals depending on the base chosen, and a given string of numerals can denote various numbers depending on the base chosen.

6 Basis Vectors

A coordinate system has associated with it *at each point* a set of *basis vectors*, one for each dimension/coordinate. We will take it as given that any vector can be written as a linear combination of basis vectors, a fact that can be found in any book on introductory linear algebra.

When the basis vectors at each point are mutually orthogonal, of unit length, and don't change their orientation from one point to the next, the associated coordinate system is called *cartesian*, and is especially easy to work with. A flying aircraft typically defines its own cartesian coordinate axes as embedded in its body, but even if the aircraft flies straight and level over Earth's curved surface, an observer at rest on Earth will see the relevant basis vectors change from point to point, and must incorporate that change if using the aircraft's coordinates.

Suppose that in three dimensions we have a set of u, v, w coordinates which may or may not be cartesian. A curve along which one of these coordinates changes while the other two are held fixed need not be a straight line. We construct the u basis vector " e_u " at any point Pin the following way. (Because P is a point, not a vector, we don't write it bold.) Start at P, and step an infinitesimal amount in u while holding v and w constant. The end of the step is the new point P + dP, where the vector pointing from start to end points is written dP(and we have signalled that it is a proper vector by writing a bold P). The basis vector e_u at P is defined to be this infinitesimal vector divided by the infinitesimal increase du in u:

$$\boldsymbol{e}_u = \mathrm{d}\boldsymbol{P}/\mathrm{d}\boldsymbol{u}, \quad \text{with } \boldsymbol{v}, \boldsymbol{w} \text{ constant.}$$
 (6.1)

This vector is the same regardless of whether du is positive or negative; that is, we could step in either direction along the relevant curve. Treating P as a function of the coordinates, we could write

$$\boldsymbol{e}_u \equiv \partial P(u, v, w) / \partial u \,. \tag{6.2}$$

With this definition (6.2), the vector that represents an infinitesimal step away from P in an arbitrary direction is written using the chain rule of partial differentiation as

$$d\boldsymbol{P} = du \frac{\partial P}{\partial u} + dv \frac{\partial P}{\partial v} + dw \frac{\partial P}{\partial w} = du \boldsymbol{e}_u + dv \boldsymbol{e}_v + dw \boldsymbol{e}_w.$$
(6.3)

(We cannot generally write a similar expression for a *non*-infinitesimal step away from P, because the basis vectors might change with position.) With S labelling the coordinate system

DST-Group-TR-3309

 $\{u, v, w\}$, the coordinate vector of this step d**P** is

$$[\mathbf{d}\boldsymbol{P}]_S = \begin{bmatrix} \mathbf{d}\boldsymbol{u} \\ \mathbf{d}\boldsymbol{v} \\ \mathbf{d}\boldsymbol{w} \end{bmatrix}.$$
(6.4)

This is a familiar result when using cartesian coordinates, but we now see that it also holds for non-cartesian coordinates. The catch is that the basis vectors e_u, e_v, e_w might not have unit length or be orthogonal, so that the dot- and cross-product operations will need to be modified for such coordinates.

As a side comment, we might write the components of a *point* as

$${}^{\prime\prime}\left[P\right]_{S} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} .{}^{\prime\prime} \tag{6.5}$$

But we must be careful to remember that this is just notation: the non-bold P here signifies that it is a point and not a vector. Equation (6.5) certainly does not say that P equates to a vector $ue_u + ve_v + we_w$, because if that were so, which e_u, e_v, e_w would we mean here, since in general they are functions of position? A similar expression *is* often written when using cartesian coordinates: you will often see a point written in such coordinates as "xi + yj + zk", where i, j, k denote e_x, e_y, e_z respectively. (On a historical note, the use of i, j, k is a quaternion convention predating vectors, which derived from the non-bold *i* representing $\sqrt{-1}$.) But this description of a point as a displacement from the coordinate origin is valid only because the cartesian basis vectors e_x, e_y, e_z don't change with position.

Given two coordinate systems u, v, w and u', v', w', the chain rule of partial differentiation says, e.g. for $e_{u'}$,

$$\boldsymbol{e}_{u'} = \frac{\partial P}{\partial u'} = \frac{\partial u}{\partial u'} \frac{\partial P}{\partial u} + \frac{\partial v}{\partial u'} \frac{\partial P}{\partial v} + \frac{\partial w}{\partial u'} \frac{\partial P}{\partial w}$$
$$= \frac{\partial u}{\partial u'} \boldsymbol{e}_u + \frac{\partial v}{\partial u'} \boldsymbol{e}_v + \frac{\partial w}{\partial u'} \boldsymbol{e}_w, \qquad (6.6)$$

which shows how to relate two sets of basis vectors. We see from (6.6) that the S coordinates of $e_{u'}$ are

$$[\mathbf{e}_{u'}]_S = \begin{bmatrix} \frac{\partial u}{\partial u'} \\ \frac{\partial v}{\partial u'} \\ \frac{\partial w}{\partial u'} \end{bmatrix} = \frac{\partial}{\partial u'} \begin{bmatrix} u \\ v \\ w \end{bmatrix}.$$
(6.7)

There is nothing mysterious about differentiating a matrix in (6.7). Differentiation uses a subtraction and a scaling, and these are both defined component-wise for a matrix. So the derivative of a matrix is simply the matrix of derivatives of its components.

The resemblance of (6.7) to (6.5) is somewhat superficial, but is potentially a cause for confusion. We must realise here that the (u, v, w) on the right-hand sides of (6.5) and (6.7) is not a coordinate vector; it is merely an array of the coordinates of the point P.

As an example, we calculate the cartesian coordinates of the polar basis vectors $\boldsymbol{e}_r, \boldsymbol{e}_{\theta}$. Begin with

$$x = r\cos\theta, \quad y = r\sin\theta.$$
 (6.8)

We could write

$$\boldsymbol{e}_{r} = \frac{\partial x}{\partial r} \, \boldsymbol{e}_{x} + \frac{\partial y}{\partial r} \, \boldsymbol{e}_{y} \,, \quad \boldsymbol{e}_{\theta} = \frac{\partial x}{\partial \theta} \, \boldsymbol{e}_{x} + \frac{\partial y}{\partial \theta} \, \boldsymbol{e}_{y} \,, \tag{6.9}$$

DST-Group-TR-3309



Figure 2: Basis vectors for polar and cartesian coordinates in two dimensions. The cartesian basis vectors have unit length and are identical everywhere. The polar basis vector \mathbf{e}_r points everywhere radially outward from the origin of the polar coordinates, and has unit length. The polar basis vector \mathbf{e}_{θ} is everywhere transverse, and is easily shown to have length r.

or—the same idea but more compact—we could apply (6.7), writing C for the cartesian coordinates:

$$[e_r]_C = \partial_r \begin{bmatrix} x\\ y \end{bmatrix} = \partial_r \begin{bmatrix} r\cos\theta\\ r\sin\theta \end{bmatrix} = \begin{bmatrix} \cos\theta\\ \sin\theta \end{bmatrix}$$
(which has unit length),
$$[e_\theta]_C = \partial_\theta \begin{bmatrix} x\\ y \end{bmatrix} = \begin{bmatrix} -r\sin\theta\\ r\cos\theta \end{bmatrix}$$
(which has length r). (6.10)

Examples of these basis vectors for polar coordinates along with basis vectors for cartesian coordinates are shown in Figure 2. The cartesian basis vectors always point in the same direction regardless of where they are situated: e_x always points right in the figure, and e_y always points up. Because of this, whenever cartesian basis vectors are drawn at all, they tend to be drawn just once, typically with their tails at the coordinate origin. In contrast, the polar basis vectors really do change from point to point. At each point, the radial vector e_r always points radially out from the origin of the polar coordinates and has unit length, and the transverse vector e_{θ} is found by rotating this e_r 90° counterclockwise and giving the result length r.

The above definition (6.2) of a basis vector \mathbf{e}_u as a partial derivative $\partial P/\partial u$ has led to the notation " $\partial/\partial u$ " sometimes being used to denote that basis vector. There is no implication in the above analysis that a basis vector is the *operator* $\partial/\partial u$: the operating has already been done on P, so to speak. Nevertheless, the modern language of differential geometry does in fact omit P, defining \mathbf{e}_u as precisely this partial derivative operator $\partial/\partial u$; and more generally, it defines a vector to be a linear combination of these operators. This particular abstraction finds no application in aerospace. To use proper vectors in real physical scenarios, define them to be geometrical arrows as we have done here, and then you will always be able to construct a meaningful and useful picture.

The handedness of a three-dimensional coordinate system u, v, w is determined by its basis

DST-Group-TR-3309

vectors e_u, e_v, e_w : if the angle between $e_u \times e_v$ and e_w is between 0 and 90°, the coordinate system is called right handed; otherwise it is left handed. Although coordinate systems can have any handedness, it is wise to ensure that all such systems being used concurrently have the same handedness. The usual convention is that they are right handed. Right handedness is built in, for example, to the standard expression for calculating a cross product from vector components, so that calculating the cross product for the left-handed case would require using a non-standard formula. Although there is nothing wrong with doing this, one must not lose one's audience along the way, and certainly it is considered very conventional and safe to settle on one convention, which is that coordinates are constructed as right handed in aerospace. In fact they are typically constructed as orthonormal: the basis vectors all have unit length and are mutually orthogonal. In the case of polar coordinates, the basis vectors are orthogonal but not all of unit length. Even so, they can be normalised, and examples of this are given in the next section and in Appendix A.1.

6.1 Example: Basis Vectors for Polar Coordinates

What are the basis vectors for a latitude/longitude/height coordinate system of the ECEF? Write⁴

 $\lambda = \text{latitude}, \quad \phi = \text{longitude}, \quad h = \text{height above sea level.}$ (6.11)

For simplicity, suppose Earth is a perfect sphere of sea-level radius R. (Earth's actual oblate spheroid shape means that the following equations don't quite apply to it, but this discussion is for illustration only. The main ideas here extend to a surface of any shape.) At any point (λ, ϕ, h) which may or may not be on Earth's surface, consider making a single step north by some infinitesimal latitude $d\lambda > 0$ (meaning longitude and height are held constant), and constructing the corresponding displacement vector d \mathbf{P} . This points north and has length $(R + h) d\lambda$. Following (6.1), divide that vector by $d\lambda$ to give \mathbf{e}_{λ} , which thus points north with length R + h. We could equally well step south instead of north, in which case the latitude increase is $d\lambda < 0$: when the south-pointing displacement vector d \mathbf{P} is divided by $d\lambda$ (which is now negative), the resulting \mathbf{e}_{λ} will again point north as it should, since it's defined to be unique at any point.

Similarly, e_{ϕ} points east with length $(R+h)\cos\lambda$, and e_h points up with length one. These vectors are easily expressed in the commonly used xyz coordinates for the ECEF that originate at Earth's centre, as shown in Figure 3. First, relate λ, ϕ, h in the usual spherical polar way to the cartesian set x, y, z via

$$x = (R+h)\cos\lambda\cos\phi,$$

$$y = (R+h)\cos\lambda\sin\phi,$$

$$z = (R+h)\sin\lambda.$$
(6.12)

Now write the proper vector e_{λ} as, following (6.6),

$$e_{\lambda} = \frac{\partial x}{\partial \lambda} e_{x} + \frac{\partial y}{\partial \lambda} e_{y} + \frac{\partial z}{\partial \lambda} e_{z}$$

= -(R+h) sin \lambda cos \phi \end{e}_{x} - (R+h) sin \lambda sin \phi \end{e}_{y} + (R+h) cos \lambda \end{e}_{z}, (6.13)

⁴The symbols in (6.11) are also used by others, and make sense for two reasons: (1) ϕ is very commonly used as the angle of longitude in spherical polar coordinates even when Earth is not present, and (2) the first two letters of "lambda" match those of "latitude". But be aware that some practitioners call longitude λ and latitude ψ , which matches no other implementation of spherical polar coordinates.



Figure 3: The geometry defining the coordinates (λ, ϕ, h) in (6.12)

and thus as a coordinate vector in xyz coordinates:

$$[\boldsymbol{e}_{\lambda}]_{xyz} = \begin{bmatrix} -(R+h)\sin\lambda\cos\phi\\ -(R+h)\sin\lambda\sin\phi\\ (R+h)\cos\lambda \end{bmatrix}.$$
(6.14)

Similar analysis gives the coordinate vectors $[e_{\phi}]_{xyz}$ and $[e_h]_{xyz}$.

The lengths and mutual angles of basis vectors are encapsulated in the *metric* coefficients of the relevant coordinate system. Define

$$g_{\alpha\beta} \equiv \boldsymbol{e}_{\alpha} \cdot \boldsymbol{e}_{\beta} \,. \tag{6.15}$$

For the above lat/long/height coordinates,

$$g_{\lambda\lambda} = (R+h)^2, \quad g_{\phi\phi} = (R+h)^2 \cos^2 \lambda, \quad g_{hh} = 1,$$
 (6.16)

and all other combinations $g_{\lambda\phi}$ etc. equal zero. Because e_{λ} , e_{ϕ} , e_{h} are mutually orthogonal, the lat/long/height coordinate system is called orthogonal. But it is not *orthonormal*, because the basis vectors are not all of unit length. And it's certainly not cartesian, because the basis vectors change direction and/or length from point to point.

The above gives mathematical meaning to the "east-north-up" coordinates frequently used in aerospace analyses: e_{ϕ} points east with length $(R + h) \cos \lambda$, e_{λ} points north with length R + h, and e_h points up with length one. But east-north-up coordinates are usually defined as an orthonormal basis, meaning their basis vectors have unit length:

> east unit-length basis vector $= u_{\phi} \equiv \frac{e_{\phi}}{|e_{\phi}|} = \frac{e_{\phi}}{(R+h)\cos\lambda}$, north unit-length basis vector $= u_{\lambda} \equiv \frac{e_{\lambda}}{|e_{\lambda}|} = \frac{e_{\lambda}}{R+h}$, up unit-length basis vector $= u_{h} \equiv \frac{e_{h}}{|e_{h}|} = e_{h}$. (6.17)

UNCLASSIFIED

11

DST-Group-TR-3309

For example, (6.14) then gives the coordinates of the north unit-length basis vector as

$$[\boldsymbol{u}_{\lambda}]_{xyz} = \begin{bmatrix} -\sin\lambda\cos\phi\\ -\sin\lambda\sin\phi\\ \cos\lambda \end{bmatrix}, \qquad (6.18)$$

and similarly for the east and up unit-length vectors, u_{ϕ} and u_h . If we work within the ECEF and agree to use the east-north-up coordinates originating at a *fixed* point such as Adelaide to describe any point (which could even be on Pluto), then those coordinates will certainly be cartesian because by definition, at any other point in space they are made to be parallel to those at Adelaide. On the other hand, if we embed the basis vectors in the body of an aircraft, then they can change from point to point along the aircraft's track if it manoeuvres, and so will not be cartesian. They will be defined only at points occupied by the aircraft; and if the aircraft visits a given point more than once, the basis vectors at that point will be updated to the latest set each time the point is occupied.

7 Relating Different Coordinate Systems

One or more observers might express a proper vector \boldsymbol{v} in multiple coordinate systems: call those systems A and B. Vector \boldsymbol{v} is written as $[\boldsymbol{v}]_A$ using A coordinates, and as $[\boldsymbol{v}]_B$ using B coordinates. How are the 3×1 matrices $[\boldsymbol{v}]_A$ and $[\boldsymbol{v}]_B$ related? I answer this question in the section that follows, and then extend the idea beyond the realm of vectors.

7.1 Relating Vector Components

The following is a standard approach of linear algebra that shows how to transform a coordinate vector between coordinate systems. It requires the basis vectors to be orthonormal—meaning of unit length and mutually orthogonal—so we remind ourselves of that by writing the orthonormal basis vectors of coordinate system A as (keeping to 2 dimensions for economy of notation) $\boldsymbol{u}_{x_A}, \boldsymbol{u}_{y_A}$, and the orthonormal basis vectors of coordinate system B as $\boldsymbol{u}_{x_B}, \boldsymbol{u}_{y_B}$. (The following can be re-expressed for a general basis, but some complexity arises, and since aerospace universally uses only orthonormal bases, we confine discussion to those. The details for a general basis are given in Appendix A.) Starting with an arbitrary proper vector \boldsymbol{v} written in two coordinate systems as

$$v = v_{x_A} u_{x_A} + v_{y_A} u_{y_A} = v_{x_B} u_{x_B} + v_{y_B} u_{y_B}, \qquad (7.1)$$

the relevant coordinate vectors of \boldsymbol{v} are

$$[\boldsymbol{v}]_{A} = \begin{bmatrix} v_{x_{A}} \\ v_{y_{A}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v} \cdot \boldsymbol{u}_{x_{A}} \\ \boldsymbol{v} \cdot \boldsymbol{u}_{y_{A}} \end{bmatrix}, \quad [\boldsymbol{v}]_{B} = \begin{bmatrix} v_{x_{B}} \\ v_{y_{B}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v} \cdot \boldsymbol{u}_{x_{B}} \\ \boldsymbol{v} \cdot \boldsymbol{u}_{y_{B}} \end{bmatrix}.$$
(7.2)

Now consider

$$\begin{bmatrix} \boldsymbol{v} \end{bmatrix}_{B} = \begin{bmatrix} \boldsymbol{v} \cdot \boldsymbol{u}_{x_{B}} \\ \boldsymbol{v} \cdot \boldsymbol{u}_{y_{B}} \end{bmatrix} = \begin{bmatrix} (v_{x_{A}}\boldsymbol{u}_{x_{A}} + v_{y_{A}}\boldsymbol{u}_{y_{A}}) \cdot \boldsymbol{u}_{x_{B}} \\ (v_{x_{A}}\boldsymbol{u}_{x_{A}} + v_{y_{A}}\boldsymbol{u}_{y_{A}}) \cdot \boldsymbol{u}_{y_{B}} \end{bmatrix}$$
$$= \begin{bmatrix} \boldsymbol{u}_{x_{A}} \cdot \boldsymbol{u}_{x_{B}} & \boldsymbol{u}_{y_{A}} \cdot \boldsymbol{u}_{x_{B}} \\ \boldsymbol{u}_{x_{A}} \cdot \boldsymbol{u}_{y_{B}} & \boldsymbol{u}_{y_{A}} \cdot \boldsymbol{u}_{y_{B}} \end{bmatrix} \begin{bmatrix} v_{x_{A}} \\ v_{y_{A}} \end{bmatrix} \equiv \mu_{B}^{A} [\boldsymbol{v}]_{A} . \tag{7.3}$$

That is, the *direction-cosine* or *orientation* matrix μ_B^A transforms coordinates from A to B:

$$[\boldsymbol{v}]_B = \mu_B^A [\boldsymbol{v}]_A, \qquad (7.4)$$

where (7.3) shows that the *columns* of μ_B^A are the basis vectors of A written in B coordinates:

$$\mu_B^A \equiv \left[\begin{bmatrix} \boldsymbol{u}_{x_A} \end{bmatrix}_B \quad \begin{bmatrix} \boldsymbol{u}_{y_A} \end{bmatrix}_B \right].$$
(7.5)

Equation (7.4) has here been derived for coordinates with orthonormal basis vectors, but in fact it holds also for general coordinates, as shown in Appendix A. Although orthonormal coordinates are the only type ever used in aerospace analysis, we will derive various results in the pages to come using (7.4), which are thus true for more general coordinate systems.

