STOKES' THEOREM: CALCULUS OF DIFFERENTIAL FORMS

THESIS

Presented to the Graduate Councıl
of Texas State University–San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Master of SCIENCE

by

Christopher E. Johnson, B.S

San Marcos, Texas
December 2004

ACKNOWLEDGMENTS

I would first like to acknowledge my debt of gratitude to the creator of this most spectacular universe for the beauty therein and for endowing mankind with reason. Humans' ability to create logical and rational thoughts allows our species to better appreciate the many splendors of life. The field of mathematics is one such splendor. Like our physical universe the basic elements in mathematics are simple yet together magnificent in their harmony. The second debt of gratitude is also to God who has surrounded me with inexpressibly profound love in the sacrifice of His son and in the persons of my parents and family, my wife, my friends, and my teachers.

My father and mother have taught me the benefits of self-sacrifice, hard work, and faith in God. Their unconditional love was instrumental in my nurturing. My grandma Johnson was also influential in my upbringing by introducing me to art and particularly music—the only manmade construct more beautiful than mathematics. My siblings and extended family have supported my endeavors. I thank my entire family for their love, kindness, and generosity.

My wife Denise has filled my life with a joy I would not have otherwise known. She is devoted to me and I to her. Her accomplishments have inspired mine. During those times of melancholy when a new insight brightened my abstract universe showing me the relative insignificance of my knowledge, she has convinced me of the worth of my pursuits. As my sun, she makes me like the full moon—bright only with her light. I thank her for her patience whilst my work has consumed me and for her unwavering support even when I seemed to be standing in the shadow of failure.

I thank my friends collectively for their many thoughtful gestures and kind words.

In regard to this thesis, my teachers are all due a special note of thanks. I am greatly indebted to Dr. Curtin for his expertise in the area of differential geometry. His class presentations have inspired me to focus on similar ideas in my future endeavors. The time commitment he gave to reading this thesis, and the subsequent comments have been invaluable.

I am likewise grateful to Dr. Singh whose extensive knowledge of mathematics and teaching philosophy has significantly shaped my development in mathematics. He cast me and my peers far off into a number of ideas, and in a spirit of competition, we worked enthusiastically to demonstrate our understanding of them. In due time, he always reeled us back. I have a lasting impression of the many ideas to which he exposed me, and I am thankful for his patience, encouragement, and wisdom.

Lastly, I am ineffably grateful to Dr. McCabe. He was the teacher who first guided me to discover that the beauty in mathematics lies not only in the constructs themselves, but in the words we use to express the constructs and precisely communicate our ideas. My journey to this point of understanding and indeed through the year of writing this thesis would not have been possible without his unending patience, dedication to his profession, and commitment to the philosophy of learning that belief comes from experience. It is with Dr. McCabe that my career in mathematics truly began.

This manuscript was submitted on 18 December 2004.

TABLE OF CONTENTS

ABSTRACT

STOKES' THEOREM: CALCULUS OF DIFFERENTIAL FORMS

by

Christopher Elliot Johnson, B.S.

Texas State University–San Marcos

December 2004

SUPERVISING PROFESSOR TERENCE MCCABE

This thesis connects a number of fields of mathematics in relation to Euclidean n-space. It defines the meanings of differentiation for functions between these spaces and gives an exposition of the inverse function theorem. One also finds the definition for integration of real valued function defined on a Euclidean n-space. These definitions of differentiation and integration are precursors to the topics of differential forms and integration of forms over chains that stand out as the main ideas developed herein. A great deal of effort is spent on developing the algebraic structure of differential forms including the non-trivial associative property of the wedge product. The final chapter ties the previous chapters together nicely in a result known as Stokes' Theorem.

CHAPTER I

INTRODUCTION

The general Stokes' Theorem is named for Sir George Gabriel Stokes (1819 – 1903). While he is not the originator of the premise and conclusion presented in the theorem that bears his name, it is not a grossly inaccurate designation. After sitting for the Mathematical Tripos, Cambridge mathematics graduates were given an opportunity to further distinguish themselves by competing for the Smith's Prize. From 1849 to 1882, it was Stokes' duty to set one paper for this competitive exam, and on it he asked for a proof of the theorem that the examinees began to refer to as Stokes' Theorem. Despite this deviation from the standard nomenclature of mathematical theorems, Stokes is indeed tied to this theorem in a more remarkable way.

G. G. Stokes published hundreds of papers on mathematics and physics, won many extraordinary awards, and worked in the most prestigious academic positions of his day. Beyond his academic accomplishments he is said to have been well-regarded by his colleagues and students. Even while in the position of Lucasian Professor of Mathematics at Cambridge he made a declaration to offer help to anyone at Cambridge who found themselves troubled by problems in mathematics. These were not empty words. He led many to be successful through encouragement and suggestion of problems. It is through his capacity as a dutiful professor that his eternal link to Stokes' Theorem was establish. Two of his most famous pupils were James Clerk Maxwell and William Thomson also called Lord Kelvin.

In fact, it was William Thomson who seems to have been the first to state the theorem in the post-script of a letter he wrote to Stokes. After Cambridge, Thomson went to Paris where his colleagues included Cauchy, Liouville, and Sturm, among others. Liouville had perhaps the most influence on Thomson for suggesting that he work to unite the ideas of Faraday, Coulomb, and Poisson. While Thomson followed the suggestion of Liouville, it is Maxwell who succeeded in creating a unified theory of electromagnetism. It is the physical phenomena summarized by Maxwell using the theorem proposed by Thomson and bearing the name of Stokes that has tied these three men together in the most interesting and profound way.

The theorem has evolved much since the time of Stokes due mainly to the advent of differential forms. Forms have been hailed as a powerful tool in making fundamentals of electromagnetic field theory intuitive. Just as vectors are important for representing displacement and velocity, differential forms are useful for representing field intensity and flux density. Furthermore, differential forms allow Stokes' Theorem to not only nicely relate the grad, curl, and div, but in fact replace the Divergence Theorem altogether.

This paper uses the construct of differential forms to express and prove a version of Stokes' Theorem that involves the all essential difficulties. It builds from scratch much of the mathematical ideas necessary to define and prove the theorem. We will follow the approach of Michael Spivak. Like most good books on mathematics, Spivak's *Calculus on Manifolds* carefully lays out a sequence of ideas while omitting much of the detail in showing their validity. This style brings enjoyment to a wide audience by allowing advanced readers to progress swiftly through the ideas without becoming bored with the intricate details and by giving the novice readers a chance to thoroughly understand the ideas through working out the details for themselves. This thesis works out many of the details.

Preliminary Notation

We start this section on notation with the real number system—which we will denote with the symbol $\mathbb{R}$—since it is the foundation for all that we will build in this paper. We mention next a generalization of the real number system often referred to as *Euclidean n-space* and denoted by $\mathbb{R}^n$. We neglect the particulars of the mathematical structure and subtleties since they are standard, but we do want to mention a notational convention. We use bold print for elements of $\mathbb{R}^n$ also called vectors, and we use superscript for the real number components of a vector. For example, whenever we write $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x}$ represents an n-tuple of real numbers $(x^1, x^2, \ldots, x^n)$ where $x^i \in \mathbb{R}$ for each $i \in \{1, 2, \ldots, n\}$.

We will use the usual absolute value symbols exclusively for real numbers. For example, we write $|x|$ for the absolute value of a real number x. For the usual Euclidean norm of a vector $\boldsymbol{x} \in \mathbb{R}^n$, we write $\|\boldsymbol{x}\|$. We will make these definition formal.

Definition 1.1 For a number $x \in \mathbb{R}$ the *absolute value* of x, written $|x|$, is equal to x whenever $x \geq 0$ and $-x$ whenever $x < 0$.

Definition 1.2 For a vector $\boldsymbol{x} \in \mathbb{R}^n$ the *norm* of $\boldsymbol{x} = (x^1, x^2, \ldots, x^n)$, written $\|\boldsymbol{x}\|$, is the real number equal to $\sqrt{\sum_{i=1}^{n} (x^i)^2}$.

We will use elementary properties of the absolute value and norm. Other more profound properties we will expound upon in a proof. As an example and exposition of our symbol conventions so far, we prove the following necessary theorem.

Theorem 1.1 Suppose $\boldsymbol{x} \in \mathbb{R}^n$. If $\boldsymbol{x} = (x^1, x^2, \ldots, x^n)$, then $\|\boldsymbol{x}\| \leq \sum_{i=1}^{n} |x^i|$ and $\sum_{i=1}^{n} |x^i| \leq \sqrt{n}\, \|\boldsymbol{x}\|$.

Proof. Let $\quad = (x^1, x^2, \ldots, x^n) \in \mathbb{R}^n$. We will show the first conclusion by exhibiting $\|\boldsymbol{x}\|^2$ as part of $\left(\sum_{i=1}^{n} |x^i|\right)^2$, where $\left(\sum_{i=1}^{n} |x^i|\right)^2$ is a sum of non-negative terms. The principle square root of each side of this inequality will yield the desired result. We expand $\left(\sum_{i=1}^{n} |x^i|\right)^2$ below.

$$
\begin{array}{ccccccc}
|x^1|\,|x^1| & + & |x^1|\,|x^2| & + \cdots + & |x^1|\,|x^n| + \\
|x^2|\,|x^1| & + & |x^2|\,|x^2| & + \cdots + & |x^2|\,|x^n| + \\
+ & & & & \\
\vdots & & & \ddots & \\
+ & & & & \\
|x^n|\,|x^1| & + & |x^n|\,|x^2| & + \cdots + & |x^n|\,|x^n|
\end{array}
$$

The terms written on the descending-diagonal in the arrangement above are exactly $\|\boldsymbol{x}\|^2$ written as $\sum\limits_{i=1}^{n} (x^i)^2 = \sum\limits_{i=1}^{n} |x^i|\,|x^i|$. We make this explanation formal. Since the terms above the descending-diagonal can be written $\sum\limits_{i=1}^{n-1}\left(|x^i| \sum\limits_{j=i+1}^{n} |x^j|\right)$, the terms below the descending-diagonal can be written $\sum\limits_{i=2}^{n}\left(|x^i| \sum\limits_{j=1}^{i-1} |x^j|\right)$, and each is non-negative then,

$$\|\boldsymbol{x}\|^2 = \sum_{i=1}^{n} |x^i|\,|x^i| \le \sum_{i=1}^{n} |x^i|\,|x^i| + \sum_{i=1}^{n-1}\left(|x^i| \sum_{j=i+1}^{n} |x^j|\right) + \sum_{i=2}^{n}\left(|x^i| \sum_{j=1}^{i-1} |x^j|\right) = \left(\sum_{i=1}^{n} |x^i|\right)^2 .$$

In particular, $\|\boldsymbol{x}\|^2 \le \left(\sum\limits_{i=1}^{n} |x^i|\right)^2$ so we arrive at our first conclusion $\|\boldsymbol{x}\| \le \sum\limits_{i=1}^{n} |x^i|$.

For the second conclusion we will exhibit $n\,\|\boldsymbol{x}\|^2 - \left(\sum\limits_{i=1}^{n} |x^i|\right)^2 \ge 0$, which quickly leads to the result. We can look at the expansion $n \sum\limits_{i=1}^{n} |x^i|\,|x^i| - \left(\sum\limits_{i=1}^{n} |x^i|\right)^2$ in the arrangement below.

$$
\begin{array}{ccccc}
|x^1|\,|x^1| - |x^1|\,|x^1| & + & |x^2|\,|x^2| - |x^1|\,|x^2| & + \cdots + & |x^n|\,|x^n| - |x^1|\,|x^n| + \\
|x^1|\,|x^1| - |x^2|\,|x^1| & + & |x^2|\,|x^2| - |x^2|\,|x^2| & + \cdots + & |x^n|\,|x^n| - |x^2|\,|x^n| + \\
+ & & & & \\
\vdots & & & \ddots & \\
+ & & & & \\
|x^1|\,|x^1| - |x^n|\,|x^1| & + & |x^2|\,|x^2| - |x^n|\,|x^2| & + \cdots + & |x^n|\,|x^n| - |x^n|\,|x^n|
\end{array}
$$

Each term on the descending diagonal of the above arrangement is identically zero and each remaining difference has a pair the sum of which makes a perfect square. For each $i, j \in \{1, 2, \ldots, n\}$ with $i \ne$, we have the difference $|x^i|\,|x^i| - |x^j|\,|x^i|$ which can be factored as $|x^i|\,(|x^i| - |x^j|)$, the difference $|x^j|\,|x^j| - |x^i|\,|x^j|$ which can be factored $-|x\ |\,(|x\ | - |x\ |)$, and the sum of these two factored differences can be further factored as $(|x^i| - |x^j|)^2$. We can formally write

$$0 \le \sum_{i=1}^{n-1}\left(\sum_{j=i+1}^{n} (|x^i| - |x^j|)^2\right) = n \sum_{i=1}^{n} |x^i|\,|x^i| - \left(\sum_{i=1}^{n} |x^i|\right)^2 .$$

In particular, $\left(\sum_{i=1}^{n} |x^i|\right)^2 \le n \sum_{i=1}^{n} |x^i|\,|x^i|$, thus follows $\sum_{i=1}^{n} |x^i| \le \sqrt{n}\,\sqrt{\sum_{i=1}^{n} |x^i|\,|x^i|} =$ $\sqrt{n}\,\|\boldsymbol{x}\|$, the second conclusion. ∎

Using the work in the proof, we can see that equality in the first conclusion holds whenever is on an axis, that is, whenever $\boldsymbol{x} = (0, 0, \ldots, x^i, \ldots, 0)$ for some $i \in \{1, 2, \ldots, n\}$; equality in the second conclusion holds whenever $\boldsymbol{x}$ is the corner of a generalized cube centered at $\boldsymbol{0} = (0, 0, \ldots, 0)$, that is, wherever $|x^i| - |x^j| = 0$ for each $i, j \in \{1, 2, \ldots, n\}$; and equality throughout holds when either $n = 1$ or $\boldsymbol{x} = \boldsymbol{0}$.