It might be thought that (6.7) can now be used to calculate the matrix μ_B^A , but that would be so only if equations relating the A coordinates to B coordinates were available—and if all basis vectors were orthonormal. In practice such equations are generally unavailable, so the matrix μ_B^A must be calculated by other means. See [4] for details.

It's clear that

$$[\boldsymbol{v}]_A = \mu_A^B \, \mu_B^A \, [\boldsymbol{v}]_A \,, \tag{7.6}$$

so that

$$\mu_A^B = \left(\mu_B^A\right)^{-1}.\tag{7.7}$$

For orthonormal coordinates, (7.3) shows that the rows of μ_B^A are the basis vectors of B written in A coordinates. From this, it's not hard to see that

$$\mu_A^B = \left(\mu_B^A\right)^t. \tag{7.8}$$

So a μ matrix for orthonormal coordinates is orthonormal: its inverse equals its transpose. This is *not* the case for non-orthonormal coordinates.

The μ matrices can be "chained" together; for example, with three coordinate systems:

$$[\boldsymbol{v}]_A = \mu_A^B [\boldsymbol{v}]_B = \mu_A^B \, \mu_B^C [\boldsymbol{v}]_C \equiv \mu_A^C [\boldsymbol{v}]_C \,, \tag{7.9}$$

showing that

$$\mu_A^C = \mu_A^B \, \mu_B^C \,, \tag{7.10}$$

and similarly for any number of coordinate systems.

7.2 Relating Rank-2 Tensor Components

Just as a proper vector \boldsymbol{v} is a geometrical object whose A-coordinates can be written as a 3×1 matrix $[\boldsymbol{v}]_A$, so too a linear operator L that transforms \boldsymbol{v} to another vector $L\boldsymbol{v}$ can be coordinatised, and is most usefully then written as a two-dimensional matrix $[L]_A$. This is a fundamental idea in linear algebra, where L might be a rotation, or a stretch. To match the geometrical picture of \boldsymbol{v} as an arrow, the operator L can be envisaged as an object in its own right that combines with a proper vector to produce a new proper vector. Whereas differential geometry has taken a proper vector (an arrow) and redefined it to be an operator, here we are taking an operator (L) and treating it as an object in its own right that "combines" with a proper vector \boldsymbol{v} to produce $L\boldsymbol{v}$.

DST-Group-TR-3309

Just as a proper vector is independent of any choice of coordinates, so too this new object L, called a *tensor*, is independent of any choice of coordinates; e.g. we can rotate an arrow in space without any reference to a coordinate system. Similar to proper vectors and coordinate vectors, we might call L a "proper tensor" to distinguish it from its coordinates $[L]_A$, a "coordinate tensor". But perhaps universally the proper tensor is simply called a tensor; and when it operates on a vector, the coordinate tensor can usefully be written as a matrix, so it tends to be called simply a matrix. The relationship between coordinates and tensor/vector is

$$[L]_A [\boldsymbol{v}]_A = [L\boldsymbol{v}]_A.$$
(7.11)

For example, if A is the xy cartesian coordinate system in the plane and L rotates a vector counter-clockwise through angle θ ,

$$[L]_{A} = \begin{bmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta \end{bmatrix}, \quad [\boldsymbol{v}]_{A} = \begin{bmatrix} x\\ y \end{bmatrix}, \quad [L\boldsymbol{v}]_{A} = \begin{bmatrix} x\cos\theta - y\sin\theta\\ x\sin\theta + y\cos\theta \end{bmatrix}.$$
(7.12)

Extending Section 7.1 we ask the question: how do the matrices $[L]_A$ and $[L]_B$ relate? Consider:

$$[L]_{B}[\boldsymbol{v}]_{B} = [L\boldsymbol{v}]_{B} \xrightarrow{(7.4)} \mu_{B}^{A} [L\boldsymbol{v}]_{A} = \mu_{B}^{A} [L]_{A} [\boldsymbol{v}]_{A} = \mu_{B}^{A} [L]_{A} \mu_{A}^{B} [\boldsymbol{v}]_{B}, \qquad (7.13)$$

and since v is arbitrary, it follows that

$$[L]_B = \mu_B^A \ [L]_A \ \mu_A^B \,. \tag{7.14}$$

Equations (7.4) and (7.14) are fundamental to any study of coordinatising tensors (see e.g. [2] for the use of these expressions in aerospace).

L is a "rank-2 tensor" because it needs two indices to coordinatise it, corresponding to its matrix representative being 2-dimensional. Proper vectors are also tensors because of their "concrete" existence independent of a coordinate system. Needing a single coordinate index to describe them (and hence being naturally written as arrays), they are called rank-1 tensors. Unlike vectors with their pictorial representation as an arrow, higher-rank tensors are perhaps not readily visualised.

The simplest tensor is a rank-0 tensor, called a *scalar*, which is a number with no coordinate dependence—so no coordinate index is necessary. The height at any point on a mountain and the temperature at any point in a room are both scalars. The word "scalar" is often loosely used to denote simply a number, but it has a more refined meaning than that: it is a value of some quantity that doesn't depend on coordinate choice.

The main feature of a tensor of any rank is that as an "object", it is independent of the coordinate system chosen to represent it. As an example with a 2-dimensional vector \boldsymbol{v} , the sum $v_{x_A}\boldsymbol{e}_{x_A} + v_{y_A}\boldsymbol{e}_{y_A}$ equals the sum $v_{x_B}\boldsymbol{e}_{x_B} + v_{y_B}\boldsymbol{e}_{y_B}$ (which equals \boldsymbol{v}), but the coordinate vector (v_{x_A}, v_{y_A}) is not required to equal (v_{x_B}, v_{y_B}) . An older definition of a tensor (which includes a vector as one possible type) is modelled on the spatial-invariance property of a vector, but this older definition acknowledges the existence of coordinate vectors but not proper vectors. So this older definition treats the tensor as comprised in some sense only of its coordinates, defining the tensor to be the aggregate of all possible coordinate sets that it can have, and stipulating that these coordinates must transform between their respective coordinate systems via, say, (7.4) or (7.14)—although expressions involving partial derivatives

are more usually used for that transform, following the analysis of Section 6. This "behaviour under a transform" definition follows from the above analysis, but when seen in isolation, it begs the question of why the coordinates *should* relate via (7.4) or (7.14). In the modern view, the tensor's coordinates certainly do depend on the coordinate system via (7.4) or (7.14), but the tensor itself does *not* depend on the coordinate system. This independence of coordinate system was known to the earliest researchers who incorporated tensors into physics, because that independence is the entire reason that tensors are useful. The old idea that a tensor is an aggregate of sets of numbers only mistakes the components for the tensor itself. The modern view defines a tensor as anything that does *not* depend on the choice of coordinates: a much simpler definition!

Remember that a matrix is not a tensor. Rather, a matrix is a convenient way to write the components of a rank-2 tensor in some chosen basis. This is just an extension of the statement that a column of three numbers is not a proper vector.

7.3 Interpretations of "Position"

As stated in Section 2, the position of a point A is not a proper vector. Instead, we can only quantify A's position relative to some other point B, in which case the displacement vector \mathbf{r}_{AB} might simply be called the "position of A", and almost certainly written as \mathbf{r}_A when the presence of B is understood. The "relativeness to B" is then implied, but we must always be aware of which point has been chosen as the reference B.

The reference point B is often the origin O of a set of cartesian coordinates that has been chosen to quantify the events of some frame. Suppose observer (i.e. frame) 1 uses a coordinate system 1, and observer 2 uses a coordinate system 2. Observer 1 chooses to define the position of A as \mathbf{r}_{AO_1} , meaning relative to the origin O_1 of coordinate system 1. Likewise, observer 2 chooses to define the position of A as \mathbf{r}_{AO_2} . These two positions are *different* proper vectors if O_1 and O_2 are different points. Moreover, each observer recognises both vectors as the real entities that they are; there is no disagreement anywhere. Using the usual way of adding vectors, the positions quantified by the two observers are related via

$$\mathbf{r}_{AO_1} = \mathbf{r}_{AO_2} + \mathbf{r}_{O_2O_1} \,, \tag{7.15}$$

as shown in Figure 4. The proper vector $r_{O_2O_1}$ describes the displacement of the origins of the two coordinate systems, but we must remember that positions have nothing to do with coordinate systems as such; it is simply the case that observers 1 and 2 deemed it useful to use the origins of their coordinate systems as their required reference points. They could have used any points whatsoever as reference points; they could even have used the same reference point.

When we switch to a new coordinate system and must transform the coordinates of a proper vector appropriately, we know that the new coordinates are obtained from a matrix multiplication with the old, by way of (7.4). But we must use care when applying this idea to the coordinates of positions. As mentioned several lines up, observer 1 will often define the position of point A as the displacement vector of A from a reference point O_1 that also serves as the origin of observer 1's coordinate system, hence writing the position of A as $[\mathbf{r}_{AO_1}]_1$. Observer 2 chooses to define the position of the same point A as its displacement vector from O_2 , the origin of observer 2's coordinate system, and hence writes A's position as $[\mathbf{r}_{AO_2}]_2$. The question is: how are these coordinate vectors $[\mathbf{r}_{AO_1}]_1$ and $[\mathbf{r}_{AO_2}]_2$ related?



Figure 4: Observers 1 and 2 are both free to choose any point relative to which they each specify the position of A as a displacement vector. But in practice, observer 1 probably chooses his own coordinate origin O_1 , and observer 2 probably likewise chooses his own coordinate origin O_2 . Even with this choice, observer 1 still has the option of choosing either $[\mathbf{r}_{AO_1}]_1$ or $[\mathbf{r}_{AO_1}]_2$ as coordinates, and observer 2 has the option of choosing either $[\mathbf{r}_{AO_2}]_1$ or $[\mathbf{r}_{AO_2}]_2$ as coordinates. Usually—but not always observer 1 chooses $[\mathbf{r}_{AO_1}]_1$ and observer 2 chooses $[\mathbf{r}_{AO_2}]_2$. These coordinate vectors are related to each other via the second line of (7.16).

Write

$$[\mathbf{r}_{AO_1}]_1 = [\mathbf{r}_{AO_2} + \mathbf{r}_{O_2O_1}]_1 = [\mathbf{r}_{AO_2}]_1 + [\mathbf{r}_{O_2O_1}]_1$$
$$= \mu_1^2 [\mathbf{r}_{AO_2}]_2 + [\mathbf{r}_{O_2O_1}]_1.$$
(7.16)

That is, the coordinates of the two observers' descriptions of the position of A are related via a matrix multiplication and a shift. There is nothing mysterious about this; it simply reflects the fact that positions are commonly referred to the origin of the coordinate system being used, and this origin might differ across different coordinate systems. It's certainly true that

$$[\mathbf{r}_{AO_2}]_1 = \mu_1^2 \, [\mathbf{r}_{AO_2}]_2 \,, \tag{7.17}$$

because a single displacement vector \mathbf{r}_{AO_2} is present here. But the simple fact is that observer 1 tends to write the position of A as $[\mathbf{r}_{AO_1}]_1$, not $[\mathbf{r}_{AO_2}]_1$. Because of this common and completely normal practice, we must carefully define what is meant by "position" or "position vector" when more than one coordinate system is in use. In general, don't assume that the coordinates of "positions" are related only by a matrix multiplication; instead, use the last line of (7.16).

8 The Velocity Vector

As discussed above, the position of a point A is defined only relative to some reference point B, and then is the proper vector \mathbf{r}_{AB} . In contrast, the velocity of A is not only defined using the reference point B, but also according to the *frame* F being used. "The velocity of A" has no a priori meaning; instead we require "the velocity of A relative to B in the frame F", \mathbf{v}_{AB}^F .

We can omit reference to B and F only when their identities are understood from the relevant context. This velocity is a proper vector due to its being defined in the usual way in the frame F with a subtraction of displacements (which are proper vectors), which can only produce a proper vector. Further, the S-coordinate vector of this velocity is $[\boldsymbol{v}_{AB}^F]_S$. This requires four entities to be specified: the point of interest A, the reference point B, the frame F, and the coordinate system S. Any supposed velocity coordinate vector for which these quantities are not known has no place in any analysis.

To show how the velocity vector is defined by a choice of frame, consider the velocity of Adelaide A relative to Perth P in the ECEF, v_{AP}^{ECEF} . Standing on Earth places us in the ECEF. We draw an arrow from Perth to Adelaide (which is the proper vector r_{AP}), and detect that it does not change with time. It follows that $v_{AP}^{\text{ECEF}} = 0$. Next consider the velocity of Adelaide relative to Perth in the ECI, v_{AP}^{ECI} . Hovering in space far from Earth, we draw the same arrow from Perth to Adelaide, r_{AP} , but now find that it rotates through 360° in one sidereal day. It follows that v_{AP}^{ECI} is not zero. We can calculate this velocity at any moment in the usual way of calculus by drawing r_{AP} at times t and t + dt, subtracting the initial vector from the final, and dividing the result by dt. We signal that this procedure is carried out within a particular frame with a superscript on the derivative d:

$$\boldsymbol{v}_{AP}^{\text{ECEF}} = \frac{\mathrm{d}^{\text{ECEF}}\boldsymbol{r}_{AP}}{\mathrm{d}t} = \boldsymbol{0}, \quad \boldsymbol{v}_{AP}^{\text{ECI}} = \frac{\mathrm{d}^{\text{ECI}}\boldsymbol{r}_{AP}}{\mathrm{d}t} \neq \boldsymbol{0}.$$
(8.1)

It might seem counterintuitive that "Adelaide can have a velocity relative to Perth", but this quoted phrase has no meaning as it stands. Adelaide may or may not have a velocity relative to Perth *in some specified frame*. Velocity is defined by watching how the displacement vector of Adelaide relative to Perth, \mathbf{r}_{AP} , changes with time, *in some specified frame*. And this vector *does* change with time in the ECI. We tend to live our day-to-day lives from the perspective of the ECEF, and because $\mathbf{v}_{AP}^{\text{ECEF}} = \mathbf{0}$, we might be inclined to think that "Adelaide has no velocity relative to Perth". But the frame must be specified for such a phrase to be meaningful.

The phrase "velocity of point A relative to point B" is sometimes assumed by non-specialists to involve "the rest frame of B" [5]. But such cannot be the case, simply because the point Bhas no unique rest frame. For suppose that B is Earth's centre: this point is at rest in the ECEF, but is also at rest in the ECI, because even though Earth rotates in the ECI, a point cannot rotate about itself; so Earth's centre is fixed in the ECI. But the velocity of Adelaide relative to Earth's centre is zero in the ECEF and clearly non-zero in the ECI. So the "velocity of Adelaide relative to Earth's centre in the rest frame of Earth's centre" is undefined.

The reference point B might be moving in the frame used to define the velocity. But all points that are fixed in that frame give rise to the same velocity when used as reference points. This is because the velocity vector is constructed by joining the *head* of a displacement vector at time t to the *head* of the new displacement vector at time t + dt, and dividing the resulting vector (arrow) by dt; the tail point shared by both displacement vectors plays no role here. We can also show this using the relevant notation in the following way. Consider a frame F, along with two points F_1, F_2 , fixed in frame F. We ask: how does the F-velocity of some point A relative to $F_1(v_{AF_1}^F)$ relate to the F-velocity of A relative to $F_2(v_{AF_2}^F)$?

$$\boldsymbol{v}_{AF_1}^F \equiv \frac{\mathrm{d}^F \boldsymbol{r}_{AF_1}}{\mathrm{d}t} = \frac{\mathrm{d}^F}{\mathrm{d}t} \left(\boldsymbol{r}_{AF_2} + \boldsymbol{r}_{F_2F_1} \right) = \boldsymbol{v}_{AF_2}^F + \frac{\mathrm{d}^F \boldsymbol{r}_{F_2F_1}}{\mathrm{d}t} \,. \tag{8.2}$$

But F_1 and F_2 are fixed in F, so certainly $d^F r_{F_2F_1}/dt = 0$. It follows that

$$\boldsymbol{v}_{AF_1}^F = \boldsymbol{v}_{AF_2}^F, \qquad (8.3)$$

UNCLASSIFIED

17

DST-Group-TR-3309

allowing us to drop all mention of a fixed reference point, and simply speak of "the velocity of A in frame F":

$$\boldsymbol{v}_A^F = \text{velocity of } A \text{ in frame } F \equiv \boldsymbol{v}_{AF_1}^F \text{ for any point } F_1 \text{ fixed in } F.$$
 (8.4)

It's now easy to see that

$$\boldsymbol{v}_{A}^{F} - \boldsymbol{v}_{B}^{F} = \boldsymbol{v}_{AF_{1}}^{F} - \boldsymbol{v}_{BF_{1}}^{F} \stackrel{(2.1)}{==} \boldsymbol{v}_{AB}^{F}, \qquad (8.5)$$

because (2.1) applies to any vectors (i.e. arrows), not just displacement vectors. So we see that the velocity of A relative to B in frame F equals the velocity of A in frame F minus the velocity of B in frame F. This relation is often used intuitively. Zipfel [2] calls v_{AB}^F the "differential velocity" of point A relative to point B in frame F, and calls v_A^F the "linear velocity" of point A in frame F. These two differently named velocities are fundamentally the same thing: they are both the velocity of A relative to some possibly moving reference point in frame F. The same comments apply to Zipfel's use of the terms "differential acceleration" and "linear acceleration".

9 The Angular Velocity Vector

All higher derivatives of velocity might now be introduced, beginning with acceleration. But I pause here to define an angular velocity vector, which be useful to describe frames that tumble relative to each other, and which therefore yield different values for an object's acceleration. Defining an angular velocity vector is rather subtle because we must establish that any object so defined has the properties that a vector should have; that is, we require to investigate what adding angular velocities should mean.

We will define an angular velocity for two distinct scenarios: (1) a point moving in three dimensions relative to a nominated point, and (2) a frame spinning relative to a nominated frame. It can be shown that an arbitrary change of orientation can always be expressed as a single rotation [6], so a frame whose orientation changes in a complex way relative to a nominated frame can always be described as having a possibly time-dependent angular velocity relative to that nominated frame. In other words, circular motion is always sufficient to describe the motion of one frame that is perhaps tumbling chaotically relative to another.

9.1 Angular Velocity of a Point Relative to Another Point in Some Frame

Begin with the usual ideas of a point moving with velocity v (a real number) in a circle of radius r, sweeping out an angle θ as it moves. The point is defined to have a real-number angular velocity of $\omega \equiv d\theta/dt = v/r$.