The next bit of notation we will introduce is that of the inner product. This brings us to the first instance of two vectors being used simultaneously. When we need to distinguish only two or a few vectors we will often use different letter symbols, but when the number of different letter symbols becomes unruly, we will instead use subscripts.

Definition 1.3 For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, the inner product of and $\boldsymbol{y}$, written $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is equal to $\sum_{i=1}^{n} x^i y^i$.

Linear Algebra & Functions

We write the usual basis for $\mathbb{R}^n$ as $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$. For example, $\boldsymbol{e}_2 = (0, 1, 0, 0) \in \mathbb{R}^4$ where the 1 is in the second position indicated by the 2 in $\boldsymbol{e}_2$.

Definition 1.4 A function $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is called *linear* if for each $\alpha, \beta \in \mathbb{R}$ and each $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ we have $\lambda(\alpha \cdot \boldsymbol{x} + \beta \cdot \boldsymbol{y}) = \alpha \cdot \lambda(\boldsymbol{x}) + \beta \cdot \lambda(\boldsymbol{y})$.

When a basis for each vector space is specified, a linear function between those vector spaces has a matrix representation. We refrain from making a theorem here, leaving it to the linear algebraists, however, the coming importance of a clear understanding of linear functions warrants an exposition of the implication for the purposes of this paper as well as an example. In the case of a linear function $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ the ith column of the $m \times n$ matrix $A = (a_{ij})$ for $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$ is determined by the m coefficients of $\lambda(\boldsymbol{e}_i)$ when it is written as a linear combination of the basis vectors. Suppose $\lambda : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ and let $\boldsymbol{x} \in \mathbb{R}^2$. Each of $\lambda(\boldsymbol{e}_1)$ and $\lambda(\boldsymbol{e}_2)$ is an element of $\mathbb{R}^3$, therefore we can uniquely write each as a sum of 3 coefficients and the three standard basis elements for $\mathbb{R}^3$. We write $\lambda(e_1) = a_{11} \cdot e_1 + a_{21} \cdot e_2 + a_{31} \cdot e_3$, $\lambda(e_2) = a_{12} \cdot e_1 + a_{22} \cdot e_2 + a_{32} \cdot e_3$, $\boldsymbol{x} = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = x^1 \cdot e_1 + x^2 \cdot e_2$, and we are set to show how it works.

We start by writing $\boldsymbol{x}$ as a linear combination of the basis elements of $\mathbb{R}^2$, $\lambda(\boldsymbol{x}) = \lambda(x^1 \cdot e_1 + x^2 \cdot e_2)$. Next we apply the linearity of λ which yields $\lambda(x^1 \cdot e_1 + x^2 \cdot e_2) = x^1 \cdot \lambda(e_1) + x^2 \cdot \lambda(e_2)$. We can substitute the unique vector representation of $\lambda(e_1)$ and $\lambda(e_2)$ from above and note the result is just the matrix product of $A \cdot \boldsymbol{x}$,

$$x^1 \cdot \lambda(e_1) + x^2 \cdot \lambda(e_2) = x^1 \cdot \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} + x^2 \cdot \begin{pmatrix} a_{12} \\ a_{22} \\ a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = A \cdot \boldsymbol{x}.$$

A is determined completely from the basis elements of $\mathbb{R}^2$ with no dependence on our arbitrary ; therefore, for each $\boldsymbol{x} \in \mathbb{R}^2$, $\lambda(\boldsymbol{x}) = A \cdot \boldsymbol{x}$. Here after if a basis is known for both the domain and codomain, we will use either representation of a linear function whenever convenient. We continue now with an important theorem about a certain type of boundedness in regard to linear functions between Euclidean vector spaces.

Theorem 1.2 If $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, then there exists a number M such that if $\in \mathbb{R}^n$ then $\|\lambda(\boldsymbol{x})\| \leq M \, \|\boldsymbol{x}\|$.

Proof. Suppose $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation. Since we have the usual basis for Euclidean vector spaces, we know there exists a matrix representation of the linear transformation and we call it $A = (a_{ij})$. We define K as $\max \{\sum_{i=1}^{m} |a_{i1}|, \sum_{i=1}^{m} |a_{i2}|, \ldots, \sum_{i=1}^{m} |a_{in}|\}$. Note K is determined from the columns in A. For each column, the absolute value of the components is summed and the largest of these sums is K.

Pick $M = \sqrt{n}\ K$ and let $\in \mathbb{R}^n$.

We write under the influence of linear transformation λ in the equivalent and more convenient form:

$$\lambda(\boldsymbol{x}) = A \cdot \boldsymbol{x} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}\, x^j\right) \boldsymbol{e}_i,$$

so $\|\lambda(\boldsymbol{x})\| = \left\|\sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij}\, x^j)\, \boldsymbol{e}_i\right\|$. By the triangle inequality and properties of normed vector spaces, $\left\|\sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij}\, x^j)\, \boldsymbol{e}_i\right\| \leq \sum_{i=1}^{m} |(\sum_{j=1}^{n} a_{ij}\, x^j)|\, \|\boldsymbol{e}_i\|$. For each $i \in \{1, 2, \ldots, m\}$, $\|\boldsymbol{e}_i\| = 1$, so $\sum_{i=1}^{m} |(\sum_{j=1}^{n} a_{ij}\, x^j)|\, \|\boldsymbol{e}_i\| = \sum_{i=1}^{m} |(\sum_{j=1}^{n} a_{ij}\, x^j)|$. Again by the triangle inequality, $\sum_{i=1}^{m} |(\sum_{j=1}^{n} a_{ij}\, x^j)| \leq \sum_{i=1}^{m} (\sum_{j=1}^{n} |a_{ij}|\, |x^j|)$. We regroup the larger sum by j and factor out $|x^j|$ to give $\sum_{i=1}^{m} (\sum_{j=1}^{n} |a_{ij}|\, |x^j|) = \sum_{j=1}^{n} (|x^j| \sum_{i=1}^{m} |a_{ij}|)$. Now for each $j \in \{1, 2, \ldots, n\}$, by our choice of K, $\sum_{i=1}^{m} |a_{ij}| \leq K$ from which follows $\sum_{j=1}^{n} (|x^j| \sum_{i=1}^{m} |a_{ij}|) \leq \sum_{j=1}^{n} |x^j|\, K = (\sum_{j=1}^{n} |x^j|)\, K$. Finally from Theorem 1.1, we have $\sum_{j=1}^{n} |x^j| \leq \sqrt{n}\, \|\boldsymbol{x}\|$, so $(\sum_{j=1}^{n} |x^j|)\, K \leq \sqrt{n}\, K\, \|\boldsymbol{x}\| = M\, \|\boldsymbol{x}\|$.