Now generalise this circular motion in a plane to arbitrary motion in three dimensions. Define the angular velocity of point P relative to point A in some frame F, to be the putative vector $\boldsymbol{\omega}_{PA}^{F}$, whose length is the rate at which the line joining A to P sweeps out angle in the plane that it is momentarily moving in, and whose direction is given by the right-hand rule choice of this plane's normal. With the component of the point's velocity \boldsymbol{v}_{PA}^{F} transverse

DST-Group-TR-3309



Figure 5: The red point P follows the red curve in space. At the moment of the picture, its velocity relative to point A in this frame F is the red vector. Point P's instantaneous angular velocity $\boldsymbol{\omega}_{PA}^{F}$ relative to point A in frame F is the blue vector, and is perpendicular to the black and the red vectors. The blue vector can be placed anywhere in the diagram, not just on the point P as done here.

to r_{PA} written as v_{\perp} , the length of this vector is, using the notation "(a, b)" for the angle between a and b,

$$\left|\boldsymbol{\omega}_{PA}^{F}\right| \equiv \frac{\left|\boldsymbol{v}_{\perp}\right|}{r_{PA}} = \frac{\boldsymbol{v}_{PA}^{F}\left|\sin(\boldsymbol{r}_{PA}, \boldsymbol{v}_{PA}^{F})\right|}{r_{PA}} = \frac{\left|\boldsymbol{r}_{PA} \times \boldsymbol{v}_{PA}^{F}\right|}{r_{PA}^{2}}.$$
(9.1)

In that case, define the angular velocity of point P relative to point A in frame F as

$$\boldsymbol{\omega}_{PA}^{F} \equiv \frac{\boldsymbol{r}_{PA} \times \boldsymbol{v}_{PA}^{F}}{r_{PA}^{2}} = \hat{\boldsymbol{r}}_{PA} \times \frac{\boldsymbol{v}_{PA}^{F}}{r_{PA}}$$
(9.2)

(where the hat denotes a unit vector), because the length of this vector agrees with (9.1), and its direction correctly gives a right-handed sense for the "instantaneous rotation" of P around A. Figure 5 shows the instantaneous angular velocity relative to point A (blue arrow) of the blue point P that is moving along some arbitrary curve. The angular velocity is continuously changing in time, and at any moment it describes the extent and orientation at which P is instantaneously "orbiting" A.

Figure 6 shows the two fields of angular velocity vectors for a selection of points of a rigidly rotating cylinder, relative to each of two choices of reference point A. The blue individual angular velocity vectors have a spread of lengths and directions.

Although ω_{PA}^F is defined in (9.2) as a cross product of vectors and is therefore a vector itself, at this stage just how we might employ it as a vector is not immediately clear.⁵ For example,

⁵You will find a distinction in the literature between "polar vectors" (our arrows) and "axial vectors", also known as "pseudo vectors", that describe cross products such as angular momentum. This distinction relates to how the world looks in a mirror and the notion of right- and left-handedness. One might say that we don't use mirrors to create new frames or coordinate systems in aerospace analysis, so we needn't distinguish between polar and axial vectors. But I see this distinction as completely artificial anyway: an exercise in creating jargon for its own sake. Clearly, a right-handed cross product becomes left handed when viewed in a mirror (unless we modify the meaning of "left handed"), and introducing terminology to describe that obvious fact gives no benefit that I have ever seen. Every discussion requiring knowledge of pseudo vectors that I am aware of can be reduced to a discussion of the notion of handedness, a topic that does not require two types of vector to be defined. Pseudo vectors are sometimes said to "generalise the cross product to higher than three dimensions", but the



Figure 6: The fields of angular velocity vectors for a selection of points (in green) in horizontal slices of a rigidly rotating vertical cylinder, relative to a choice of two reference points A, marked in each plot as a red star. The cylinder's bulk rotation is indicated by the magenta shaded arrows, with its axis of rotation the vertical magenta dashed line. Left: Point A lies at the centre of the picture on the vertical rotation axis. The vector field is cylindrically symmetric around the rotation axis. Right: Point A lies in the central plane of the picture, but now to the right of the vertical rotation axis. The lengths of the vectors in the right-hand plot have been increased for clarity. Note that these two plots describe the same object rotating. Their details are not important; the main point to note is that the field of angular velocity vectors is not simple.

 $\omega_{PA}^F + \omega_{AB}^F$ does not generally equal ω_{PB}^F . (This is easily seen by calculating $\omega_{PA}^F + \omega_{AB}^F$ and ω_{PB}^F using (9.2) for a very simple scenario.) Perhaps the main use of definition (9.2) is that it serves to introduce the concept of angular momentum, which appears in Newton's laws. Staying within the realm of classical mechanics, define the angular momentum of particle P (of mass m) relative to a point A in frame F as

$$\boldsymbol{L}_{PA}^{F} \equiv \boldsymbol{r}_{PA} \times m\boldsymbol{v}_{PA}^{F} = mr_{PA}^{2} \,\boldsymbol{\omega}_{PA}^{F} \,. \tag{9.3}$$

Now define the total angular momentum of a (not necessarily rigid) body B of point masses m_i relative to point A in frame F to be the sum of their individual angular momenta:

$$\boldsymbol{L}_{BA}^{F} \equiv \sum_{i} \boldsymbol{L}_{iA}^{F} = \sum_{i} \boldsymbol{r}_{iA} \times m_{i} \boldsymbol{v}_{iA}^{F} = \sum_{i} m_{i} r_{iA}^{2} \boldsymbol{\omega}_{iA}^{F}.$$
(9.4)

Some algebraic manipulation shows that a body's total angular momentum can be written as the sum of an "orbital" part (the angular momentum of a particle of mass $\sum_{i} m_{i}$ at the

cross product easily generalises to higher dimensions based on the theory of determinants, without requiring any concept of a pseudo vector [7]. The study of such vector behaviour is in fact more readily handled using "multivectors", which are geometrical objects built from vectors; for example, a planar area combined with a direction is a "bivector".



Figure 7: The vectors used in deriving (9.8). \mathbf{r}_{PA} and \mathbf{v}_{PA}^{F} are at right angles to each other.

particles' centre of mass, relative to A in frame F), and a "spin" part (the total angular momentum of all of the particles relative to their centre of mass, in frame F):

$$\boldsymbol{L}_{BA}^{F} = \underbrace{\boldsymbol{L}_{MA}^{F}}_{\text{orbital part, where } M \text{ is}}_{\substack{\text{a particle with total}\\\text{mass of body at body's}}} + \underbrace{\sum_{i} \boldsymbol{L}_{i,\text{CM}}^{F}}_{\text{spin part, where CM is}}$$
(9.5)

The sum (9.4) will be easy to evaluate when all of the ω_{iA}^F are parallel. This idea becomes more tractable for a rigid body, which is where the concept of total angular momentum finds its main use. Consider for a moment a rigid body rotating right-handed about a unit vector \boldsymbol{n} in frame F. Denote by p the point on the rotation axis closest to a given body point P. What is the value of ω_{Pp}^F ?

$$\boldsymbol{\omega}_{Pp}^{F} = \left|\boldsymbol{\omega}_{Pp}^{F}\right| \boldsymbol{n} = \left| \boldsymbol{\widehat{r}}_{Pp} \times \frac{\boldsymbol{v}_{Pp}^{F}}{r_{Pp}} \right| \boldsymbol{n} = \frac{\boldsymbol{v}_{Pp}^{F}}{r_{Pp}} \boldsymbol{n}.$$
(9.6)

But for circular motion the speed of rotation is proportional to distance from the axis, so it must be that v_{Pp}^F/r_{Pp} is a constant independent of the location of P. Hence ω_{Pp}^F is independent of P. So call it simply Ω^F :

"angular velocity of rigid body" =
$$\mathbf{\Omega}^F \equiv \boldsymbol{\omega}_{Pp}^F$$
 for any P . (9.7)

We see that the length of Ω^F is v_{Pp}^F/r_{Pp} , the angle turned through by the body in unit time, and its direction is the rotation axis in a right-handed sense.

Moment of Inertia

Refer to Figure 7, which shows a point P (blue dot) in orbit around an axis on which is a point A. The velocity of P relative to A is given by

DST-Group-TR-3309

$$\boldsymbol{v}_{PA}^F = \boldsymbol{\Omega}^F \times \boldsymbol{r}_{PA} \,, \tag{9.8}$$

which can be shown in the following way. First, the figure shows that $\Omega^F \times \mathbf{r}_{PA}$ has the correct direction: tangential to the rotation in the direction of motion. Second, the length of $\Omega^F \times \mathbf{r}_{PA}$ is

$$\begin{aligned} \left| \mathbf{\Omega}^{F} \times \mathbf{r}_{PA} \right| &= \left| \mathbf{\Omega}^{F} \right| \left| \mathbf{r}_{PA} \right| \sin \theta \\ &= \left| \mathbf{\Omega}^{F} \right| r_{\perp} = v_{PA} \,, \end{aligned} \tag{9.9}$$

as required. So in the rigid body of the previous discussion, the velocity of any point P relative to a point A where A is located on the body's rotation axis (and not necessarily in the plane of P's rotation) is given by (9.8). This means that when A is on the rotation axis, the angular velocity relative to A of any point in the body is

$$\boldsymbol{\omega}_{PA}^{F} \stackrel{(\underline{9.2})}{=} \frac{\boldsymbol{r}_{PA} \times \boldsymbol{v}_{PA}^{F}}{r_{PA}^{2}} \stackrel{(\underline{9.8})}{=} \frac{\boldsymbol{r}_{PA} \times \left(\boldsymbol{\Omega}^{F} \times \boldsymbol{r}_{PA}\right)}{r_{PA}^{2}} \quad (A \text{ on rotation axis})$$
$$= \frac{-\boldsymbol{r}_{PA} \times \left(\boldsymbol{r}_{PA} \times \boldsymbol{\Omega}^{F}\right)}{r_{PA}^{2}} \quad (A \text{ on rotation axis}). \tag{9.10}$$

The cross product is used so frequently in rotational theory that I simplify its notation by defining a new object v^{\times} for any vector v, such that for any vector a,

$$\boldsymbol{v}^{\times}\boldsymbol{a} \equiv \boldsymbol{v} \times \boldsymbol{a} \,. \tag{9.11}$$

We can coordinatise \boldsymbol{v}^{\times} to produce $[\boldsymbol{v}^{\times}]_S$, or equivalently $[\boldsymbol{v}]_S^{\times}$, by defining $[\boldsymbol{v}^{\times}]_S [\boldsymbol{a}]_S$ to equal $[\boldsymbol{v} \times \boldsymbol{a}]_S$. That is,

$$[\boldsymbol{v}^{\times}]_{S} [\boldsymbol{a}]_{S} \equiv [\boldsymbol{v} \times \boldsymbol{a}]_{S} = [\boldsymbol{v}]_{S} \times [\boldsymbol{a}]_{S} = \begin{bmatrix} v_{x} \\ v_{y} \\ v_{z} \end{bmatrix} \times \begin{bmatrix} a_{x} \\ a_{y} \\ a_{z} \end{bmatrix} = \begin{bmatrix} v_{y}a_{z} - v_{z}a_{y} \\ v_{z}a_{x} - v_{x}a_{z} \\ v_{x}a_{y} - v_{y}a_{x} \end{bmatrix}$$
$$= \begin{bmatrix} 0 & -v_{z} & v_{y} \\ v_{z} & 0 & -v_{x} \\ -v_{y} & v_{x} & 0 \end{bmatrix} \begin{bmatrix} a_{x} \\ a_{y} \\ a_{z} \end{bmatrix},$$
(9.12)

from which it follows that

$$[\boldsymbol{v}^{\times}]_{S} \equiv [\boldsymbol{v}]_{S}^{\times} = \begin{bmatrix} v_{x} \\ v_{y} \\ v_{z} \end{bmatrix}^{\times} \equiv \begin{bmatrix} 0 & -v_{z} & v_{y} \\ v_{z} & 0 & -v_{x} \\ -v_{y} & v_{x} & 0 \end{bmatrix}.$$
(9.13)

As a side comment, due to its central role in rotational theory the cross-product matrix also appears frequently in the quantum mechanics of angular momentum, where the following matrices appear:

$$J_x \equiv i\hbar \begin{bmatrix} 1\\0\\0 \end{bmatrix}^{\times}, \quad J_y \equiv i\hbar \begin{bmatrix} 0\\1\\0 \end{bmatrix}^{\times}, \quad J_z \equiv i\hbar \begin{bmatrix} 0\\0\\1 \end{bmatrix}^{\times}, \quad (9.14)$$

where \hbar is a constant. These matrices form one set of solutions to the equation $J_x J_y - J_y J_x = i\hbar J_z$ and its cyclic permutations [8].

Because v^{\times} combines with a vector to produce another vector, v^{\times} is a tensor. With this notation, (9.10) is written more compactly as⁶

$$\boldsymbol{\omega}_{PA}^{F} = \frac{-\boldsymbol{r}_{PA}^{\times 2} \,\boldsymbol{\Omega}^{F}}{r_{PA}^{2}} = -\hat{\boldsymbol{r}}_{PA}^{\times 2} \,\boldsymbol{\Omega}^{F} \quad (A \text{ on rotation axis}), \tag{9.16}$$

where $\hat{\boldsymbol{r}}$ is the unit vector \boldsymbol{r}/r .

The angular momentum of a point mass P (of mass m) relative to point A on the rotation axis and in frame F is now

$$\boldsymbol{L}_{PA}^{F} \stackrel{(9.3)}{=} mr_{PA}^{2} \boldsymbol{\omega}_{PA}^{F} \stackrel{(9.16)}{=} -m\boldsymbol{r}_{PA}^{\times 2} \boldsymbol{\Omega}^{F} \equiv I_{PA} \boldsymbol{\Omega}^{F}, \qquad (9.17)$$

where I_{PA} is the particle's moment of inertia relative to A. (Note that the moment of inertia is defined relative to a point, not relative to an axis as sometimes thought: see the end of this section for a further comment.) Recalling (9.4), when several particles are present forming a (not necessarily rigid) body B, their total angular momentum relative to on-axis point A in frame F is

$$\boldsymbol{L}_{BA}^{F} = \sum_{i} -m_{i}\boldsymbol{r}_{iA}^{\times 2} \boldsymbol{\Omega}^{F} = \left(\sum_{i} -m_{i}\boldsymbol{r}_{iA}^{\times 2}\right) \boldsymbol{\Omega}^{F}.$$

$$\equiv I_{BA}, \text{ moment of inertia}$$
of whole body *B* relative to
on-axis point *A*

$$(9.18)$$

The expression for a body B's moment of inertia I_{BA} relative to on-axis point A is neatly written using the cross notation in (9.18), and we see that I_{BA} is a tensor. As usual, an expression such as (9.18) converts to coordinates S as

$$\left[\boldsymbol{L}_{BA}^{F}\right]_{S} = \left[I_{BA}\right]_{S} \left[\boldsymbol{\Omega}^{F}\right]_{S}, \quad (A \text{ on axis})$$

$$(9.19)$$

where

$$[I_{BA}]_S = \sum_i -m_i \left[\boldsymbol{r}_{iA}^{\times} \right]_S^2 \quad (A \text{ on axis}).$$

$$(9.20)$$

Because a "cross matrix" is skew symmetric (i.e. transposing it changes its sign), its square is symmetric. Hence the moment of inertia matrix $[I_{BA}]_S$ is symmetric.

Equation (9.20) is a useful, compact, and unconventional expression for the moment of inertia of a body relative to a point A situated on the rotation axis. We can check that it reduces to a recognisable expression for a very simple, symmetrical case: that of a particle at a distance r from the origin O, orbiting the z axis in the xy plane; note that the origin is on the z axis, as required for (9.19) and (9.20) to apply. Referring to these last two equations, set the cartesian axes S to be at rest in the F frame. The particle's moment of inertia is not

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{a}^{\times} (\boldsymbol{b}^{\times} \boldsymbol{c}) = (\boldsymbol{a}^{\times} \boldsymbol{b}^{\times}) \boldsymbol{c}$$

and $(\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c} = (\boldsymbol{a}^{\times} \boldsymbol{b})^{\times} \boldsymbol{c}$. (9.15)

⁶The cross product is of course not associative, but use of the cross-product tensor v^{\times} in fact *is* associative. If this seems counterintuitive, compare and contrast the expressions

DST-Group-TR-3309

determined by its axis of rotation, but with an eye to eventually considering motion in the xy plane, set the particle to lie at (x, y, 0). Its moment of inertia relative to the origin equals

$$[I_{mO}]_S = -m \begin{bmatrix} \mathbf{r}_{mO}^{\times} \end{bmatrix}_S^2 = -m \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}^{\times 2} = -m \begin{bmatrix} 0 & 0 & y \\ 0 & 0 & -x \\ -y & x & 0 \end{bmatrix}^2 = m \begin{bmatrix} y^2 & -xy & 0 \\ -xy & x^2 & 0 \\ 0 & 0 & r^2 \end{bmatrix}.$$
 (9.21)

[Although I have set z = 0 for the particle to save some computation, (9.21) is valid for rotation about any axis passing through the origin: even if that axis is not the z axis, (9.21) will give the instantaneous angular momentum via (9.19) for any rotation when the particle passes through the xy plane.] Now choose the orbit to have angular rate ω right-handed around the z axis. Equation (9.19) gives

$$\begin{bmatrix} \boldsymbol{L}_{mO}^{F} \end{bmatrix}_{S} = m \begin{bmatrix} y^{2} & -xy & 0 \\ -xy & x^{2} & 0 \\ 0 & 0 & r^{2} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \omega \end{bmatrix} = mr^{2}\omega \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$
 (9.22)

This is the expected well-known result: for this highly symmetrical case, the angular momentum is parallel to the angular velocity, and equals mr^2 times the angular velocity.⁷

For an advanced example, we calculate the moment of inertia of a two-particle body B relative to some arbitrarily chosen point A. The body B is composed of a particle with mass 1 at position (1, 2, 3) and a particle with mass 2 at position (4, 5, 6). What is B's moment of inertia relative to the point A = (-1, 4, 2) when this two-particle body is spinning rigidly on an axis passing through A? (Remember that these equations have been derived only for the interesting case of A lying on the rotation axis.) Equation (9.20) gives

$$[I_{BA}]_{S} = -\left(\begin{bmatrix} 1\\2\\3 \end{bmatrix} - \begin{bmatrix} -1\\4\\2 \end{bmatrix} \right)^{\times 2} - 2\left(\begin{bmatrix} 4\\5\\6 \end{bmatrix} - \begin{bmatrix} -1\\4\\2 \end{bmatrix} \right)^{\times 2}$$
$$= -\begin{bmatrix} 2\\-2\\1 \end{bmatrix}^{\times 2} - 2\begin{bmatrix} 5\\1\\4 \end{bmatrix}^{\times 2} = -\begin{bmatrix} 0 & -1 & -2\\1 & 0 & -2\\2 & 2 & 0 \end{bmatrix}^{2} - 2\begin{bmatrix} 0 & -4 & 1\\4 & 0 & -5\\-1 & 5 & 0 \end{bmatrix}^{2}$$
$$= \begin{bmatrix} 39 & -6 & -42\\-6 & 87 & -6\\-42 & -6 & 60 \end{bmatrix}.$$
(9.23)

This matrix is only valid at the moment when the particles occupy the stated positions. An instant later they will have moved, and a new $[I_{BA}]_S$ must be calculated.