We have shown that given a linear transformation $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ we were able to choose a number M so that for each $\boldsymbol{x} \in \mathbb{R}^n$ $\|\lambda(\boldsymbol{x})\| \leq M\, \|\boldsymbol{x}\|$. ∎

CHAPTER II

DIFFERENTIATION AND INTEGRATION

In this chapter, we build on our previous understanding of the calculus of function of real variables. We start with a discussion of differentiation of functions between various Euclidean n-spaces. We then spend a good deal of energy on the Inverse Function Theorem as it is one of those ubiquitous theorems underlying much of the applications of calculus In the final section we investigate classic notions of integration, and in particular the Fundamental Theorem of Calculus in one dimension as it is this theorem which Stokes' Theorem is an analogy to in higher dimensions.

A thorough understanding of this chapter is essential for the rest of this thesis. We have in fact constructed rigorous proofs for each of the results within this chapter, however, we refrain from presenting a good number of these as they would distract from the goal of this paper. We include statements of definitions and theorems as reference, and we give an occasional proof.

Differentiation

We first ask what differentiable could mean in an abstract setting such as $\mathbb{R}^n$. We recall in the real number setting a real valued function being differentiable at a number p means there is a line that is the best approximation to at p. Generalized, a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ being differentiable at a point $\boldsymbol{p}$ means there is a linear function that is the best approximation to at $\boldsymbol{p}$.

Definition 2.1 A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is *differentiable* at $\boldsymbol{p} \in \mathbb{R}^n$ if there exists a linear transformation $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ so that if $\varepsilon > 0$ then there is a $\delta > 0$ so that if $\boldsymbol{x} \in \mathbb{R}^n$ with $||\boldsymbol{x} - \boldsymbol{a}|| < \delta$ then $\frac{||[f(x)-f(p)]-[\lambda(x)-\lambda(p)]||}{||x-p||} < \varepsilon$.

One often substitutes the vector $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{p}$ into Definition 2.1 and carries out some simplifications. It is often convenient to define differentiable as satisfying the following limit:

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{||f(p+h)-f(p)-\lambda(h)||}{||h||} = 0.$$

A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable on $A \subseteq \mathbb{R}^n$ if f is differentiable at $\boldsymbol{p}$ for each $\boldsymbol{p} \in A$.

Theorem 2.1 If a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ there is a unique linear transformation $\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ so that if $\varepsilon > 0$ then there is a $\delta > 0$ so that if $\boldsymbol{x} \in \mathbb{R}^n$ with $||\boldsymbol{x} - \boldsymbol{a}|| < \delta$ then $\frac{||[f(x)-f(p)]-[\lambda(x)-\lambda(p)]||}{||x-p||} < \varepsilon$. We denote λ by $Df(\boldsymbol{p})$ and call it the *derivative* of f at $\boldsymbol{p}$.

We have seen that whenever a linear transformation is defined between two vector spaces each with a basis, then there is a matrix representation for that linear transformation.

Definition 2.2 For differentiable function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ at $\boldsymbol{p} \in \mathbb{R}^n$ the matrix representation of $Df(\boldsymbol{p}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is called the *Jacobian Matrix* and is denoted $f'\ (\boldsymbol{p})$.

Theorem 2.2 If a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ then f is continuous at $\boldsymbol{p}$.

Proof. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be differentiable at $\boldsymbol{p} \in \mathbb{R}^n$. Since $Df(\boldsymbol{p})$ is the derivative of f at $\boldsymbol{p}$, then we need only show $\lim_{\boldsymbol{h}\to\boldsymbol{0}} ||Df(\boldsymbol{p})\,(\boldsymbol{h})|| = 0$. Since $Df(\boldsymbol{p})$ is a linear transformation, then by Theorem 1.2 pick a number M such that $||Df(\boldsymbol{p})\,(\boldsymbol{h})|| \leq M\,||\boldsymbol{h}||$ for all $\boldsymbol{h} \in \mathbb{R}^n$. Let $\varepsilon > 0$. Pick $\delta > 0$ such that $\delta = \frac{\varepsilon}{M}$. Let $\boldsymbol{h} \in \mathbb{R}^n$ such that $||\boldsymbol{h}|| < \delta$. We have $||Df(\boldsymbol{p})\,(\boldsymbol{h})|| \leq M\,||\boldsymbol{h}||$, $M\,||\boldsymbol{h}|| < M\delta = \varepsilon$. ∎

Theorem 2.3 If a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ and $g : \mathbb{R}^m \longrightarrow \mathbb{R}^l$ is differentiable at $f(\boldsymbol{p}) \in \mathbb{R}^m$, then $g \circ f : \mathbb{R}^n \longrightarrow \mathbb{R}^l$ is differentiable at $\boldsymbol{p}$, and $D(g \circ f)\,(\boldsymbol{p}) = D(g(f(\boldsymbol{p}))) \circ Df(\boldsymbol{p})$.

Theorem 2.4 If a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is constant, then $Df(\boldsymbol{p}) = 0$.

Theorem 2.5 If a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, then $Df(\boldsymbol{p}) = f$.

From consideration in linear algebra, we have seen for $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ there are m unique component functions $f^{(i)} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, and we can use this to define partial derivatives

Theorem 2.6 A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ if and only if $f^i : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable at $\boldsymbol{p}$ for each $i \in \{1, 2, \ldots, m\}$.

Theorem 2.7 If functions $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ are each differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ then $f + g$ is differentiable at $\boldsymbol{p}$, and $D(f + g)\,(\boldsymbol{p}) = Df(\boldsymbol{p}) + Dg(\boldsymbol{p})$.

Theorem 2.8 If functions $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ are each differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ then $f \cdot g$ is differentiable at $\boldsymbol{p}$, and $D(f \cdot g)\,(\boldsymbol{p}) = g(\boldsymbol{p})\,Df(\boldsymbol{p}) + f(\boldsymbol{p})\,Dg(\boldsymbol{p})$.

Theorem 2.9 If functions $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and each of f and g is differentiable at $\boldsymbol{p} \in \mathbb{R}^n$ and $g(\boldsymbol{p}) \neq \boldsymbol{0}$ then f / g is differentiable at $\boldsymbol{p}$, and $D(f / g)\,(\boldsymbol{p}) = \frac{g(\boldsymbol{p})\,Df(\boldsymbol{p}) - f(\boldsymbol{p})\,Dg(\boldsymbol{p})}{[g(\boldsymbol{p})]^2}$.

Definition 2.3 If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a function the ith *partial derivative* of f at $\boldsymbol{p} \in \mathbb{R}^n$ is $\lim_{x^i \to p^i} \frac{f(p^1, \ \ ,p^i - x^i, \ldots, p^n) - f(p^i, \ldots, p^n)}{p^i - x^i}$, if the limit exists, and is denoted $\ _i f(\boldsymbol{p})$.

Theorem 2.10 If $D_i\,(D_j\,f(\boldsymbol{p}))$ and $\ _j\,(D_i\,f(\boldsymbol{p}))$ are continuous in an open set containing $\boldsymbol{p}$, then $D_i\,(D_j\,f(\boldsymbol{p})) = D_j\,(D_i\,f(\boldsymbol{p}))$.