Equation (9.19) shows that the total angular momentum of this "rigid" two-particle body will not generally be parallel to its angular velocity. In an engineering context, the fact that these two vectors are not parallel is associated with a sideways torque on the rotation axis that stresses the bearings holding the spinning body, even to the point of rapid bearing failure. But even the most asymmetrical body's moment of inertia I_{BA} will have three eigenvectors, so that the body's total angular momentum when spinning around an axis defined by one of

⁷The moment of inertia of a point mass m is sometimes described as the number mr^2 instead of the matrix in (9.21). We see here the role of mr^2 in the wider scheme. It's also wise to recognise that our particle's moment of inertia is *not* equal to mr^2 times the identity matrix.

these eigenvectors passing through A will be parallel to its angular velocity. Here the body spins smoothly without stressing its bearings, which is of course the most desirable situation for engineers. It is perhaps a surprising result that even the most asymmetrical body always has these three axes about which it will spin smoothly—and it turns out that these axes are always orthogonal.

The moment of inertia is often mistakenly thought to be defined relative to an axis. Moment of inertia is defined relative to a point, not an axis; its value does not depend on any choice of axis. Recalling (9.3), angular momentum " $\mathbf{r} \times m\mathbf{v}$ " is defined relative to the point from which \mathbf{r} emanates, so if the " $\mathbf{L} = I\mathbf{\Omega}$ " expression for the angular momentum \mathbf{L} of a rigid body is to hold, then the moment of inertia I must be defined relative to the same point, since the angular velocity $\mathbf{\Omega}$ can hold no information about that point. Alternatively, suppose that the moment of inertia was indeed defined relative to an axis. In that case, wouldn't we be better off just multiplying in advance the 3×3 matrix in (9.23) by that axis expressed as a unit vector, to arrive at a more portable 3×1 column instead of a matrix? Three numbers would be simpler to handle than nine, and then, for a given body with scalar spin rate Ω , we would then only need to multiply this column by Ω . But, of course, this is not done in practice precisely because the moment of inertia is *not* defined relative to an axis.

The widespread belief that I refers to a specific axis probably arose because the eigenvectors of I tend to define the axes of spin for commonly used bodies in engineering such as wheels. But even here a reference point must still be specified, because for example a wheel spinning at the end of an axle presents different dynamics to a wheel spinning in the middle of that same axle. See page 35 for a related comment about torque.

9.2 Angular Velocity of a Frame Relative to Another Frame

The above discussion showed that a single angular velocity Ω^F could be defined for the rigid body as a whole, and the angular velocity ω_{PA}^F in frame F of any point P relative to an *on-rotation-axis* point A was then given by applying (9.16) to Ω^F . This suggests a second type of angular velocity: that of one frame relative to another, which will be useful in the next section. If frame F spins within frame G, we calculate Ω^G for the lattice of frame F, and call the result the angular velocity of frame F relative to frame G: Ω^{FG} .

But in what sense is this "vector" a real vector? Can two such vectors be added? It turns out that they can, so that for example

$$\mathbf{\Omega}^{AB} + \mathbf{\Omega}^{BC} = \mathbf{\Omega}^{AC} \,. \tag{9.24}$$

This is proved in Section 6 of [4]. It rests on the fact that rotating a vector by $d\alpha$ around unit-vector \boldsymbol{a} is effected by pre-multiplying the vector by $1 + d\alpha \boldsymbol{a}^{\times}$; so combining two infinitesimal rotations is equivalent to applying a single rotation of $1 + (d\alpha \boldsymbol{a} + d\beta \boldsymbol{b})^{\times}$, which involves the sum of $d\alpha \boldsymbol{a}$ and $d\beta \boldsymbol{b}$. That this process happens in a time dt is equivalent to adding angular velocity vectors. This also proves that infinitesimal rotations commute: that is, turning an object through an infinitesimal angle $d\alpha$ around one axis, then through an infinitesimal angle $d\beta$ around a possibly different axis, gives the same result as swapping those two rotations. In combining two angular velocities, we consider a body as undergoing two infinitesimal rotations in time dt. We could never represent angular velocity as a vector if the two infinitesimal rotations did not commute—because vector addition is commutative.

Setting C to A in (9.24) gives

$$\mathbf{\Omega}^{AB} = -\mathbf{\Omega}^{BA}, \qquad (9.25)$$

DST-Group-TR-3309

which makes sense intuitively from the definition of Ω^{AB} .

One important instance of the angular velocity occurs for an aircraft undergoing roll/pitch/yaw motion. Convention sets the aircraft's x axis to be its roll axis (forward through its nose), y to be its pitch axis (out along the starboard wing) and z to be its yaw axis (the below-fuselage direction). The changing orientation of the aircraft is comprised of a roll angular velocity, a pitch angular velocity, and a yaw angular velocity. Because these angular velocities add as vectors, the aircraft's (B) overall angular velocity relative to the world W is their sum:

$$\mathbf{\Omega}^{BW} = \text{roll vector} + \text{ pitch vector} + \text{ yaw vector.}$$
(9.26)

The roll, pitch, and yaw rates are conventionally labelled p, q, r respectively, so

$$\left[\mathbf{\Omega}^{BW}\right]_{B} = [\operatorname{roll}]_{B} + [\operatorname{pitch}]_{B} + [\operatorname{yaw}]_{B} = \begin{bmatrix} p\\0\\0 \end{bmatrix} + \begin{bmatrix} 0\\q\\0 \end{bmatrix} + \begin{bmatrix} 0\\q\\0 \end{bmatrix} + \begin{bmatrix} 0\\0\\r \end{bmatrix} = \begin{bmatrix} p\\q\\r \end{bmatrix}.$$
(9.27)

This expression is used frequently in 6-degree-of-freedom analyses of aircraft.

All frames agree that the angular velocity of frame F relative to frame G is the vector $\mathbf{\Omega}^{FG}$, so there is no need to specify which frame is "doing the observing". This agreement can be seen as follows. Imagine F and G represented by two boxes in contact along a common face. All frames agree on the physical reality of the axis of relative rotation of the boxes. As box Frotates relative to box G, it scratches grooves in the surface of box G that have a physical reality. All frames agree on the length of any groove gouged out in a given time interval. Hence all agree on the length and direction of $\mathbf{\Omega}^{FG}$.

10 Coordinates and the Time Derivative

To calculate time derivatives such as velocity and acceleration, we naturally prefer to manipulate coordinates (coordinate vectors) rather than draw and manipulate arrows (proper vectors). But combining the processes of differentiating and coordinatising demands careful inspection. The "coordinates of a time derivative" and the "time derivative of coordinates" can easily be confused, but these two quantities are generally not the same. Start with a proper vector \boldsymbol{v} that varies with time, and consider:

Coordinates of a time derivative: Calculate the F-frame time derivative of v, then express this in coordinates S:

proper vector =
$$\boldsymbol{v}$$
;
its time derivative in $F = \frac{\mathrm{d}^F \boldsymbol{v}}{\mathrm{d}t}$;
S-coordinates of time derivative = $\left[\frac{\mathrm{d}^F \boldsymbol{v}}{\mathrm{d}t}\right]_S$. (10.1)

Time derivative of coordinates: Now reverse the order of differentiating and coordinatising:

proper vector =
$$\boldsymbol{v}$$
;
its S-coordinate vector = $[\boldsymbol{v}]_S$;
time derivative of coordinate vector = $\frac{\mathbf{d}[\boldsymbol{v}]_S}{\mathbf{d}t}$. (10.2)

Recall that " d^{F}/dt " acts here on a proper vector (an arrow), whereas "d/dt" differentiates each element of a coordinate vector (a column of coordinates).

The two columns of numbers $[d^F \boldsymbol{v}/dt]_S$ and $d[\boldsymbol{v}]_S/dt$ are generally different. To see how they might be related, write a vector \boldsymbol{v} as a linear combination of basis vectors, where the following xyz coordinates needn't be cartesian:

$$\boldsymbol{v} = v_x \boldsymbol{e}_x + v_y \boldsymbol{e}_y + v_z \boldsymbol{e}_z \,. \tag{10.3}$$

Differentiate this in some frame F, writing e.g. \dot{v}_x to mean $d(v_x)/dt$:

$$\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} = \dot{v}_{x}\boldsymbol{e}_{x} + v_{x}\frac{\mathrm{d}^{F}\boldsymbol{e}_{x}}{\mathrm{d}t} + \dot{v}_{y}\boldsymbol{e}_{y} + v_{y}\frac{\mathrm{d}^{F}\boldsymbol{e}_{y}}{\mathrm{d}t} + \dot{v}_{z}\boldsymbol{e}_{z} + v_{z}\frac{\mathrm{d}^{F}\boldsymbol{e}_{z}}{\mathrm{d}t}.$$
(10.4)

If the basis vectors $\{e_x, e_y, e_z\}$ don't change with time in frame F, then (10.4) reduces to

$$\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} = \dot{v}_{x}\boldsymbol{e}_{x} + \dot{v}_{y}\boldsymbol{e}_{y} + \dot{v}_{z}\boldsymbol{e}_{z}, \qquad (10.5)$$

and if this basis set that doesn't change in F is called f, then

$$\begin{bmatrix} \frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} \end{bmatrix}_{f} = \begin{bmatrix} \dot{v}_{x} \\ \dot{v}_{y} \\ \dot{v}_{z} \end{bmatrix} = \frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} v_{x} \\ v_{y} \\ v_{z} \end{bmatrix} = \frac{\mathrm{d}[\boldsymbol{v}]_{f}}{\mathrm{d}t}.$$
 (10.6)

So for this special coordinate system f that is at rest in frame F,

$$\left[\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t}\right]_{f} = \frac{\mathrm{d}[\boldsymbol{v}]_{f}}{\mathrm{d}t}.$$
(10.7)

Equation (10.7) shows that the coordinates of the derivative do equal the derivative of the coordinates for the special case when the coordinate basis vectors don't change with time in the chosen frame. In the following section I take the next step of expressing $d^F \boldsymbol{v}/dt$ in arbitrary coordinates S.

10.1 Three Forms of the Time Derivative of a Vector

Here we derive the following three forms of the S-coordinates of the F-frame time derivative of an arbitrary vector \boldsymbol{v} :

$$\left[\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t}\right]_{S} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + \mu_{S}^{f}\dot{\mu}_{f}^{S}\right)[\boldsymbol{v}]_{S} \quad \text{(orientation-matrix form)}$$
(10.8a)

$$= \left(\frac{\mathrm{d}}{\mathrm{d}t} + \left[\mathbf{\Omega}^{SF}\right]_{S}^{\times}\right) [\boldsymbol{v}]_{S} \quad (\text{angular-velocity form}) \tag{10.8b}$$

$$= \left(\frac{\mathrm{d}}{\mathrm{d}t} + \Gamma_S^F\right) [\boldsymbol{v}]_S \quad \text{(basis-vector form)}. \tag{10.8c}$$

The second term in each of the parentheses, $\mu_S^f \dot{\mu}_f^S$, $[\mathbf{\Omega}^{SF}]_S^{\times}$, and Γ_S^F , when multiplied by $[\boldsymbol{v}]_S$, each quantify the extent to which "differentiating \boldsymbol{v} " and "coordinatising \boldsymbol{v} " do not commute.

DST-Group-TR-3309

10.1.1 Direction-Cosine Form (10.8a)

Equation (10.8a) is easily derived for orientation matrices in the following way. Recall (10.7), which expresses an F-frame time derivative in a special set of coordinates f, whose basis vectors don't change with time in F. How might we express $d^F v/dt$ in arbitrary coordinates S—which might even be non-cartesian? Begin with

$$\left[\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t}\right]_{S} = \mu_{S}^{f} \left[\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t}\right]_{f} = \mu_{S}^{f} \frac{\mathrm{d}[\boldsymbol{v}]_{f}}{\mathrm{d}t}$$
(10.9)

for arbitrary v. We now have an ordinary derivative d/dt—but it has come at the price of a switch to f-coordinates. So convert the last coordinate vector back to S-coordinates, writing

$$\left[\frac{\mathrm{d}^{F} \boldsymbol{v}}{\mathrm{d}t} \right]_{S} = \mu_{S}^{f} \frac{\mathrm{d}}{\mathrm{d}t} \left(\mu_{f}^{S} \left[\boldsymbol{v} \right]_{S} \right) = \mu_{S}^{f} \left(\dot{\mu}_{f}^{S} \left[\boldsymbol{v} \right]_{S} + \mu_{f}^{S} \frac{\mathrm{d} \left[\boldsymbol{v} \right]_{S}}{\mathrm{d}t} \right)$$
$$= \mu_{S}^{f} \dot{\mu}_{f}^{S} \left[\boldsymbol{v} \right]_{S} + \frac{\mathrm{d} \left[\boldsymbol{v} \right]_{S}}{\mathrm{d}t} ,$$
(10.10)

so that we arrive at (10.8a):

$$\left[\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t}\right]_{S} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + \mu_{S}^{f}\dot{\mu}_{f}^{S}\right)[\boldsymbol{v}]_{S}.$$
(10.11)

At first sight it might appear anomalous that a specific choice of coordinates f for frame F appears on the right-hand side of (10.11) but not the left-hand side. We might infer that for two choices of coordinates f, f', whose basis vectors don't change with time in F, that

$$\mu_S^f \dot{\mu}_f^S = \mu_S^{f'} \dot{\mu}_{f'}^S \,. \tag{10.12}$$

This is certainly true. Prove (10.12) by first referring to (7.10) to write

$$\mu_f^S = \mu_f^{f'} \mu_{f'}^S \,. \tag{10.13}$$

Time-differentiate both sides:

$$\dot{\mu}_{f}^{S} = \dot{\mu}_{f}^{f'} \mu_{f'}^{S} + \mu_{f}^{f'} \dot{\mu}_{f'}^{S} .$$
(10.14)

Because basis sets f and f' are fixed relative to each other, their relative orientation $\mu_f^{f'}$ doesn't depend on time, so $\dot{\mu}_f^{f'} = 0$ and (10.14) becomes

$$\dot{\mu}_f^S = \mu_f^{f'} \dot{\mu}_{f'}^S \,. \tag{10.15}$$

Now pre-multiply both sides of (10.15) by μ_S^f , using (7.10) again:

$$\mu_{S}^{f} \dot{\mu}_{f}^{S} = \mu_{S}^{f} \mu_{f}^{f'} \dot{\mu}_{f'}^{S} = \mu_{S}^{f'} \dot{\mu}_{f'}^{S}, \qquad (10.16)$$

which is just (10.12), as required.

Another useful expression involving $\dot{\mu}$ is formed by differentiating the expression $\mu_B^A \mu_A^B = 1$ with respect to time, to arrive at

$$\dot{\mu}_B^A \,\mu_A^B = -\mu_B^A \,\dot{\mu}_A^B \,. \tag{10.17}$$

10.1.2 Angular Velocity Form (10.8b)

Equation (10.8b) can be derived by scrutinising the picture of \boldsymbol{v} evolving. Place ourselves in frame F and watch the coordinate system S rotating around an axis whose direction might also be continuously changing. An observer attached to S defines a frame S, which rotates relative to frame F with some angular velocity $\boldsymbol{\Omega}^{SF}$. In a time dt, we in frame F watch how \boldsymbol{v} evolves. We notice that an infinitesimal increment in \boldsymbol{v} can be considered as the sum of two steps:

Increment of \boldsymbol{v} in F = Increment of \boldsymbol{v} in S

+ increment provided by rotation of S in F. (10.18)

By virtue of its rotation alone, \boldsymbol{v} will increment by $\boldsymbol{\Omega}^{SF} dt \times \boldsymbol{v}$: this follows from (9.8). Writing the increment of \boldsymbol{v} in frame F as $d^F \boldsymbol{v}$ and similarly for frame S, (10.18) becomes

$$d^{F}\boldsymbol{v} = d^{S}\boldsymbol{v} + \boldsymbol{\Omega}^{SF} dt \times \boldsymbol{v}, \qquad (10.19)$$

so that

$$\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} = \frac{\mathrm{d}^{S}\boldsymbol{v}}{\mathrm{d}t} + \boldsymbol{\Omega}^{SF} \times \boldsymbol{v} \,. \tag{10.20}$$

Coordinatising in S, and then recalling (10.7) as it applies to S, converts (10.20) to

$$\begin{bmatrix} \frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} \end{bmatrix}_{S} = \frac{\mathrm{d}[\boldsymbol{v}]_{S}}{\mathrm{d}t} + \left[\boldsymbol{\Omega}^{SF} \times \boldsymbol{v}\right]_{S} = \frac{\mathrm{d}[\boldsymbol{v}]_{S}}{\mathrm{d}t} + \left[\boldsymbol{\Omega}^{SF}\right]_{S}^{\times} [\boldsymbol{v}]_{S}$$
$$= \left(\frac{\mathrm{d}}{\mathrm{d}t} + \left[\boldsymbol{\Omega}^{SF}\right]_{S}^{\times}\right) [\boldsymbol{v}]_{S},$$
(10.21)

which is (10.8b).

Equations (10.8a) and (10.8b) together yield

$$\mu_S^f \dot{\mu}_f^S = \left[\mathbf{\Omega}^{SF} \right]_S^{\times}, \tag{10.22}$$

which is useful because it relates the slightly mysterious-looking matrix $\mu_S^f \dot{\mu}_f^S$ to the very concrete notion of the instantaneous angular velocity of frame S relate to frame F, expressed in S-coordinates. It trivially produces the standard expression for how the orientation matrix evolves:

$$\dot{\mu}_f^S = \mu_f^S \left[\mathbf{\Omega}^{SF} \right]_S^{\times}. \tag{10.23}$$

This important equation appears in books on aerospace dynamics, such as [9] and [10].