Theorem 2.11 Let $A \subset \mathbb{R}^n$. If the maximum or minimum of $f : A \longrightarrow \mathbb{R}$ occurs at a point $\boldsymbol{p}$ in the interior of A and $D_i\,f(\boldsymbol{p})$ exists, then $D_i\,f(\boldsymbol{p}) = 0$.

Theorem 2.12 If a function $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $\boldsymbol{p} \in \mathbb{R}^n$, then $D_j f^i\,(\boldsymbol{p})$ exists $i \in \{1, 2, \dots, n\}$ and $j \in \{1, 2, \dots, m\}$, and $f'\,(\boldsymbol{p})$ is the $m \times n$ matrix

$$\begin{pmatrix} D_1 f^1\,(\boldsymbol{p}) & D_2 f^1\,(\boldsymbol{p}) & \cdots & D_n f^1\,(\boldsymbol{p}) \\ D_1 f^2\,(\boldsymbol{p}) & D_2 f^2\,(\boldsymbol{p}) & \cdots & D_n f^2\,(\boldsymbol{p}) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f^m\,(\boldsymbol{p}) & D_2 f^m\,(\boldsymbol{p}) & \cdots & D_n f^m\,(\boldsymbol{p}) \end{pmatrix}.$$

Definition 2.4 A function $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ is *continuously differentiable* at $\boldsymbol{p} \in \mathbb{R}^n$ if there exists an open set $A \subset \mathbb{R}^n$ with $\boldsymbol{p} \in A$ such that if $\boldsymbol{x} \in A$, then $\ _j f^i\,(\boldsymbol{x})$ exists and $D_j f^i\,(\boldsymbol{p})$ is continuous for $j \in \{1, 2, \dots, n\}$ and $i \in \{1, 2, \dots, m\}$.

Theorem 2.13 If a function $f:\mathbb{R}^n \longrightarrow \mathbb{R}^m$ is continuously differentiable at $\boldsymbol{p} \in \mathbb{R}^n$, then $\ f(p)$ exists.

The Inverse Function Theorem

In this section we begin by discussing some aspects of the Inverse Function Theorem. This ranges from importance of the several hypotheses to examples of its uses. This is a constructive and considerably more accessible albeit long proof compared to the relatively short and less insightful versions of the proof that involve Banach's Fixed Point Lemma. We first consider a concise statement for the purposes of informal presentation.

Inverse Function Theorem: Suppose $f:\mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a continuously differentiable function on some open subset O of $\mathbb{R}^n$. If $p \in O$ and $\det\,[f'(p)] \neq 0$, then there exists open sets $\mathcal{U}$ and $\mathcal{V}$ of $\mathbb{R}^n$ so that: $p \in \mathcal{U}$, $f(\mathcal{U}) = \mathcal{V}$, $f^{-1}:\mathcal{V} \longrightarrow \mathcal{U}$ exists and $f^{-1}(\mathcal{V}) = \mathcal{U}$, and $(f^{-1})'\,(q) = [f' \circ f^{-1}(q)]^{-1}$ for each $q \in \mathcal{V}$.

We first make several remarks on various of the hypotheses and conclusions. Our goal is to establish some bijection and that only has a chance if the domain and codomain have the same dimension. Secondly, we have the important hypothesis of continuously differentiable on an open set. To illustrate this, suppose we reword the theorem so that f is only defined on O and differentiable only at $p \in O$. Define the function

$f:(-\pi, \pi) \longrightarrow (-\pi, \pi)$ by

$$f(x) = \begin{cases} \sin x & \text{if } x \text{ is rational} \\ x & \text{if } x \text{ is irrational.} \end{cases}$$

This function is differentiable only at 0 and its derivative is 1, not zero, but there is no open subset $\mathcal{U}$ so that is bijective on that set because for any irrational number y in the range there is a rational and irrational in the domain that both map to y, so one can never pick out a set on which f^{-1} is a well defined function.

Note also that hypotheses may be sufficient, but not necessary for a bijective function to exist. For example $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = x^3$ has $f'(0) = 0$ and yet $f^{-1} : \mathbb{R} \longrightarrow \mathbb{R}$ defined of course by $f(x) = \sqrt[3]{x}$ is a well defined function. What does necessarily follow though is that f^{-1} cannot be differentiable at 0.

To build a familiarity for the general problem, we will examine the Inverse Function Theorem in the setting of functions of real variables. We will maintain the wording of the more general setting. The wording will seem awkward since open connected proper subsets in the setting of the real line have the simpler terminology of segments or rays, still the consistent wording will serve to connected the simpler case to the general case.

Theorem 2.14 Suppose $f : \mathbb{R} \longrightarrow \mathbb{R}$ is a continuously differentiable function on some open subset O of $\mathbb{R}$. If $p \in O$ and $f'(p) \neq 0$, then there exists open sets $\mathcal{U}$ and $\mathcal{V}$ of $\mathbb{R}$ so that:

1. $p \in \mathcal{U}$,
2. $f(\mathcal{U}) = \mathcal{V}$,
3. $f^{-1} : \mathcal{V} \longrightarrow \mathcal{U}$ exists and $f^{-1}(\mathcal{V}) = \mathcal{U}$, and
4. $(f^{-1})'(q) = [f' \circ f^{-1}(q)]^{-1}$ for each $q \in \mathcal{V}$.

Proof. Let $p \in O$ such that $f'(p) \neq 0$. We will assume that $f'(p) > 0$, the other case following similarly. Since f' is continuous on O and $\frac{1}{2} f'(p) > 0$, then by a definition of continuous there is a number $\delta > 0$ so that if $x \in O$ with $0 < |x - p| < \delta$ then $|f'(x) - f'(p)| < \frac{1}{2} f'(p)$. We pick $\mathcal{U} = O \bigcap (p - \delta, p + \delta)$. Since $p \in O$ and $p \in (p - \delta, p + \delta)$ and each of these is open, then $p \in \mathcal{U}$ and $\mathcal{U}$ is open as the finite intersection of open sets is open.

Since is differentiable at each $x \in O$, then by Theorem 2.2, is continuous on O. In particular since $\mathcal{U} \subseteq O$, then is continuous on $\mathcal{U}$, and by previous theorems from topology, the continuous image of a segment (an open compact, connected set) is a segment. Since $\mathcal{U}$ is a segment and is continuous on $\mathcal{U}$, then we pick $\mathcal{V} = f(\mathcal{U})$ and from topology $\mathcal{V}$ is necessarily a segment and hence an open set.

We next show is invertible. Let $x, y \in \mathcal{U}$ such that $x < y$. Since is continuous on $[x, y]$ and differentiable on (x, y), then by the mean value theorem for derivatives of real value functions, there exists $s \in \mathcal{U}$ so that $f'(s)(y - x) = f(y) - f(x)$.

Since $s \in \mathcal{U}$, then $|s - p| < \delta$, so $|f'(s) - f'(p)| < \frac{1}{2} f'(p)$. Since $|f'(s) - f'(p)| < \frac{1}{2} f'(p)$, then by a property of absolute value inequalities, $-\frac{1}{2} f'(p) < f'(s) - f'(p) < \frac{1}{2} f'(p)$, which from algebra follows $\frac{1}{2} f'(p) < f'(s) < \frac{3}{2} f'(p)$. Indeed, since $0 < \frac{1}{2} f'(p)$ and $\frac{1}{2} f'(p) < f'(s)$, then $f'(s) > 0$.