10.1.3 Changing Basis-Vector Form (10.8c)

Suppose we write a vector v as a linear combination of the basis vectors $\{e_x, e_y, e_z\}$, which needn't be orthonormal; in other words, the xyz coordinates needn't be the usual cartesian sort. In this section we use a language of components that is common in tensor analysis:

$$\boldsymbol{v} = v^{x}\boldsymbol{e}_{x} + v^{y}\boldsymbol{e}_{y} + v^{z}\boldsymbol{e}_{z} \equiv v^{\alpha}\boldsymbol{e}_{\alpha}, \qquad (10.24)$$

where summation over any repeated up-down index (here α) is assumed. Now write

$$\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} = \frac{\mathrm{d}^{F}(v^{\alpha}\boldsymbol{e}_{\alpha})}{\mathrm{d}t}.$$
(10.25)

DST-Group-TR-3309

As usual, d^F/dt simply means "differentiate with respect to time, remembering that we are in frame F", so the usual product rule of differentiation applies:

$$\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} = \frac{\mathrm{d}v^{\alpha}}{\mathrm{d}t}\boldsymbol{e}_{\alpha} + v^{\alpha}\frac{\mathrm{d}^{F}\boldsymbol{e}_{\alpha}}{\mathrm{d}t},\tag{10.26}$$

where dv^{α}/dt denotes the time derivative of v^{α} (a number), not the α -component of the time derivative of v or v (which would require a "d^F", not just a "d").

Now, $d^F e_{\alpha}/dt$ is another proper vector, so it too can be written as a linear combination of the S-basis vectors as

$$\frac{\mathrm{d}^{F}\boldsymbol{e}_{\alpha}}{\mathrm{d}t} = \Gamma_{S\alpha}^{Fx} \boldsymbol{e}_{x} + \Gamma_{S\alpha}^{Fy} \boldsymbol{e}_{y} + \Gamma_{S\alpha}^{Fz} \boldsymbol{e}_{z} = \Gamma_{S\alpha}^{F\beta} \boldsymbol{e}_{\beta}, \qquad (10.27)$$

for some numbers $\Gamma_{S\alpha}^{Fx}$, $\Gamma_{S\alpha}^{Fy}$, $\Gamma_{S\alpha}^{Fz}$. Then (10.26) becomes

$$\frac{\mathrm{d}^{F}\boldsymbol{v}}{\mathrm{d}t} = \frac{\mathrm{d}v^{\alpha}}{\mathrm{d}t}\boldsymbol{e}_{\alpha} + v^{\alpha}\,\Gamma_{S\alpha}^{F\beta}\boldsymbol{e}_{\beta} \quad (\text{now swap dummy indices }\alpha \text{ and }\beta)$$

$$= \frac{\mathrm{d}v^{\alpha}}{\mathrm{d}t}\boldsymbol{e}_{\alpha} + v^{\beta}\,\Gamma_{S\beta}^{F\alpha}\,\boldsymbol{e}_{\alpha} = \left(\frac{\mathrm{d}v^{\alpha}}{\mathrm{d}t} + v^{\beta}\,\Gamma_{S\beta}^{F\alpha}\right)\boldsymbol{e}_{\alpha}\,. \tag{10.28}$$

In other words,

which is (10.8c).

We see here that the following quantities are identical: $\mu_S^f \dot{\mu}_f^S$ (relating to the orientation matrix and how it evolves), $[\mathbf{\Omega}^{SF}]_S^{\times}$ (relating to the instantaneous angular velocity of S in F), and Γ_S^F (relating to the way the basis vectors of S evolve in F).

Equation (10.8a) uses the orientation matrix μ_S^f that specifies how the basis vectors of f relate to those of S; (10.8b) uses Ω^{SF} , the angular velocity vector of the basis-vector set S in frame F, and (10.8c) uses component knowledge of how the S-basis vectors evolve in frame F. When the S-coordinates are set identical to the f-coordinates, it's clear that

$$\mu_{S}^{f} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{\Omega}^{SF} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Gamma_{S}^{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (10.30)$$

in which case

$$\mu_{S}^{f} \dot{\mu}_{f}^{S} = \begin{bmatrix} \mathbf{\Omega}^{SF} \end{bmatrix}_{S}^{\times} = \Gamma_{S}^{F} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$
 (10.31)

Then each of (10.8a)–(10.8c) reduces to (10.7) as expected.

The above analysis of vectors can also be extended to higher-order tensors. In particular, Appendix B derives the time derivative of a rank-2 tensor, producing (B3), the rank-2 version of (10.8b).

The signature of whether the coordinates of a time derivative equal the time derivatives of the coordinates is whether the basis vectors of the coordinate system in use are changing with time. They need not be changing at any particular point in space; rather, it's sufficient only that the basis vectors forming the continuous parade "encountered" along its track by an airborne vehicle change with time.

Zipfel [2] emphasises the distinction between the coordinates of a derivative and the derivative of coordinates by calling $d^F \boldsymbol{v}/dt$ the "rotational time derivative", written as $D^F \boldsymbol{v}$. Names aside, the primary point is that $d^F \boldsymbol{v}/dt$ is not a new kind of time derivative; it is simply the time derivative of a proper vector \boldsymbol{v} calculated in frame F. All derivatives must be calculated in *some* frame, and it's only because we are generally using more than one frame in aerospace scenarios that we must explicitly denote the frame currently in use.

Equation (10.8c) is actually just an instance of the "covariant derivative" of tensor analysis, which I rederive in more standard language in Appendix C. In tensor analysis the covariant derivative applies to all coordinates, not just time, and is nothing more than the usual partial derivative written in a way that incorporates how the basis vectors change with time and position, which enables us to work with coordinates only and omit explicit reference to those basis vectors.

Omitting basis vectors makes sense computationally in that we tend to compute with coordinates only, since these are easily written as arrays of numbers in any programming language. Nonetheless, I think that omitting basis vectors from the start (which is the traditional approach to tensor analysis) is pedagogically risky because basis vectors connect coordinate vectors to proper vectors. Tensor analysis has traditionally focussed on components only, and this is why the covariant derivative exists: to incorporate the effect of the changing basis vectors automatically and invisibly. But this very convenience continues to support the idea that a vector is *only* a set of components. Perhaps a vector was once only a set of components, or rather an aggregate of such components across all conceivable coordinate systems, possessing a particular transformation property across those coordinate systems. But this idea of an aggregate of sets of numbers belies the intuitive idea that a vector is simply an arrow, a single coordinate-independent object. This is precisely why we profit from distinguishing a coordinate vector from a proper vector; and basis vectors are the enabler here. Some authors [11] describe covariant differentiation as an operation that is *constructed* to satisfy the rules of tensor analysis, understandably seen as arcane by many. While it certainly can be constructed that way, it is really not something arbitrarily constructed in order "to work"; instead, it is simply a convenient way to write a partial derivative without explicit reference to basis vectors [12]. It emerges naturally when we differentiate a proper vector while constructing a coordinate vector.

It's evident, then, from any one of (10.8a)-(10.8c) that the S-coordinate vector of the time derivative, $[d^F \boldsymbol{v}/dt]_S$, does not generally equal the time derivative of the coordinate vector, $d[\boldsymbol{v}]_S/dt$ —nor should they generally be equal, given that no frame appears in the latter quantity. Clearly, the frame used to calculate a derivative need have nothing to do with the coordinates that might be chosen to express the vector. For example, calculate the ground velocity of an aircraft in the body coordinates of a rotating satellite. The ground velocity of the aircraft A is $\boldsymbol{v}_A^{\text{ECEF}}$ or equivalently $\boldsymbol{v}_{AB}^{\text{ECEF}}$, where B is any point at rest in the ECEF.

DST-Group-TR-3309

The satellite coordinates S of this velocity are

$$\left[\boldsymbol{v}_{AB}^{\text{ECEF}}\right]_{S} \equiv \left[\frac{\mathrm{d}^{\text{ECEF}}\boldsymbol{r}_{AB}}{\mathrm{d}t}\right]_{S} \stackrel{(\underline{10.8a})}{=} \left(\frac{\mathrm{d}}{\mathrm{d}t} + \mu_{S}^{E}\dot{\mu}_{E}^{S}\right)[\boldsymbol{r}_{AB}]_{S}, \qquad (10.32)$$

where E is any coordinates that don't change with time in the ECEF. In particular, we see here explicit reference to the changing matrix μ_E^S that quantifies the orientation of the satellite relative to Earth. Examples of calculating μ can be found in [4].

The above discussion clears up a particular confusion that exists regarding the nature of velocity. The argument goes something like this, and note that I am deliberately using vague wording within the quotes that follow in order to mimic the vagueness of the typical argument: "Consider two coordinate systems S, T, whose relation to each other varies with time, written as the time-dependent matrix μ_T^S . A position in S is related to a position in T via multiplying by μ_T^S . When we differentiate these positions to give velocities, an extra term involving $\dot{\mu}_T^S$ appears, whose presence implies that the two velocities are no longer related by simply multiplying by μ_T^S . So velocity cannot really be a vector." But velocity is a proper vector: its components do transform simply via multiplying by μ_T^S , just as for a displacement vector. Where does the argument go wrong?

The argument goes wrong because its wording is too vague to mean anything. We need only define everything carefully to see where the problem lies. The position (i.e. displacement vector) of some point A relative to some reference point B is r_{AB} , whose components are related in the two coordinate systems via

$$[\mathbf{r}_{AB}]_T = \mu_T^S \, [\mathbf{r}_{AB}]_S \,. \tag{10.33}$$

Differentiate each side to give

$$\frac{\mathrm{d}}{\mathrm{d}t}[\mathbf{r}_{AB}]_T = \dot{\mu}_T^S [\mathbf{r}_{AB}]_S + \mu_T^S \frac{\mathrm{d}}{\mathrm{d}t} [\mathbf{r}_{AB}]_S \,. \tag{10.34}$$

So far we have derivatives of coordinate vectors; but the original argument discusses "velocity", so we must rewrite (10.34) in terms of proper vectors, using (10.7). Choose a frame in which the coordinate system S doesn't change: this can be called S also. Similarly, choose a frame in which T doesn't change, to be called T. Then (10.34) becomes

$$\left[\frac{\mathrm{d}^{T}\boldsymbol{r}_{AB}}{\mathrm{d}t}\right]_{T} = \dot{\mu}_{T}^{S}\left[\boldsymbol{r}_{AB}\right]_{S} + \mu_{T}^{S}\left[\frac{\mathrm{d}^{S}\boldsymbol{r}_{AB}}{\mathrm{d}t}\right]_{S}.$$
(10.35)

In velocity notation, this is

$$\left[\boldsymbol{v}_{AB}^{T}\right]_{T} = \dot{\mu}_{T}^{S} \left[\boldsymbol{r}_{AB}\right]_{S} + \mu_{T}^{S} \left[\boldsymbol{v}_{AB}^{S}\right]_{S} . \tag{10.36}$$

Being proper vectors, velocities can be treated as arrows (just like displacement vectors), so their components transform in the usual way of (7.4). Equation (10.36) then becomes

$$\left[\boldsymbol{v}_{AB}^{T}\right]_{T} = \dot{\mu}_{T}^{S} \left[\boldsymbol{r}_{AB}\right]_{S} + \left[\boldsymbol{v}_{AB}^{S}\right]_{T} . \tag{10.37}$$

Now we see where the original vaguely worded argument in italics above has gone awry. The $\dot{\mu}$ matrix does not interfere with any notion of velocity being a proper vector; rather, $\dot{\mu}$ simply relates the *T* coordinates of two *different* velocity vectors, \boldsymbol{v}_{AB}^T and \boldsymbol{v}_{AB}^S . This underlines the importance of being clear about which frame is being used to calculate velocity, and more generally of using a clear language to describe kinematic scenarios.

11 The Acceleration Vector

An acceleration vector can be defined as

$$\boldsymbol{a}_{AB}^{FG} \equiv \frac{\mathrm{d}^{F}}{\mathrm{d}t} \frac{\mathrm{d}^{G}}{\mathrm{d}t} \boldsymbol{r}_{AB} = \frac{\mathrm{d}^{F}}{\mathrm{d}t} \boldsymbol{v}_{AB}^{G} , \qquad (11.1)$$

denoting how frame F measures the frame-G velocity vector to change. But to apply Newton's laws it suffices to consider only one frame, so we define a more restricted version of (11.1) as

$$\boldsymbol{a}_{AB}^{F} \equiv \boldsymbol{a}_{AB}^{FF} = \frac{\mathrm{d}^{F}}{\mathrm{d}t} \boldsymbol{v}_{AB}^{F} = \left(\frac{\mathrm{d}^{F}}{\mathrm{d}t}\right)^{2} \boldsymbol{r}_{AB} \equiv \frac{\mathrm{d}^{FF}}{\mathrm{d}t^{2}} \boldsymbol{r}_{AB}, \qquad (11.2)$$

where d^{FF}/dt^2 is shorthand for $(d^F/dt)^2$.

Suppose F_1 and F_2 are fixed points in frame F. Then analogously to (8.3), write

$$\boldsymbol{a}_{AF_{1}}^{F} = \frac{\mathrm{d}^{F}}{\mathrm{d}t} \boldsymbol{v}_{AF_{1}}^{F} \xrightarrow{(\boldsymbol{8}.\boldsymbol{3})} \frac{\mathrm{d}^{F}}{\mathrm{d}t} \boldsymbol{v}_{AF_{2}}^{F} = \boldsymbol{a}_{AF_{2}}^{F}, \qquad (11.3)$$

which prompts us to define the acceleration of a point in frame F as

 $\boldsymbol{a}_{A}^{F} = ext{acceleration of } A ext{ in frame } F \equiv \boldsymbol{a}_{AF_{1}}^{F} ext{ for any point } F_{1} ext{ fixed in } F.$ (11.4)

Consider now two frames, F and G. How does the F-frame acceleration of point A relative to point B compare with the G-frame acceleration of A relative to B? We can apply (10.20) provided we know the relative angular velocity of F relative to G. Suppose then, that the two frames have no relative angular velocity; it follows that $d^F \boldsymbol{v}/dt = d^G \boldsymbol{v}/dt$ for any vector \boldsymbol{v} . In that case,

$$\boldsymbol{a}_{AB}^{G} = \frac{\mathrm{d}^{GG}\boldsymbol{r}_{AB}}{\mathrm{d}t^{2}} = \frac{\mathrm{d}^{FF}\boldsymbol{r}_{AB}}{\mathrm{d}t^{2}} = \boldsymbol{a}_{AB}^{F}, \qquad (11.5)$$

so that both frames measure the same value for the acceleration of A relative to B, independently of how those frames are moving relative to each other.

Next we ask a related but different question: how does the acceleration of A in frame $F(\mathbf{a}_A^F)$ compare with the acceleration of A in frame $G(\mathbf{a}_A^G)$ when the two frames (as above) have no relative angular velocity? For this, single out a fixed point F_0 in frame F and a fixed point G_0 in frame G and write

$$\boldsymbol{a}_{A}^{F} = \boldsymbol{a}_{AF_{0}}^{F} = \frac{\mathrm{d}^{FF}\boldsymbol{r}_{AF_{0}}}{\mathrm{d}t^{2}} = \frac{\mathrm{d}^{GG}}{\mathrm{d}t^{2}} \left(\boldsymbol{r}_{AG_{0}} + \boldsymbol{r}_{G_{0}F_{0}}\right) = \boldsymbol{a}_{A}^{G} + \frac{\mathrm{d}^{GG}\boldsymbol{r}_{G_{0}F_{0}}}{\mathrm{d}t^{2}}.$$
 (11.6)

The frames' two values of the acceleration of A differ by the relative acceleration of the two frames. Suppose F and G separate by at most a constant velocity: $d^{G} r_{G_{0}F_{0}}/dt = \text{constant}$. Then (11.6) produces

$$\boldsymbol{a}_A^F = \boldsymbol{a}_A^G. \tag{11.7}$$

That is, the acceleration of a point A is the same in two frames whose relative acceleration and angular velocity are zero.

The notation here addresses all manner of similar questions. For example, how does the F-acceleration of A relative to a point fixed in the G frame $(\mathbf{a}_{AG_0}^F)$ compare with the G-acceleration of A (\mathbf{a}_A^G) ? When the frames have no relative angular velocity, these two accelerations are equal:

$$\boldsymbol{a}_{AG_0}^F = \frac{\mathrm{d}^{FF} \boldsymbol{r}_{AG_0}}{\mathrm{d}t^2} = \frac{\mathrm{d}^{GG} \boldsymbol{r}_{AG_0}}{\mathrm{d}t^2} = \boldsymbol{a}_A^G.$$
(11.8)

For any arbitrary relative linear and orientational motion of the frames, we need only return to (10.20) to calculate the details.

DST-Group-TR-3309

12 Inertial Frames

Recalling the analysis that produced (11.7), we see that if two frames F and G maintain a constant relative orientation ($\Omega^{FG} = \mathbf{0}$) and a constant relative velocity ($\mathrm{d}^F \mathbf{r}_{F_0G_0}/\mathrm{d}t = \mathrm{constant}$) then (11.7) holds: a body's acceleration in frame F will equal its acceleration in frame G.

Define an inertial frame I to be one in which a free body (one subject to no forces) has no acceleration. It follows that the free body will also have no acceleration in any other frame that maintains a constant relative orientation and constant relative velocity to I. So this second frame will also be inertial, and we arrive at the following statement: any frame that maintains a constant relative orientation and constant relative velocity to an inertial frame will be inertial as well.

Inertial frames are of especial interest in analysing a system's dynamics, because they allow Newton's laws of motion to be applied with a minimal number of forces. In particular, Newton's second law "force = mass \times acceleration" stated in an inertial frame I for a possibly time-dependent mass m is

$$\boldsymbol{F}^{I} = \frac{\mathrm{d}^{I}}{\mathrm{d}t} (m\boldsymbol{v}_{m}^{I}), \qquad (12.1)$$

where I can be *any* inertial frame.