We have $f'(s) > 0$ and since $y > x$ then $y - x > 0$, so $f'(s)(y - x) > 0$. Now $f'(s)(y - x) = f(y) - f(x)$ and $f'(s)(y - x) > 0$, so $f(y) - f(x) > 0$ and in particular $f(y) > f(x)$. We have show for arbitrary $x, y \in \mathcal{U}$ with $x < y$, that $f(x) < f(y)$. This is to say that is strictly increasing on $\mathcal{U}$ or is injective with domain $\mathcal{U}$ and range $\mathcal{V}$. Since $\mathcal{V} = f(\mathcal{U})$, then is surjective. Finally since $f : \mathcal{U} \longrightarrow \mathcal{V}$ is a bijection, then there exists a function $f^{-1} : \mathcal{V} \longrightarrow \mathcal{U}$ so that $f^{-1}(\mathcal{V}) = \mathcal{U}$.

It remains to show that f^{-1} is continuously differentiable on $\mathcal{V}$. Let $q \in \mathcal{V}$, then there is $p \in \mathcal{U}$ such that $f(p) = q$. Since $\lim_{x \to p} \left| \frac{f(x) - f(p)}{x - p} - f'(p) \right| = 0$ and for each $x \in \mathcal{U}$, $f(x) \neq f(p)$, then theorems from analysis give, $\lim_{x \to p} \left| \frac{x - p}{f(x) - f(p)} - \frac{1}{f'(p)} \right| = 0$. Since $\lim_{x \to p} \left| \frac{x - p}{f(x) - f(p)} - \frac{1}{f'(p)} \right| = 0$ and is continuous, then $\lim_{y \to q} \left| \frac{f^{-1}(y) - f^{-1}(q)}{y - q} - \frac{1}{f'(f^{-1}(q))} \right| = 0$. This established the theorem for the special case of a function of a real variable. ∎

The generalization of our simple case is hardly trivial. In fact, about the only aspect that maintains a semblance of the previous case is the statement of the theorem. Therein change each $\mathbb{R}$ to $\mathbb{R}^n$ and the generalized notion of differentiable requires $f'(p) \neq 0$ to become $\det[f'(p)] \neq 0$. We first prove a lemma.

Theorem 2.15 Let $A \subset \mathbb{R}^n$ be a rectangle and let $f : A \longrightarrow \mathbb{R}^n$ be continuously differentiable. If there is a number M such that for $i, j \in \{1, 2, \ldots, n\}$, $|D_j f^i(\boldsymbol{x})| \leq M$ for all $\quad \in \mathrm{Int}(A)$, then $|f(\boldsymbol{x}) - f(\boldsymbol{y})| \leq n^2 M |\boldsymbol{x} - \boldsymbol{y}|$ for all $\quad , \boldsymbol{y} \in A$.

Proof. Suppose there is a number M such that for $i, j \in \{1, 2, \ldots, n\}$, $|D_j f^i(\boldsymbol{x})| \leq M$ for all $\boldsymbol{x} \in \mathrm{Int}(A)$. To more clearly explain the proof, we prove the theorem for $n = 3$ before giving a more concise proof of the general case.

Let each of $\boldsymbol{p}$ and $\boldsymbol{q}$ be elements of the interior of A. Without loss of generality we will assume $p^i < q^i$ for each $i \in \{1, 2, 3\}$ allowing us to write $[p^i, q^i]$ without ambiguity. Recall for each $\boldsymbol{x} \in A$ we can write

$$f(\boldsymbol{x}) = \begin{pmatrix} f^1(\boldsymbol{x}) \\ f^2(\boldsymbol{x}) \\ f^3(\boldsymbol{x}) \end{pmatrix}$$

for unique functions $f^i : \mathbb{R}^3 \longrightarrow \mathbb{R}$ with $i \in \{1, 2, 3\}$, and since $Df(\boldsymbol{x})$ exists for each $\boldsymbol{x} \in \mathrm{Int}(A)$ then by Theorem 2.12 $Df(\boldsymbol{x})$ has the Jacobian Matrix form

$$f'(\boldsymbol{x}) = \begin{pmatrix} D_1 f^1(\boldsymbol{x}) & D_2 f^1(\boldsymbol{x}) & D_3 f^1(\boldsymbol{x}) \\ D_1 f^2(\boldsymbol{x}) & D_2 f^2(\boldsymbol{x}) & D_3 f^2(\boldsymbol{x}) \\ D_1 f^3(\boldsymbol{x}) & D_2 f^3(\boldsymbol{x}) & D_3 f^3(\boldsymbol{x}) \end{pmatrix} = (D_j f^i(\boldsymbol{x})).$$

For each $i \in \{1, 2, 3\}$, we can write $f^i(\boldsymbol{q}) - f^i(\boldsymbol{p})$ in an equivalent manner that defines three real valued functions on $[p^i, q^i] \subset \mathbb{R}$ giving us the power to apply the mean value theorem. Note

$$\begin{aligned} f^1(\boldsymbol{q}) - f^1(\boldsymbol{p}) &= f^1(q^1, p^2, p^3) - f^1(p^1, p^2, p^3) \\ &\quad + f^1(q^1, q^2, p^3) - f^1(q^1, p^2, p^3) \\ &\quad + f^1(q^1, q^2, q^3) - f^1(q^1, q^2, p^3) \end{aligned}$$

since the first term of each difference has its additive inverse in the following difference, the only exception is the last difference. What is left is precisely the left-hand-side. There are three functions $f^1(\,\cdot\,, p^2, p^3) : [p^1, q^1] \longrightarrow \mathbb{R}$, $f^1(q^1, \,\cdot\,, p^3) : [p^2, q^2] \longrightarrow \mathbb{R}$, and $f^1(q^1, q^2, \,\cdot\,) : [p^3, q^3] \longrightarrow \mathbb{R}$, and for each $i \in \{1, 2, 3\}$

and each $x \in (p^i, q^i)$, each of $D_1 f^1(x, p^2, p^3)$, $D_2 f^1(q^1, x, p^3)$, and $D_3 f^1(q^1, q^2, x)$ exists. We can therefore apply the mean value theorem, applicable only to functions from $\mathbb{R}$ to $\mathbb{R}$, and choose $z_{11} \in [p^1, q^1]$, $z_{12} \in [p^2, q^2]$, and $z_{13} \in [p^3, q^3]$ so that

$$|f^1(q^1, p^2, p^3) - f^1(p^1, p^2, p^3)| = |q^1 - p^1|\,|D_1 f^1(z_{11}, p^1, p^3)|,$$
$$|f^1(q^1, q^2, p^3) - f^1(q^1, p^2, p^3)| = |q^2 - p^2|\,|D_2 f^1(q^1, z_{12}, p^3)|, \text{ and}$$
$$|f^1(q^1, q^2, q^3) - f^1(q^1, q^2, p^3)| = |q^3 - p^3|\,|D_3 f^1(q^1, q^2, z_{13})|.$$