Suppose we have a constant mass m accelerated by a force F^{I} in an inertial frame I, so that (12.1) becomes

$$\boldsymbol{F}^{I} = m\boldsymbol{a}_{m}^{I} \,. \tag{12.2}$$

We wish to work with Newton's laws in a rotating frame R, and so wish to write Newton's second law in the non-inertial frame R by equating a possibly different force \mathbf{F}^R to $m\mathbf{a}_m^R$. The question is, what is \mathbf{F}^R ? We need only calculate \mathbf{a}_m^R , which we can do either directly by converting d^R/dt to d^I/dt , or indirectly by starting with (12.2) and converting d^I/dt to d^R/dt . This second way turns out to be slightly faster than the first. For convenience set $D^A \equiv d^A/dt$, and begin with

$$\boldsymbol{a}_{m}^{I} = (\mathbf{D}^{I})^{2} \boldsymbol{r}_{mI} = (\mathbf{D}^{I})^{2} (\boldsymbol{r}_{mR} + \boldsymbol{r}_{RI}) = (\mathbf{D}^{R} + \boldsymbol{\Omega}^{RI^{\times}})^{2} \boldsymbol{r}_{mR} + \boldsymbol{a}_{R}^{I}.$$
(12.3)

Now expand the parentheses to give

$$\boldsymbol{a}_{m}^{I} = \boldsymbol{a}_{m}^{R} + \mathbf{D}^{R} \left(\boldsymbol{\Omega}^{RI} \times \boldsymbol{r}_{mR} \right) + \boldsymbol{\Omega}^{RI} \times \mathbf{D}^{R} \boldsymbol{r}_{mR} + \boldsymbol{\Omega}^{RI} \times (\boldsymbol{\Omega}^{RI} \times \boldsymbol{r}_{mR}) + \boldsymbol{a}_{R}^{I}$$
$$= \boldsymbol{a}_{m}^{R} + \left(\mathbf{D}^{R} \boldsymbol{\Omega}^{RI} \right) \times \boldsymbol{r}_{mR} + 2\boldsymbol{\Omega}^{RI} \times \boldsymbol{v}_{m}^{R} + \boldsymbol{\Omega}^{RI} \times (\boldsymbol{\Omega}^{RI} \times \boldsymbol{r}_{mR}) + \boldsymbol{a}_{R}^{I}.$$
(12.4)

Rearrange:

$$\boldsymbol{a}_{m}^{R} = \boldsymbol{F}^{I}/m \underbrace{-\left(\boldsymbol{D}^{R}\boldsymbol{\Omega}^{RI}\right) \times \boldsymbol{r}_{mR}}_{\text{angular}} \underbrace{-2\boldsymbol{\Omega}^{RI} \times \boldsymbol{v}_{m}^{R}}_{\text{Coriolis}} \underbrace{-\boldsymbol{\Omega}^{RI} \times \left(\boldsymbol{\Omega}^{RI} \times \boldsymbol{r}_{mR}\right)}_{\text{centrifugal}} \underbrace{-\boldsymbol{a}_{R}^{I}}_{\text{frame}}$$
$$\equiv \boldsymbol{F}^{R}/m \,. \tag{12.5}$$

We see here the four pseudo forces that must be included in the rotating-frame analysis: a term measuring the relative angular acceleration of the two frames, Coriolis and centrifugal terms, and a term for the relative linear acceleration of the two frames. These forces define \mathbf{F}^{R} , the frame-R version of the inertial-frame force \mathbf{F}^{I} .

13 Torque and Angular Momentum

The torque τ_{PA}^F in frame F on a particle P relative to some point A is defined as the frame-F rate of increase of the frame-F angular momentum of P relative to A:

$$\boldsymbol{\tau}_{PA}^{F} \equiv \frac{\mathrm{d}^{F} \boldsymbol{L}_{PA}^{F}}{\mathrm{d}t} \,. \tag{13.1}$$

Applying (9.3) gives

$$\boldsymbol{\tau}_{PA}^{F} = \frac{\mathrm{d}^{F}}{\mathrm{d}t} \left(\boldsymbol{r}_{PA} \times m\boldsymbol{v}_{PA}^{F} \right)$$
$$= \underbrace{\boldsymbol{v}_{PA}^{F} \times m\boldsymbol{v}_{PA}^{F}}_{\mathrm{the zero vector}} + \boldsymbol{r}_{PA} \times \frac{\mathrm{d}^{F}}{\mathrm{d}t} m\boldsymbol{v}_{PA}^{F}. \tag{13.2}$$

Select a point F_0 at rest in frame F, so that

$$\boldsymbol{\tau}_{PA}^{F} = \boldsymbol{r}_{PA} \times \frac{\mathrm{d}^{F}}{\mathrm{d}t} m \left(\boldsymbol{v}_{PF_{0}}^{F} + \boldsymbol{v}_{F_{0}A}^{F} \right)$$
$$= \boldsymbol{r}_{PA} \times \boldsymbol{f}_{P}^{F} + \boldsymbol{r}_{PA} \times \frac{\mathrm{d}^{F}}{\mathrm{d}t} m \boldsymbol{v}_{F_{0}A}^{F}, \qquad (13.3)$$

where \mathbf{f}_P^F is the force on the particle in frame F. Then provided that point A has at most a constant velocity in frame F, the last term above vanishes, and we have

$$\boldsymbol{\tau}_{PA}^F = \boldsymbol{r}_{PA} \times \boldsymbol{f}_P^F \,. \tag{13.4}$$

When many particles are present that form a body B (not necessarily rigid), a short calculation gives the rate of change of the body's total angular momentum as

$$\frac{\mathrm{d}^{F}\boldsymbol{L}_{BA}^{F}}{\mathrm{d}t} = \sum_{i} \boldsymbol{r}_{iA} \times \boldsymbol{f}_{i}^{F,\mathrm{ext}} + \sum_{ij} \boldsymbol{r}_{iA} \times \boldsymbol{f}_{ij}^{F}, \qquad (13.5)$$

where $\mathbf{f}_i^{F,\text{ext}}$ is the "external" force acting on particle *i* (the force applied in *F* by everything except the body itself), and \mathbf{f}_{ij}^F is the "internal" force exerted by particle *j* on particle *i*. For the common case where all internal forces are central (i.e. radial), the second term $\sum_{ij} \mathbf{r}_{iA} \times \mathbf{f}_{ij}^F$ equals zero, so that

$$\frac{\mathrm{d}^{F}\boldsymbol{L}_{BA}^{F}}{\mathrm{d}t} = \sum_{i} \boldsymbol{r}_{iA} \times \boldsymbol{f}_{i}^{F,\mathrm{ext}} \quad (\mathrm{central \ forces \ internally})$$
$$\xrightarrow{(13.4)} \sum_{i} \boldsymbol{\tau}_{iA}^{F,\mathrm{ext}} \equiv \boldsymbol{\tau}_{BA}^{F,\mathrm{ext}}, \text{ the torque on the body.} \quad (13.6)$$

This is the torque law as it's usually applied. The torque is usually easiest to compute in an inertial frame, where there are no pseudo forces to include.

Together with angular momentum and moment of inertia, torque is defined relative to a point, not an axis (see page 25 for prior comments). To appreciate this idea, consider increasing the spin rate of a wheel by holding and twisting only the end of its axle and applying a torque. (Use an inertial frame to remove the complicating influence of gravity.)

DST-Group-TR-3309

The $\mathbf{r} \times \mathbf{f}$ term in (13.6) shows that the wheel's distance from the end of the axle affects the torque we must apply: after all, if it is very far from us, then we are effectively accelerating many distant point masses at the end of a rigid arm, and $\mathbf{r} \times \mathbf{f}$ has a large magnitude for each of these masses. Alternatively, imagine the wheel is severely unbalanced and spinning slowly, and we are trying to keep it rotating about a fixed axle. As it spins, it will pull sideways on the axle, which we must counteract by applying a torque vector that is not parallel to the axle. The further the wheel is from us, the harder we must work to counteract its tendency to want to move sideways: we must apply a greater torque to keep the axle pointing in a fixed direction.

14 Final Comments

Aerospace modelling typically combines many frames and coordinate systems following specifications that do not always make use of "nice" points such as centres of mass or origins of coordinate systems. Even making sense of simple kinematic scenarios requires a sound footing for the concepts of vectors, frames, and coordinates. The heavy use of sub- and superscripts in this document might at first seem excessive, but they come into their own when we move beyond simple calculations: describing complex scenarios that depend on several entities would quickly become unwieldy if we did not consistently indicate those entities in the notation. The payoff for doing so is that we are able to make computational sense of what could otherwise be a prohibitively complex environment.

I thank Peter Zipfel for discussions on this subject over some years. His book on aerospace modelling [2] includes a great many numerical examples that extend the analyses presented in this report.

15 References

[1] See "Row and column vectors" (https://en.wikipedia.org/wiki/Row_and_column_vectors), which doesn't distinguish coordinate vectors from proper vectors; while Wikipedia is not being treated as a credible citation source here, it can certainly be indicative of widespread opinion, incorrect as well as correct. General discussions and differences of opinion on vectors and related concepts are found at https://www.physicsforums.com and http://physics.stackexchange.com. E.g. "Is the covariance or contravariance of vectors/tensors something that can be visualized?" (http://physics.stackexchange.com/questions/87775/ is-the-covariance-or-contravariance-of-vectors-tensors-something-that-can-be-vi?) does not distinguish coordinate vectors from proper vectors; neither does "Is it foolish to distinguish between covariant and contravariant vectors?" (http://physics.stackexchange.com/questions/105347/ is-it-foolish-to-distinguish-between-covariant-and-contravariant-vectors?) (with differences of opinion therein). Admissions of confusion are found at http://www.physicsoverflow.org/24929/ is-partial-derivative-a-vector-or-dual-vector?merged=24735. The two pages "Gradient is covariant or contravariant?" (http://physics.stackexchange.com/

questions/126740/gradient-is-covariant-or-contravariant?) and "Is force a contravariant vector or a covariant vector (or either)?"

(http://physics.stackexchange.com/questions/131348/

is-force-a-contravariant-vector-or-a-covariant-vector-or-either?) obfuscate simple concepts while struggling to give any coherency to the subject. See also "What the hell are vector components?" (https://www.physicsforums.com/threads/

what-the-hell-are-vector-components.372639) and the various conflicting answers given therein. "What is a tensor?"

(https://www.physicsforums.com/threads/what-is-a-tensor.723969) highlights the natural confusion caused by the comment found in Zee's book *Einstein Gravity in a Nutshell* (Princeton University Press, 2013) that "a tensor is something that transforms like a tensor". The page "Vector notation"

(https://en.wikipedia.org/wiki/Vector_notation) confuses "polar vectors" (see the footnote on page 19 of this report) with polar coordinates. "Is partial derivative a vector or dual vector?" (http://physics.stackexchange.com/questions/144089/

is-partial-derivative-a-vector-or-dual-vector? and

http://www.physicsoverflow.org/24929/

is-partial-derivative-a-vector-or-dual-vector?merged=24735) show different approaches that rely on formalism more than explanation, as well as giving admissions of confusion. "How can a set of components fail to make up a vector?"

(http://physics.stackexchange.com/questions/168300/

how-can-a-set-of-components-fail-to-make-up-a-vector?) gives several definitions of a vector, again confusing coordinate vectors with proper vectors.

"Physicists' definition of vectors based on transformation laws"

(http://physics.stackexchange.com/questions/241610/

physicists-definition-of-vectors-based-on-transformation-laws/241633#

241633) shows mixed attempts to relate index-transformation ideas to more physical concepts. This page emphasises the basic distinction being discussed in this report, that coordinate vectors are distinct from proper vectors.

- [2] P.H. Zipfel (2014) Modeling and Simulation of Aerospace Vehicle Dynamics, 3rd ed., AIAA Inc., Virginia. Zipfel's notation differs slightly from that of this report, where he expresses e.g. a change of coordinates (written $[\boldsymbol{v}]_A = \mu_A^B[\boldsymbol{v}]_B$ in this report) as $[\boldsymbol{v}]^A = [T]^{AB}[\boldsymbol{v}]^B$. When introducing time rates of change, his " $[d\boldsymbol{s}/dt]^A$ " specifies no frame for the time derivative of \boldsymbol{s} , and instead needs to be interpreted as $d[\boldsymbol{s}]^A/dt$ in his notation, or $d[\boldsymbol{s}]_A/dt$ in our notation.
- [3] G.L. Bradley (1975) A Primer of Linear Algebra (Prentice Hall College) is an excellent textbook on linear algebra that distinguishes carefully between vector components (coordinate vectors) and vectors as elements of a vector space (visualised as arrows and called proper vectors in this report), and describes how different coordinate systems are related.
- [4] D. Koks (2012), A Pseudo-Reversing Theorem for Rotation and its Application to Orientation Theory. DSTO-TR-2675, Melbourne, Vic., Defence Science and Technology Organisation (Australia).
- [5] Such wording can easily be found with an appropriate text search of the internet.
- [6] M.D. Shuster (1993) A Survey of Attitude Representations, Journal of the Astronautical Sciences, 41 (4), pp. 439–517. See equation (117).

DST-Group-TR-3309

- [7] D. Koks (2006) *Explorations in Mathematical Physics*, Springer, New York. See Section 2.4.2.
- [8] J.J. Sakurai (1985) *Modern Quantum Mechanics*, Addison-Wesley. The angular momentum matrices can be found in all books on advanced quantum mechanics.
- [9] J.C. van der Ha and M.D. Shuster (2009) A Tutorial on Vectors and Attitude, *IEEE Control Systems Magazine*.
- [10] J.B. Kuipers (2002) Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace and Virtual Reality, Princeton University Press. See also [2].
- [11] A selection over some decades: P. Richards (1959) Manual of Mathematical Physics, Pergamon Press, Chapter 21; A. Eddington (1923) The Mathematical Theory of Relativity, Cambridge University Press, Section 29; S. Weinberg (1972) Gravitation and Cosmology, John Wiley and Sons, Chapter 4; A. Kyrala (1967) Theoretical Physics: Applications of Vectors, Matrices, Tensors and Quaternions, W.B. Saunders Company, Section 6.6.
- [12] One book that treats covariant differentiation similarly to this report is B. Schutz (1988) A First Course In General Relativity, Cambridge University Press.
- [13] W. Duncan et al. (1960), *Mechanics of Fluids*, Edward Arnold Publishing. The majority, if not all, books on mathematical fluid dynamics introduce D/Dt in the same way.

Appendix A: Relating Vector Components for Non-Orthonormal Coordinates

The analysis of Section 7 must be modified for coordinate systems that are not orthonormal (meaning their basis vectors are not orthonormal). Non-orthonormal bases are probably never used in aerospace dynamics; they are usually encountered in general discussions of tensor analysis, so this section is included for theoretical interest only and uses the standard "up/down" index notation of tensor analysis. In particular, the notation of Section 7 becomes modified, in that a vector v is now written in component form for coordinate system S as

$$\boldsymbol{v} = v^1 \boldsymbol{e}_1 + v^2 \boldsymbol{e}_2 + v^3 \boldsymbol{e}_3 \equiv v^\alpha \boldsymbol{e}_\alpha, \qquad (A1)$$

where the summation " $\alpha = 1, 2, 3$ " over a dummy index such as α is implied when that index appears once as a superscript and once as a subscript. The metric tensor g is defined as in (6.15), having $\alpha\beta$ component

$$g_{\alpha\beta} \equiv \boldsymbol{e}_{\alpha} \cdot \boldsymbol{e}_{\beta} \,. \tag{A2}$$

So $g_{\alpha\beta}$ is the $\alpha\beta$ component of the matrix $[g]_S$. Now define a new set of numbers $g^{\alpha\beta}$:

$$g^{\alpha\beta} \equiv \alpha\beta$$
 component of matrix $([\boldsymbol{g}]_S)^{-1}$, (A3)

and use these to define a new basis set $\{e^1, e^2, e^3\}$ which is called "dual" to $\{e_1, e_2, e_3\}$:

$$\boldsymbol{e}^{\alpha} \equiv g^{\alpha\beta} \boldsymbol{e}_{\beta} \,. \tag{A4}$$

Although the original basis $\{e_1, e_2, e_3\}$ is not necessarily orthonormal, this new basis introduces an inter-basis orthonormality:

$$\boldsymbol{e}^{\alpha} \cdot \boldsymbol{e}_{\beta} = g^{\alpha \mu} \boldsymbol{e}_{\mu} \cdot \boldsymbol{e}_{\beta} = g^{\alpha \mu} g_{\mu \beta} = \delta^{\alpha}_{\beta} \equiv \begin{cases} 1 & \alpha = \beta , \\ 0 & \text{otherwise.} \end{cases}$$
(A5)

For example, given coordinates u, v, w that are not necessarily orthogonal, the basis vector e^u will be orthogonal to e_v and e_w , the basis vector e^v will be orthogonal to e_u and e_w , and the basis vector e^w will be orthogonal to e_u and e_v . Similarly, the basis vector e_u will be orthogonal to e^v and e^w , and so on. It will be the case that $e^u = e_u$, $e^v = e_v$, and $e^w = e_w$ if and only if the basis $\{e_u, e_v, e_w\}$ is orthonormal. We see that for cartesian coordinates x, y, z, the two bases are identical: $e^x = e_x$ etc.

The new basis vectors are not necessarily orthonormal: an analysis similar to that of (A5) gives

$$\boldsymbol{e}^{\alpha} \cdot \boldsymbol{e}^{\beta} = g^{\alpha\beta} \,. \tag{A6}$$

The relationship between the two bases is symmetrical, in that

$$g_{\alpha\beta} \,\boldsymbol{e}^{\beta} = g_{\alpha\beta} \, g^{\beta\mu} \boldsymbol{e}_{\mu} = \delta^{\mu}_{\alpha} \, \boldsymbol{e}_{\mu} = \boldsymbol{e}_{\alpha} \,. \tag{A7}$$

That is, equations (A4) and (A7) are

$$e^{\alpha} = g^{\alpha\beta} e_{\beta}, \quad e_{\alpha} = g_{\alpha\beta} e^{\beta}.$$
 (A8)

We see that the dual of a dual basis is just the original basis. It's now straightforward to show that

$$v^{\alpha} = \boldsymbol{v} \cdot \boldsymbol{e}^{\alpha} \,. \tag{A9}$$

UNCLASSIFIED

39

DST-Group-TR-3309

We have also

$$\boldsymbol{v} \cdot \boldsymbol{e}_{\alpha} = v^{\beta} \boldsymbol{e}_{\beta} \cdot \boldsymbol{e}_{\alpha} = v^{\beta} g_{\beta\alpha} \equiv v_{\alpha} \,, \tag{A10}$$

again giving a symmetry in the notation:

$$v^{\alpha} = \boldsymbol{v} \cdot \boldsymbol{e}^{\alpha}, \quad v_{\alpha} = \boldsymbol{v} \cdot \boldsymbol{e}_{\alpha}.$$
 (A11)

Also,

$$\boldsymbol{v} = v^{\alpha} \boldsymbol{e}_{\alpha} = v^{\alpha} g_{\alpha\beta} \boldsymbol{e}^{\beta} = v_{\beta} \, \boldsymbol{e}^{\beta} \,, \tag{A12}$$

from which we see that just as the v^{α} are the components of v over the e_{α} basis, so too the v_{α} are the components of v over the e^{α} basis.

The significance of the dual basis is that the e_{α} component of v is $v \cdot e^{\alpha}$, and not in general $v \cdot e_{\alpha}$. Also, the nabla operator ∇ can be expressed using this dual basis as a sum over α :

$$\nabla = e^{\alpha} \frac{\partial}{\partial x^{\alpha}} \equiv e^{\alpha} \partial_{\alpha} \,. \tag{A13}$$

This expression holds in *any* coordinate system. It's often written in cartesian coordinates as a definition

$$\nabla \equiv \left(\partial_x, \partial_y, \partial_z\right)^{\prime\prime}, \tag{A14}$$

which (a) omits the basis vectors, and (b) does not suggest how ∇ might be written in arbitrary coordinates. Nabla is also sometimes written in general coordinates as

$$\ \, \overset{``}{\nabla} \equiv \sum_{\alpha} \boldsymbol{e}_{\alpha} \frac{\partial}{\partial x^{\alpha}} \, \overset{''}{} \quad (\text{Wrong!}),$$
 (A15)

which uses the wrong basis vectors. The general expression $\nabla = e^{\alpha} \partial_{\alpha}$ uses the dual basis, and can be used to compute the divergence and curl in arbitrary coordinates by calculating " $\nabla \cdot$ " and " $\nabla \times$ ".⁸

It can easily be shown that any orthonormal basis is identical to its dual, in which case the above raised/lowered indices need not be used, and we can revert to the simpler notation of Section 7.