Since $|D_j f^i(\boldsymbol{x})| \le M$ for all $\boldsymbol{x} \in \mathrm{Int}(A)$, then

$$|f^1(q^1, p^2, p^3) - f^1(p^1, p^2, p^3)| \le |q^1 - p^1|\,M,$$
$$|f^1(q^1, q^2, p^3) - f^1(q^1, p^2, p^3)| \le |q^2 - p^2|\,M, \text{ and}$$
$$|f^1(q^1, q^2, q^3) - f^1(q^1, q^2, p^3)| \le |q^3 - p^3|\,M.$$

From earlier considerations, without the absolute value, each of the differences on the left adds to precisely $f^1(\boldsymbol{q}) - f^1(\boldsymbol{p})$, then $|f^1(\boldsymbol{q}) - f^1(\boldsymbol{p})| \le \sum_{i=1}^{3} |q^i - p^i|\,M$ and since $|q^i - p^i| \le \|\boldsymbol{q} - \boldsymbol{p}\|$ then $\sum_{i=1}^{3} |q^i - p^i|\,M \le 3M\,\|\boldsymbol{q} - \boldsymbol{p}\|$ and so $|f^1(\boldsymbol{q}) - f^1(\boldsymbol{p})| \le 3M\,\|\boldsymbol{q} - \boldsymbol{p}\|$. The same conclusion can be drawn for f^2 and f^3. From Theorem 1.1 we have $\|f(\boldsymbol{q}) - f(\boldsymbol{p})\| \le \sum_{i=1}^{3} |f^i(\boldsymbol{q}) - f^i(\boldsymbol{p})| \le \sum_{i=1}^{3} 3M\,\|\boldsymbol{q} - \boldsymbol{p}\|$. Since there are three terms in this sum, then $\|f(\boldsymbol{q}) - f(\boldsymbol{p})\| \le 3^2 M\,\|\boldsymbol{q} - \boldsymbol{p}\|$.

The general case follows similarly. Let $i \in \{1, 2, \dots, n\}$. We write

$$f^i(\boldsymbol{q}) - f^i(\boldsymbol{p}) = \sum_{j=1}^{n} f^i(q^1, \dots, q^j, p^{j+1}, \dots, p^n) - f^i(q^1, \dots, q^{j-1}, p^j, \dots, p^n) \qquad (1)$$

where the first term of each difference has its additive inverse in the difference that follows in the sum; the only exception is the last difference. Just as in the example, we have n functions from a compact and connected subset of $\mathbb{R}$ to $\mathbb{R}$. We can therefore apply the mean value theorem and choose z_{ij} in $[p^j, q^j]$ or $[q^j, p^j]$, which ever makes sense, to conclude

$$\begin{aligned}|f^i(q^1, \dots, q^j, p^{j+1}, \dots, p^n) - f^i(q^1, \dots, q^{j-1}, p^j, \dots, p^n)| \\ = |q^j - p^j|\,|D_j f^i(q^1, \dots, q^{j-1}, z_{ij}, p^{j+1}, \dots, p^n)|.\end{aligned} \qquad (2)$$

Since $|D_j f^i(\boldsymbol{x})| \le M$ for all $\boldsymbol{x} \in \text{Int}(A)$, then from (2)

$$|f^i(q^1, \ldots, q^j, p^{j+1}, \ldots, p^n) - f^i(q^1, \ldots, q^{j-1}, p^j, \ldots, p^n)| \le |q^j - p^j|\, M. \tag{3}$$

From (1) and (3), we have $|f^i(\boldsymbol{q}) - f^i(\boldsymbol{p})| \le \sum_{j=1}^{n} |q^j - p^j|\, M$, and since $|q^j - p^j| \le \|\boldsymbol{q} - \boldsymbol{p}\|$ for each $j \in \{1, 2, \ldots, n\}$, then

$$|f^i(\boldsymbol{q}) - f^i(\boldsymbol{p})| \le nM\, \|\boldsymbol{q} - \boldsymbol{p}\|. \tag{4}$$

Finally, from Theorem 1.1, $\|f(\boldsymbol{q}) - f(\boldsymbol{p})\| \le \sum_{i=1}^{n} |f^i(\boldsymbol{q}) - f^i(\boldsymbol{p})|$, then with (4) we establish the result $\|f(\boldsymbol{q}) - f(\boldsymbol{p})\| \le n^2 M\, \|\boldsymbol{q} - \boldsymbol{p}\|$. ∎

Theorem 2.16 Suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a continuously differentiable function on some open subset O of $\mathbb{R}^n$. If $\boldsymbol{p} \in O$ and $\det[f'(\boldsymbol{p})] \neq 0$, then there exists open sets $\mathcal{U}$ and $\mathcal{V}$ of $\mathbb{R}^n$ so that:

1. $\boldsymbol{p} \in \mathcal{U}$,
2. $f(\mathcal{U}) = \mathcal{V}$,
3. $f^{-1} : \mathcal{V} \longrightarrow \mathcal{U}$ exists and $f^{-1}(\mathcal{V}) = \mathcal{U}$, and
4. $(f^{-1})'(\boldsymbol{q}) = [f' \circ f^{-1}(\boldsymbol{q})]^{-1}$ for each $\boldsymbol{q} \in \mathcal{V}$.

Proof Let $\boldsymbol{p} \in O$ such that $\det[f'(\boldsymbol{p})] \neq 0$. We will call the linear transformation $f'(\boldsymbol{p})$ by λ. To further simplify the situation we note that λ^{-1} exists since $\det[f'(\boldsymbol{p})] \neq 0$ and by the chain rule $D(\lambda^{-1} \circ f)(\boldsymbol{p}) = D(\lambda^{-1})(f(\boldsymbol{p})) \circ Df(\boldsymbol{p}) = \lambda^{-1} \circ Df(\boldsymbol{p}) = \lambda^{-1} \circ \lambda$. Since the theorem being true for $\lambda^{-1} \circ f$ will imply the theorem is true for f, since λ^{-1} is a linear transformation, then we can assume that $Df(\boldsymbol{p})$ is the identity transformation.

First we show there exists a closed rectangle $\mathcal{W}_1$ with $\boldsymbol{p} \in \text{Int}(\mathcal{W}_1)$ so that for each $\boldsymbol{x} \in \mathcal{W}_1$ with $\boldsymbol{x} \neq \boldsymbol{p}$, $f(\boldsymbol{x}) \neq f(\boldsymbol{p})$.