Using the above notation, we can now reformulate the analysis of Section 7 for a general basis. Consider coordinate systems A and A' (these names are chosen here so that components are more economically written than in Section 7: those for A have no prime, and those for A' have a prime), and use just two dimensions for economy. Write an arbitrary proper vector \boldsymbol{v} as

$$\boldsymbol{v} = v^{\alpha} \boldsymbol{e}_{\alpha} = v^{\beta'} \boldsymbol{e}_{\beta'} \,. \tag{A19}$$

⁸For example, the divergence is

$$\nabla \cdot \boldsymbol{v} = (\boldsymbol{e}^{\alpha} \,\partial_{\alpha}) \cdot (v^{\beta} \boldsymbol{e}_{\beta}) = \boldsymbol{e}^{\alpha} \cdot \left[\partial_{\alpha} (v^{\beta} \boldsymbol{e}_{\beta})\right]. \tag{A16}$$

Appendix C shows that in the language of tensor analysis,

$$\partial_{\alpha}(v^{\beta}\boldsymbol{e}_{\beta}) = v^{\beta}{}_{;\alpha}\boldsymbol{e}_{\beta}, \qquad (A17)$$

in which case

$$\nabla \cdot \boldsymbol{v} = \boldsymbol{e}^{\alpha} \cdot \boldsymbol{v}^{\beta}_{;\alpha} \boldsymbol{e}_{\beta} = \delta^{\alpha}_{\beta} \, \boldsymbol{v}^{\beta}_{;\alpha} = \boldsymbol{v}^{\alpha}_{;\alpha} \,. \tag{A18}$$

This last expression $v^{\alpha}_{\ :\alpha}$ is how the divergence is usually written in tensor analysis.

Now consider

$$[\boldsymbol{v}]_{A'} = \begin{bmatrix} v^{1'} \\ v^{2'} \end{bmatrix} = \begin{bmatrix} \boldsymbol{v} \cdot \boldsymbol{e}^{1'} \\ \boldsymbol{v} \cdot \boldsymbol{e}^{2'} \end{bmatrix} = \begin{bmatrix} v^{\alpha} \boldsymbol{e}_{\alpha} \cdot \boldsymbol{e}^{1'} \\ v^{\alpha} \boldsymbol{e}_{\alpha} \cdot \boldsymbol{e}^{2'} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}_1 \cdot \boldsymbol{e}^{1'} & \boldsymbol{e}_2 \cdot \boldsymbol{e}^{1'} \\ \boldsymbol{e}_1 \cdot \boldsymbol{e}^{2'} & \boldsymbol{e}_2 \cdot \boldsymbol{e}^{2'} \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} \equiv \mu_{A'}^A [\boldsymbol{v}]_A, \quad (A20)$$

where

$$\mu_{A'}^{A} \equiv \begin{bmatrix} \boldsymbol{e}_{1} \cdot \boldsymbol{e}^{1'} & \boldsymbol{e}_{2} \cdot \boldsymbol{e}^{1'} \\ \boldsymbol{e}_{1} \cdot \boldsymbol{e}^{2'} & \boldsymbol{e}_{2} \cdot \boldsymbol{e}^{2'} \end{bmatrix},\tag{A21}$$

In other words, as an extension to the idea of Section 7, the columns of $\mu_{A'}^A$ are the loweredindex basis vectors \boldsymbol{e}_{α} of A written in raised-index A' coordinates (found from dot products with $\boldsymbol{e}^{\beta'}$):

$$\mu_{A'}^A = \begin{bmatrix} [\boldsymbol{e}_1]_{A'} & [\boldsymbol{e}_2]_{A'} \end{bmatrix}$$
(A22)

and the rows of $\mu_{A'}^A$ are the raised-index basis vectors of A' written in lowered-index A coordinates.

Inverting $[\boldsymbol{v}]_{A'} = \mu_{A'}^A [\boldsymbol{v}]_A$ shows that $\mu_{A'}^A$ is still the inverse of $\mu_A^{A'}$ as in Section 7; but these two matrices are no longer in general the transposes of each other.

Finally, equation (7.14) still holds, because the analysis in (7.13) is unchanged when using non-orthonormal coordinates.

A.1 Example: Calculating ∇ in Spherical Polar Coordinates

For an uncommon example of the utility of the dual basis, we calculate ∇ in spherical polar coordinates r, θ, ϕ . Without even needing to know how these coordinates are defined, the answer is simply

$$\nabla = \boldsymbol{e}^{\alpha} \,\partial_{\alpha} = \boldsymbol{e}^{r} \,\partial_{r} + \boldsymbol{e}^{\theta} \,\partial_{\theta} + \boldsymbol{e}^{\phi} \,\partial_{\phi} \,. \tag{A23}$$

But we tend to work either with the basis $\{e_r, e_\theta, e_\phi\}$ or with the normalised basis $\{\hat{e}_r, \hat{e}_\theta, \hat{e}_\phi\}$. To find these, apply the discussion of Section 6 to write

$$\boldsymbol{e}_{\alpha} = \frac{\partial x}{\partial \alpha} \, \boldsymbol{e}_{x} + \frac{\partial y}{\partial \alpha} \, \boldsymbol{e}_{y} + \frac{\partial z}{\partial \alpha} \, \boldsymbol{e}_{z} \quad \text{for } \alpha = r, \theta, \phi \,.$$
 (A24)

Now we need to know how the spherical polar coordinates relate to x, y, z. The standard approach begins with $r - r \sin \theta \cos \phi$

$$\begin{aligned} x &= r \sin \theta \cos \phi \,, \\ y &= r \sin \theta \sin \phi \,, \\ z &= r \cos \theta \,. \end{aligned} \tag{A25}$$

Working in cartesian coordinates "C" with $s_{\theta} \equiv \sin \theta$, $c_{\theta} \equiv \cos \theta$ and similarly for ϕ , we then have

$$[\boldsymbol{e}_r]_C = \frac{\partial}{\partial r} \begin{bmatrix} x\\ y\\ z \end{bmatrix} = \begin{bmatrix} s_\theta c_\phi\\ s_\theta s_\phi\\ c_\theta \end{bmatrix}, \quad [\boldsymbol{e}_\theta]_C = \begin{bmatrix} r c_\theta c_\phi\\ r c_\theta s_\phi\\ -r s_\theta \end{bmatrix}, \quad [\boldsymbol{e}_\phi]_C = \begin{bmatrix} -r s_\theta s_\phi\\ r s_\theta c_\phi\\ 0 \end{bmatrix}.$$
(A26)

Relate the usual basis to the dual basis via (A4), for which is needed the metric elements. The dot product is independent of coordinate system and is most simply calculated in cartesian coordinates in the usual way; so write

$$g_{rr} = \boldsymbol{e}_r \cdot \boldsymbol{e}_r = [\boldsymbol{e}_r]_C \cdot [\boldsymbol{e}_r]_C \xrightarrow{\text{(A26)}} 1, \qquad (A27)$$

UNCLASSIFIED

41

DST-Group-TR-3309

and similarly

$$g_{\theta\theta} = r^2, \quad g_{\phi\phi} = r^2 \sin^2 \theta,$$
 (A28)

with all other $g_{\alpha\beta}$ equal to zero.⁹ The metric matrix is then diagonal, in which case in-voke (A3):

$$g^{rr} = \frac{1}{g_{rr}}, \quad g^{\theta\theta} = \frac{1}{g_{\theta\theta}}, \quad g^{\phi\phi} = \frac{1}{g_{\phi\phi}},$$
 (A29)

with all other $g^{\alpha\beta}$ equal to zero. Then (A4) becomes

$$e^{r} = g^{r\alpha} e_{\alpha} = g^{rr} e_{r} = \frac{e_{r}}{g_{rr}} = e_{r},$$

$$e^{\theta} = g^{\theta\theta} e_{\theta} = \frac{e_{\theta}}{g_{\theta\theta}} = \frac{e_{\theta}}{r^{2}},$$

$$e^{\phi} = g^{\phi\phi} e_{\phi} = \frac{e_{\phi}}{g_{\phi\phi}} = \frac{e_{\phi}}{r^{2} \sin^{2} \theta}.$$
(A30)

The normalised basis is

$$\widehat{\boldsymbol{e}}_{r} = \frac{\boldsymbol{e}_{r}}{|\boldsymbol{e}_{r}|} = \frac{\boldsymbol{e}_{r}}{\sqrt{g_{rr}}} = \boldsymbol{e}_{r} , \quad \widehat{\boldsymbol{e}}_{\theta} = \frac{\boldsymbol{e}_{\theta}}{\sqrt{g_{\theta\theta}}} = \frac{\boldsymbol{e}_{\theta}}{r} , \quad \widehat{\boldsymbol{e}}_{\phi} = \frac{\boldsymbol{e}_{\phi}}{\sqrt{g_{\phi\phi}}} = \frac{\boldsymbol{e}_{\phi}}{r\sin\theta} , \quad (A31)$$

and similarly

$$\widehat{\boldsymbol{e}}^r = \boldsymbol{e}^r, \quad \widehat{\boldsymbol{e}}^\theta = r \boldsymbol{e}^\theta, \quad \widehat{\boldsymbol{e}}^\phi = r \sin \theta \, \boldsymbol{e}^\phi.$$
 (A32)

Finally, (A23) can be expressed in various ways as

$$\nabla = \mathbf{e}^r \,\partial_r + \mathbf{e}^\theta \,\partial_\theta + \mathbf{e}^\phi \,\partial_\phi$$

$$= \widehat{\mathbf{e}}^r \,\partial_r + \frac{\widehat{\mathbf{e}}^\theta}{r} \,\partial_\theta + \frac{\widehat{\mathbf{e}}^\phi}{r\sin\theta} \,\partial_\phi$$

$$= \mathbf{e}_r \,\partial_r + \frac{\mathbf{e}_\theta}{r^2} \,\partial_\theta + \frac{\mathbf{e}_\phi}{r^2\sin^2\theta} \,\partial_\phi$$

$$= \widehat{\mathbf{e}}_r \,\partial_r + \frac{\widehat{\mathbf{e}}_\theta}{r} \,\partial_\theta + \frac{\widehat{\mathbf{e}}_\phi}{r\sin\theta} \,\partial_\phi \,. \tag{A33}$$

The last two forms in (A33) are seen frequently in the literature, being usually derived in other ways that are sometimes economical and other times laborious.

A.2 Comment on the Vector Components v_{α}

The notation of this appendix embodies the idea that every coordinate system has associated with it two natural bases, $\{e_{\alpha}\}$ and $\{e^{\alpha}\}$, which are identical if and only if the coordinates are orthonormal. A vector v has component v^{α} for the basis vector e_{α} , and component v_{α} for the basis vector e^{α} . (The pairing of down with up and up with down indices here is simply conventional; convention could just as well have paired down with down and up with up.) This pairing of bases is used extensively in crystallography.

⁹I have calculated g_{rr} using $[e_r]_C$ here, but it should be clear that it has unit length from its definition in (6.1) or (6.2). Similarly, $g_{\theta\theta}$ and $g_{\phi\phi}$ can be calculated from (6.1) or (6.2) by studying the geometry of the polar coordinate system.

Now recall page 22, on which we defined a tensor v^{\times} that acts on a vector a to produce the cross-product vector $v \times a$. The tensor v^{\times} was conveniently coordinatised as a 2 × 2 matrix. We might choose to do the same with the dot product, defining a tensor " v^{\bullet} " that acts on a vector a to produce the dot product $v \cdot a$. This tensor is conveniently coordinatised as the 1 × 3 matrix $[v_1 v_2 v_3]$, so that

$$\boldsymbol{v}^{\bullet}\boldsymbol{a} \equiv \boldsymbol{v} \cdot \boldsymbol{a} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} a^1 \\ a^2 \\ a^3 \end{bmatrix} = v_{\alpha} a^{\alpha} \,. \tag{A34}$$

In this view, the dot product has been absorbed into the vector v to create the new tensor v^{\bullet} . But whereas the notation v^{\times} is useful for a cross product, it's not clear whether there is anything to be gained by introducing a new tensor v^{\bullet} .

However, modern differential geometry does introduce a new object along the same lines as v^{\bullet} : just as (A5) says that $e^{\alpha} \cdot e_{\beta} = \delta^{\alpha}_{\beta}$, differential geometry introduces a new object called a "one-form" or "differential form" as a function of vectors and written variously as $\tilde{\omega}^{\alpha}$ or $d\tilde{x}^{\alpha}$, so that $\tilde{\omega}^{\alpha}(e_{\beta}) \equiv \delta^{\alpha}_{\beta}$. The numbers v_{α} are then set to be the coefficients of a linear combination of basis one-forms: thus we find ourselves dealing with the one-form $v_{\alpha}\tilde{\omega}^{\alpha}$. The rationale for this idea is that it does not require a dot product to be defined, which appears as early as (A2). The upshot is that one dot product is replaced by an infinite set of one-forms—along with several chapters of the required theory needing inclusion in textbooks on the subject. Additionally, well-established higher-order vector concepts such as bivectors become replaced by "two-forms", and so on. My own view is that Ockham would not be impressed by this introduction of such complexity, both notational and pedagogical. *All* of physics presupposes the idea of measuring length, and this is precisely what a dot product does. So physics cannot dispense with the dot product, and this implies that one-forms are not necessary in the subject.

Of course, one-forms and their higher-order relatives might be defined as an abstract mathematical exercise, but this does not guarantee they will be necessary or indeed useful for anything outside the realm of abstraction for its own sake. Every calculation that I have ever seen that uses one-forms with an end result not involving one-forms can be done more simply using vectors and a dot product. The standard and conventional mantras of one-forms state, for example, that "one-forms form the rigor behind infinitesimals". Does this phrase have any depth? Note that (A13) makes it obvious that $\nabla x^{\alpha} = e^{\alpha}$ for any coordinate system, so e.g. $\nabla r = e^r$, $\nabla \theta = e^{\theta}$, and so on. I believe that this simple identification of ∇ with the dual basis is fundamentally the reason why infinitesimals, which are related to ∇ via standard vector expressions such as $df(\boldsymbol{x}) = \nabla f \cdot d\boldsymbol{x}$, are sometimes misidentified (via the dual basis) as one-forms. Infinitesimals and one-forms both use linear notation, but that does not imply any connection between them; and I am very sure that no user of one-forms could even remotely begin to use them as an aid to envisaging the small displacement $v \, dt$ undergone by a flying aircraft in time dt. Another mantra is "one-forms render the curl operator easier to understand". But the curl operator is straightforwardly written in any coordinates by replacing an expression such as " $\nabla \times A$ " with " $e^{\alpha} \partial_{\alpha} \times A$ ", meaning " $e^{\alpha} \times \partial_{\alpha} A$ ", which is not mysterious.

Although n-forms appear in various fields of physics today, I have never seen anything emerge from their use that sheds light for me on anything that is not defined in terms of n-forms. Vector integral theorems such as those of Gauss and Stokes are certainly not simplified when remoulded into form language. It seems that in recent years the use of n-forms has transformed older straightforward vector calculus into something abstruse and convoluted

DST-Group-TR-3309

that one is obliged to master only to gain acceptance into the appropriate academic clique. In particular, I think that the use of forms has made general relativity inaccessible to a great number of students who would have no problem understanding the subject if it were presented using vectors. The only real use for n-forms that I am aware of is to produce more theorems about n-forms. But that is a very circular state of affairs.

Appendix B: Time Derivative of a Rank-2 Tensor

Equations (10.8a)–(10.8c) show how the coordinates of the time derivative of a vector relate to the time derivative of the coordinates of a vector. Here I derive the analogous expression for a rank-2 tensor L. That is, how does $[d^A L/dt]_B$ relate to $d[L]_B/dt$? The calculation is very similar to (10.9) and (10.10), but now using (7.14) instead of (7.4):

$$\begin{bmatrix} \frac{\mathrm{d}^{A}L}{\mathrm{d}t} \end{bmatrix}_{B} \stackrel{(7.14)}{==} \mu_{B}^{A} \begin{bmatrix} \frac{\mathrm{d}^{A}L}{\mathrm{d}t} \end{bmatrix}_{A} \mu_{B}^{B} = \mu_{B}^{A} \frac{\mathrm{d}[L]_{A}}{\mathrm{d}t} \mu_{A}^{B}$$

$$= \mu_{B}^{A} \frac{\mathrm{d}}{\mathrm{d}t} \left(\mu_{A}^{B} [L]_{B} \mu_{B}^{A} \right) \mu_{A}^{B}$$

$$= \mu_{B}^{A} \left(\dot{\mu}_{A}^{B} [L]_{B} \mu_{B}^{A} + \mu_{A}^{B} \frac{\mathrm{d}[L]_{B}}{\mathrm{d}t} \mu_{B}^{A} + \mu_{A}^{B} [L]_{B} \dot{\mu}_{B}^{A} \right) \mu_{A}^{B}$$

$$= \mu_{B}^{A} \dot{\mu}_{A}^{B} [L]_{B} + \frac{\mathrm{d}[L]_{B}}{\mathrm{d}t} + [L]_{B} \dot{\mu}_{B}^{A} \mu_{A}^{B}$$

$$\stackrel{(10.17)}{==} \mu_{B}^{A} \dot{\mu}_{A}^{B} [L]_{B} + \frac{\mathrm{d}[L]_{B}}{\mathrm{d}t} - [L]_{B} \mu_{B}^{A} \dot{\mu}_{A}^{B}. \tag{B1}$$

Now use (10.22) to write $\mu_B^A \dot{\mu}_A^B = \left[\mathbf{\Omega}^{BA} \right]_B^{\times}$, and hence (B1) becomes

$$\left[\frac{\mathrm{d}^{A}L}{\mathrm{d}t}\right]_{B} = \left[\mathbf{\Omega}^{BA}\right]_{B}^{\times} \left[L\right]_{B} + \frac{\mathrm{d}[L]_{B}}{\mathrm{d}t} - \left[L\right]_{B} \left[\mathbf{\Omega}^{BA}\right]_{B}^{\times},\tag{B2}$$

or finally, using the well-known "commutator bracket" $[P,Q] \equiv PQ - QP$,

$$\left[\frac{\mathrm{d}^{A}L}{\mathrm{d}t}\right]_{B} = \frac{\mathrm{d}[L]_{B}}{\mathrm{d}t} + \left[\left[\mathbf{\Omega}^{BA}\right]_{B}^{\times}, [L]_{B}\right].$$
(B3)

In particular, set $L = \mathbf{\Omega}^{BA^{\times}}$ in (B3) to produce

$$\left[\frac{\mathrm{d}^{A}\boldsymbol{\Omega}^{BA^{\times}}}{\mathrm{d}t}\right]_{B} = \frac{\mathrm{d}[\boldsymbol{\Omega}^{BA^{\times}}]_{B}}{\mathrm{d}t} = \left[\frac{\mathrm{d}^{B}\boldsymbol{\Omega}^{BA^{\times}}}{\mathrm{d}t}\right]_{B}.$$
(B4)

That is, A and B both measure the same value for how their relative angular velocity Ω^{BA} changes with time. But in fact we already know this last point: we could have derived it by differentiating the vector Ω^{BA} rather than the tensor $\Omega^{BA^{\times}}$:

$$\frac{\mathrm{d}^{A} \mathbf{\Omega}^{BA}}{\mathrm{d}t} \stackrel{(10.20)}{=} \left(\frac{\mathrm{d}^{B}}{\mathrm{d}t} + \mathbf{\Omega}^{BA} \right) \mathbf{\Omega}^{BA} = \frac{\mathrm{d}^{B} \mathbf{\Omega}^{BA}}{\mathrm{d}t} + \mathbf{\Omega}^{BA} \times \mathbf{\Omega}^{BA}$$
$$= \frac{\mathrm{d}^{B} \mathbf{\Omega}^{BA}}{\mathrm{d}t} \,. \tag{B5}$$

DST-Group-TR-3309

This page is intentionally blank.

Appendix C: Two Derivatives found in Tensor Analysis and Fluid Flow

In line with the discussion of Section 10 in which various expressions for a time derivative were given, this appendix explores two derivatives that have acquired their own names in their respective fields and are traditionally presented as more advanced than the ordinary and partial derivatives of calculus. But they are not more advanced at all; they are simply ordinary and partial derivatives, in the same way that the time derivative of Section 10 is an ordinary time derivative.

C.1 The Covariant Derivative

We have already seen the first of these derivatives: the "covariant derivative". The calculation of Section 10.1.3 is very similar to one encountered in tensor analysis, especially in the four spacetime dimensions of relativity. To see how, again write a vector \boldsymbol{v} as a linear combination of basis vectors that are not necessarily orthonormal. Again we use a language of components that is common in tensor analysis:

$$\boldsymbol{v} = v^{\alpha} \boldsymbol{e}_{\alpha} \,, \tag{C1}$$

where summation over the repeated up-down index (here α) is assumed, and in particular α sums over one time dimension labelled 0 and three space dimensions labelled 1,2,3. In relativity each frame is given its own coordinates, so that e.g. in three frames labelled as unprimed, primed, and barred, v is written as

$$\boldsymbol{v} = v^{\alpha} \boldsymbol{e}_{\alpha} = v^{\alpha'} \boldsymbol{e}_{\alpha'} = v^{\bar{\alpha}} \boldsymbol{e}_{\bar{\alpha}} \,. \tag{C2}$$

Letting the coordinates denote the frame in this way means we can dispense with a frame notation such as "F", and write the partial derivative with respect to any one of the coordinates of, say, the unprimed frame (label it x^{β}) as simply

$$\frac{\partial \boldsymbol{v}}{\partial x^{\beta}} = \frac{\partial \left(v^{\alpha} \boldsymbol{e}_{\alpha}\right)}{\partial x^{\beta}} = \frac{\partial v^{\alpha}}{\partial x^{\beta}} \boldsymbol{e}_{\alpha} + v^{\alpha} \frac{\partial \boldsymbol{e}_{\alpha}}{\partial x^{\beta}}.$$
(C3)

We will use the standard "comma" notation, where a subscripted ", β " denotes $\partial/\partial x^{\beta}$. Equation (C3) is then

$$\boldsymbol{v}_{,\beta} = (v^{\alpha} \boldsymbol{e}_{\alpha})_{,\beta} = v^{\alpha}_{,\beta} \, \boldsymbol{e}_{\alpha} + v^{\alpha} \, \boldsymbol{e}_{\alpha,\beta}. \tag{C4}$$

Similar to Section 10.1.3, expand $e_{\alpha,\beta}$ as a linear combination of basis vectors:

$$\boldsymbol{e}_{\alpha,\beta} = \Gamma^{\mu}_{\ \alpha\beta} \, \boldsymbol{e}_{\mu} \tag{C5}$$

for four numbers $\Gamma^{\mu}_{\ \alpha\beta}$ known as *Christoffel symbols*. Equation (C4) becomes

$$\boldsymbol{v}_{,\beta} = \boldsymbol{v}^{\alpha}_{,\beta} \boldsymbol{e}_{\alpha} + \boldsymbol{v}^{\alpha} \Gamma^{\mu}_{\ \alpha\beta} \boldsymbol{e}_{\mu} \quad (\text{now swap dummy indices } \alpha \text{ and } \mu)$$
$$= \left(\boldsymbol{v}^{\alpha}_{,\beta} + \boldsymbol{v}^{\mu} \Gamma^{\alpha}_{\ \mu\beta}\right) \boldsymbol{e}_{\alpha} \equiv \boldsymbol{v}^{\alpha}_{\ ;\beta} \boldsymbol{e}_{\alpha} \,. \tag{C6}$$

So, provided that we remember to carry out this procedure of *covariant differentiation*,

$$v^{\alpha}_{;\beta} \equiv v^{\alpha}_{\ \beta} + v^{\mu} \Gamma^{\alpha}_{\ \mu\beta} \,, \tag{C7}$$

UNCLASSIFIED

47

DST-Group-TR-3309

we can ignore the existence of the basis entirely, and focus only on components. In other words, the covariant derivative allows us to "pretend" that the basis vectors don't change, so that $\partial(v^{\alpha} \boldsymbol{e}_{\alpha})/\partial x^{\beta} = v^{\alpha}_{\ \beta} \boldsymbol{e}_{\alpha}$. We then have:

 $v^{\alpha}_{\ \ \ \beta} [\equiv (v^{\alpha})_{\ \ \beta}]$ is the β partial derivative of the α coordinate of the vector \boldsymbol{v} , and $v^{\alpha}_{\ \ \ \beta} [\equiv (\boldsymbol{v}_{\ \ \beta})^{\alpha}]$ is the α coordinate of the β partial derivative of the vector \boldsymbol{v} . (C8)

Widespread in the field of general relativity is the practice of mimicking the idea that $v^{\alpha}_{,\beta}$ is the same as $\partial_{\beta} v^{\alpha}$ by writing the semicolon using operator notation: thus $v^{\alpha}_{;\beta}$ becomes " $\nabla_{\beta} v^{\alpha}$ ". I see this as unfortunate because it expresses covariant differentiation as an operator (on v^{α} in this case), which is incorrect. The reason is that whereas the comma in (C8) is simply the partial derivative operator acting on v^{α} , the semicolon is *not* an operator acting on v^{α} , due to the presence of all components of \boldsymbol{v} on the right-hand side of (C7). So a particular v^{α} might be zero everywhere and yet have a non-zero covariant derivative. For example, consider the two-dimensional radial vector field $\boldsymbol{v} = v^r \boldsymbol{e}_r$ in polar coordinates r, θ . Here $v^{\theta} = 0$ everywhere and so certainly $v^{\theta}_{,\theta} \equiv \partial_{\theta} v^{\theta} = 0$, but it turns out that $v^{\theta}_{;\theta} = v^r/r$, an expression that we could hardly obtain by applying any operator to zero.

Referring to (C8), one might ask why the notation $v^{\alpha}_{;\beta}$ is necessary at all, given that $(\boldsymbol{v}_{,\beta})^{\alpha}$ works just as well. Traditionally (at least in the tensor calculus of relativity), the coordinatevector notation v^{α} has almost always been used to represent the proper vector \boldsymbol{v} ; that is, the same notation v^{α} has traditionally been used for proper vectors and coordinate vectors. But if v^{α} is used to denote both the α -element of the coordinate vector and the proper vector \boldsymbol{v} itself, then what does $v^{\alpha}_{,\beta}$ mean—does it denote the β -derivative of the α -element, or does it denote the β -derivative of the proper vector \boldsymbol{v} , or perhaps the α -element of that derivative? Equation (C8) distinguishes these quantities. Presumably $v^{\alpha}_{;\beta}$ is a historical artifact designed to provide the necessary clarity. Our practice in this report has been always to distinguish proper vectors \boldsymbol{v} from coordinate vectors. The α -element of the coordinate vector is v^{α} , and the entire coordinate vector might be written as v^{α} when there is no likelihood of confusion and we are making a point of conforming to standard tensor language. But we write the proper vector always as \boldsymbol{v} .

The calculations in this appendix have been very straightforward (and not novel: see e.g. [12]). All we have done is differentiate a vector, and the covariant derivative has emerged naturally as a result. As pointed out on page 31, covariant differentiation is often described in textbooks as an artificially constructed operation, with the Christoffel symbols $\Gamma^{\mu}_{\ \alpha\beta}$ being introduced as correction terms that must be added to the partial derivative in (C7) to render it compliant with the behaviour expected of vector components. But we see here that rather than being a separate construction of tensor analysis, the covariant derivative is simply a convenient way to calculate the components of the usual partial derivative without having to think about basis vectors or write them down. It is useful without being esoteric.

C.2 A Derivative Used in Fluid Flow

The study of fluid flow makes frequent use of the time derivative in a form known by any of a host of names in the literature, such as "material derivative" or "convective derivative". This derivative is simply the time derivative of some property of a physical element of the fluid. By way of introduction, picture a solid object with a temperature distribution that is unchanging in time, but which varies from place to place on the object. If we sample the temperature with

a sensor that moves over the object, the temperature that the sensor sees will change with time and be path dependent; and the faster the sensor moves on a given path, the faster will be the temperature changes. If the object's temperature distribution is T(x, y, z) then $\partial T/\partial t = 0$, but clearly d(sensor temp)/dt $\neq 0$. And yet the sensor temperature is clearly just T at the position of the sensor. The notational problem here is that simply writing "dT/dt $\neq 0$ " does not indicate that the temperature change is due to the motion of the sensor. We might instead write "dT_{sensor}/dt $\neq 0$ ", and although $T_{sensor} = T$ at all points, the subscript "sensor" is not redundant because it indicates that we are following the motion of a sensor.

In the study of fluid flow, consider the temperature of a stream of water, and suppress the z coordinate to save space in what follows. Taken as a whole body of water, at any point in time and space the water has a temperature distribution T(t, x, y). The body of water itself might be warmed by the sun, accounting for the time dependence in T(t, x, y). But as the water flows, it might encounter space-fixed sources of heat transfer, such as geothermal activity or icebergs attached to the river bank. These account for the space dependence of T(t, x, y). We are typically interested not so much in the overall temperature distribution T, as in the temperature of a particular *element* of water while this element follows the stream. This element is chosen at some place (x, y, z), so its position is a function of (x, y, z); but once chosen, its temperature has only time dependence. So we write this temperature of the *choice* of element has less of our focus than the time dependence of the *temperature* of that element.

In particular, ask how this temperature $T_{\rm el}$ changes with time. The moving element samples any possible time dependence of the water body's temperature distribution, as well as encountering different temperatures in different regions of the water body. The element moves with velocity \boldsymbol{v} , expressed in our chosen coordinates as (v_x, v_y) [fluid physics typically employs lowered indices for its vector components, unlike the usage of Section C.1]. We calculate the rate of increase of $T_{\rm el}$ by simply following the element as it moves by an amount $\boldsymbol{v} \, dt$ in a time dt, subtracting initial $T_{\rm el}$ (i.e. initial T) from final $T_{\rm el}$ (i.e. final T) and dividing by the time interval:

$$\frac{\mathrm{d}T_{\mathrm{el}}}{\mathrm{d}t} = \frac{T(t+\mathrm{d}t, x+v_x\mathrm{d}t, y+v_y\mathrm{d}t) - T(t, x, y)}{\mathrm{d}t} \,. \tag{C9}$$

The first term in the numerator above can be Taylor expanded, leading to

$$\frac{\mathrm{d}T_{\mathrm{el}}}{\mathrm{d}t} = \frac{T(t, x, y, z) + \frac{\partial T}{\partial t} \mathrm{d}t + \frac{\partial T}{\partial x} v_x \mathrm{d}t + \frac{\partial T}{\partial y} v_y \mathrm{d}t - T(t, x, y)}{\mathrm{d}t}$$
$$= \frac{\partial T}{\partial t} + \frac{\partial T}{\partial x} v_x + \frac{\partial T}{\partial y} v_y = \frac{\partial T}{\partial t} + v \cdot \nabla T \,. \tag{C10}$$

We could have found the same expression by reasoning that as the element moves for a time dt, its amount of temperature increase due to the sun is $\partial T/\partial t dt$, and its amount of temperature increase due to space-fixed sources is the increase it picks up by moving through space, being e.g. $\partial T/\partial x dx$ in the x direction. Its overall temperature increase is the sum of these:

$$dT_{\rm el} = \frac{\partial T}{\partial t} dt + \frac{\partial T}{\partial x} dx + \dots, \qquad (C11)$$

from which (C10) follows. (This is, of course, precisely what a Taylor expansion is doing.)

Although the temperature $T_{\rm el}$ of the element at any moment and point equals the temperature distribution T evaluated at that point, by using the symbol $T_{\rm el}$ I emphasise that we are

DST-Group-TR-3309

following the progress of a particular element as it moves in the flow. The literature of fluid dynamics typically takes a different course [13]; it solely uses "T", and then introduces a new derivative notation for $dT_{\rm el}/dt$, calling it for example DT/Dt and writing (C10) as

I suggest that using the same symbol T for both $T_{\rm el}$ and T is not productive, because such use defocusses attention from the fact that we are following the evolution of a specific fluid element, and it leads to D/Dt being treated as a new operator deserving of its own name such as material derivative. But there is no new calculus present in D/Dt that must be learned by the aspiring fluid dynamicist. The material derivative is just the time derivative (d/dt) of some scalar property of a moving element, and as such it surely does not require a special name. I think the focus should be on what is being differentiated with d/dt (the temperature $T_{\rm el}$ of an element as distinct from the overall distribution of temperature T over space and time), rather than on the idea that there is somehow a more advanced version of differentiation present here, which there is not.

On a final segue, the above ideas are also used when applying Newton's laws to an element of the fluid to derive its motion, given a knowledge of all forces present. In a given frame (whose presence is understood and so not indicated in the equations that follow) we wish to write the acceleration of a fluid element. Obviously, the velocity $v_{\rm el}$ of a fluid element at any time and place equals the velocity v of the fluid at that time and place, but we can usefully retain $v_{\rm el}$ to remind ourselves that we are applying Newton's "force equals mass times acceleration" to an element, and so write the acceleration of the element as $dv_{\rm el}/dt$:

$$\frac{\mathrm{d}\boldsymbol{v}_{\mathrm{el}}}{\mathrm{d}t} = \frac{\boldsymbol{v}(t+\mathrm{d}t,x+v_{x}\mathrm{d}t,y+v_{y}\mathrm{d}t)-\boldsymbol{v}(t,x,y)}{\mathrm{d}t}$$

$$= \frac{\boldsymbol{v}(t,x,y,z)+\frac{\partial \boldsymbol{v}}{\partial t}\mathrm{d}t+\frac{\partial \boldsymbol{v}}{\partial x}v_{x}\mathrm{d}t+\frac{\partial \boldsymbol{v}}{\partial y}v_{y}\mathrm{d}t-\boldsymbol{v}(t,x,y)}{\mathrm{d}t}$$

$$= \frac{\partial \boldsymbol{v}}{\partial t}+\frac{\partial \boldsymbol{v}}{\partial x}v_{x}+\frac{\partial \boldsymbol{v}}{\partial y}v_{y} \equiv \frac{\partial \boldsymbol{v}}{\partial t}+(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}.$$
(C13)

So we see that the acceleration $d\mathbf{v}_{\rm el}/dt$ of the fluid element turns out to be the perhaps unfamiliar-looking $\partial \mathbf{v}/\partial t + (\mathbf{v}\cdot\nabla)\mathbf{v}$, which is slightly different to the expression in (C12) because now a vector is being operated on instead of the scalar in (C12). But the underlying ideas are straightforward, and again there is no new derivative here. Setting this acceleration equal to the total force on the element is the entry point to the subject of advanced fluid dynamics. For example, $(\mathbf{v}\cdot\nabla)\mathbf{v}$ can be re-expressed using $\nabla \times \mathbf{v}$, and from this curl emerges the study of vorticity within the fluid.

Page classification: UNCLASSIFIED

Г

DEFENCE SCIEN DOCU	1. DLM	/CAVEAT (OF DOCUMENT)							
2. TITLE			3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED RE-						
On the Use of Vectors, Refer	cence Frames, and Coordina	ate	PORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO						
Systems in Aerospace Analys	DOCUMENT CLASSIFICATION)								
			Document (U)						
			Title (U)						
	Abstract (U)								
4. AUTHOR			5. CORPORATE AUTHOR						
Don Koks	Don Koks				Defence Science and Technology Group				
			PO Box 1500						
	Edinburgh, SA 5111, Australia								
6a. DST Group NUMBER	6b. AR NUMBER		6c. TYPE OF REPOR	RТ	7. DOCUMENT DATE				
DST-Group-TR-3309	016-726	16-726			January 2017				
8. Objective ID	9. TASK NUMBER	10). TASK SPONSOR						
	ADF XX/XXX								
13. DST Group Publications Re	epository		14. RELEASE AUTHORITY						
http://dspace.dsto.defenc	e.gov.au/dspace		Chief, Cyber and Electronic Warfare Division						
15. SECONDARY RELEASE S	STATEMENT OF THIS DOC	UMI	ENT						
Approved for Public Release									
OVERSEAS ENQUIRIES OUTSIDI EDINBURGH, SA 5111	E STATED LIMITATIONS SHOU	ILD I	BE REFERRED THROUGH	H DOCUM	ENT EXCHANGE, PO BOX 1500,				
16. DELIBERATE ANNOUNC	EMENT								
No Limitations									
17. CITATION IN OTHER DO	CUMENTS								
No Limitations									
18. RESEARCH LIBRARY THESAURUS									
vector analysis, reference frames, coordinate systems, aerospace analysis, tensors									
19. ABSTRACT									
This report describes the core foundational concepts of aerospace modelling, an understanding of which is necessary for									
the analysis of complex environments that hold many entities, each of which might employ a separate reference frame and									
coordinate system. I begin by defining vectors, frames, and coordinate systems, and then discuss the quantities that allow									
Newton's laws to be applied to a complex scenario. In particular, I explain the crucial distinction between the "coordinates									
of the time-derivative of a vector" and the "time-derivative of the coordinates of a vector". I finish by drawing a parallel									

Page classification: UNCLASSIFIED

dynamics, and make some comments on various supposedly different derivatives as found in the literature.

between this aerospace language and the notation found in seemingly unrelated areas such as relativity theory and fluid