Assume for each closed rectangle $\mathcal{W}$ with $\boldsymbol{p} \in \text{Int}(\mathcal{W})$, there is an $\boldsymbol{x} \in \mathcal{W}_1$ with $\boldsymbol{x} \neq \boldsymbol{p}$ so that $f(\boldsymbol{x}) = f(\boldsymbol{p})$. By the assumption, for each $i \in \mathbb{N}$, pick $\boldsymbol{x}_i \in [p^1 - \frac{1}{i}, p^1 + \frac{1}{i}] \times \ldots \times [p^n - \frac{1}{i}, p^n + \frac{1}{i}]$ different from $\boldsymbol{p}$ so that $f(\boldsymbol{x}_i) = f(\boldsymbol{p})$. We have then a sequence $\{\boldsymbol{x}_i\}_{i=1}^{\infty}$ such that $\lim_{i\to\infty} \boldsymbol{x}_i = \boldsymbol{p}$.

Now since each of f and λ is continuous at $\boldsymbol{p}$ and since $\lim_{i\to\infty} \boldsymbol{x}_i = \boldsymbol{p}$, then

$\lim_{i\to\infty} f(x_i) = f(p)$ and $\lim_{i\to\infty} \lambda(x_i) = \lambda(p)$. Since the norm is also a continuous operator then $\lim_{i\to\infty} \frac{\|f(p+x_i-p)-f(p)-\lambda(p-x_i)\|}{\|x_i-p\|} = \lim_{i\to\infty} \frac{\|\lambda(p-x_i)\|}{\|x_i-p\|}$. Now since $f'(p)$ exists then the limit on the left hand side is exactly zero, so we have $\lim_{i\to\infty} \frac{\|\lambda(p-x_i)\|}{\|x_i-p\|} = 0$. Now since for each $i \in \mathbb{N}$, $\frac{1}{\|x_i-p\|} \in \mathbb{R}$ and since λ is linear, then $\lim_{i\to\infty} \left\|\lambda\left(\frac{p-x_i}{\|x_i-p\|}\right)\right\| = 0$.

Define the sequence $\{z_i\}_{i=1}^{\infty}$ for each $i \in \mathbb{N}$ by $z_i \in \frac{-\ i}{\|x_i-p\|}$. Since for each $i \in \mathbb{N}$, $\|z_i\| = 1$, then $\{z_i\}_{i=1}^{\infty}$ is a sequence in the compact set $B = \{x \in \mathbb{R}^n \mid 1 = |x|\}$, thus there is a subsequence $\{z_{i_j}\}_{j=1}^{\infty}$ such that $\lim_{j\to\infty} z_{i_j}$ exists, call it z. Since B is closed and for each $j \in \mathbb{N}$, $z_{i_j} \in B$, then $z \in B$. Since $z \in B$, then $|z| = 1$ and hence $z \neq \mathbf{0}$.

Since $\lim_{j\to\infty} z_{i_j} =\ $, $\lim_{j\to\infty} \|\lambda(z_{i_j})\| = 0$, and λ is continuous, then $\lambda(z) = \mathbf{0}$. Since λ is a linear transformation then $\lambda(z) = \mathbf{0}$ implies that $z = \mathbf{0}$. We have shown then that $z = \mathbf{0}$ at the same time $z \neq \mathbf{0}$. Thus our assumption must be false, and we conclude its negation.

There exists a closed rectangle $\mathcal{W}_1$ with $p \in \mathrm{Int}(\mathcal{W}_1)$ so that for each $x \in \mathcal{W}_1$ with $x \neq p$, $f(x) \neq f(p)$.

Next we show there exists a closed rectangle $\mathcal{W}_2$ with $p \in \mathrm{Int}(\mathcal{W}_2)$ so that for each $x \in \mathcal{W}_2$ $\det[f'(x)] \neq 0$.

We use that $\det : \mathbb{R}^n \times \ldots n \text{ times} \ldots \times \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous and that $f'(x)$ exists for each $x \in O$. Since $\det[f'(p)] \neq 0$ then $\frac{1}{2}\,|\det[f'(p)]| > 0$. Since det is continuous and $\frac{1}{2}\,|\det[f'(p)]| > 0$ there is a number $\delta > 0$ so that if $x \in O$ and $\|x - p\| < \delta$, then $|\det[f'(x)] - \det[f'(p)]| < \frac{1}{2}\,|\det[f'(p)]|$ and in particular $\det[f'(x)] \neq 0$.

We pick $\mathcal{W}_2 = \left\{x \in O \,\middle|\, |p^i - x^i| \leq \frac{\delta}{\sqrt{2n}} \text{ for each } i = 1, 2, \ldots, n\right\}$. This is also written as the rectangle $\left[p^i - \frac{\delta}{\sqrt{2n}}, p^i + \frac{\delta}{\sqrt{2n}}\right] \times \ldots \times \left[p^n - \frac{\delta}{\sqrt{2n}}, p^n + \frac{\delta}{\sqrt{2n}}\right]$. Finally, if $x \in \mathcal{W}_2$, then $\|p - x\| = \sqrt{\sum_{i=1}^{n} (p^i - x^i)^2}$ and since for each $i \in \{1, 2, \ldots, n\}$, $|p^i - x^i| < \frac{\delta}{\sqrt{2n}}$, then $\sqrt{\sum_{i=1}^{n} (p^i - x^i)^2} \leq \sqrt{\sum_{i=1}^{n} \left(\frac{\delta}{\sqrt{2n}}\right)^2} = \frac{\delta}{2} < \delta$, so $\|p - x\| < \delta$ and $\det[f'(x)] \neq 0$. We have established a closed rectangle $\mathcal{W}_2$ with $p \in \mathrm{Int}(\mathcal{W}_2)$ so that for each $\ \in \mathcal{W}_2$ $\det[f'(x)] \neq 0$

Next we will show there exists a closed rectangle $\mathcal{W}_3$ with $\boldsymbol{p} \in \text{Int}(\mathcal{W}_3)$ so that for each $\boldsymbol{x} \in \mathcal{W}_3$ $|D_j f^i(\boldsymbol{x}) - D_j f^i(\boldsymbol{p})| < \frac{1}{2n^2}$ for each $i, j \in \{1, 2, \ldots, n\}$.

We use that is continuously differentiable on O. Since the notion of continuous is independent of the norm given to a space, for each $\boldsymbol{x} \in O$, we define $\|f'(\boldsymbol{x})\| = \sum_{i=1}^{n} (\sum_{j=1}^{n} |D_j f^i(\boldsymbol{x})|)$ where $D_j f^i$ is the entry in the ith row, jth column of $f'(\boldsymbol{x})$. Since $\frac{1}{2n^2} > 0$ and is continuously differentiable, there is a number $\delta > 0$ so that if $\boldsymbol{x} \in O$ with $\|\boldsymbol{x} - \boldsymbol{p}\| < \delta$, then $\|f'(\boldsymbol{x}) - f'(\boldsymbol{p})\| = \sum_{i=1}^{n} (\sum_{j=1}^{n} |D_j f^i(\boldsymbol{x}) - D_j f^i(\boldsymbol{p})|) < \frac{1}{2n^2}$. Since this is a finite sum of non-negative numbers, then each term must be smaller than $\frac{1}{2n^2}$. Using δ we pick $\mathcal{W}_3$ in the identical way we chose $\mathcal{W}_2$. Let $\mathcal{W}_3 = \left\{\boldsymbol{x} \in O \,\middle|\, |p^i - x^i| \leq \frac{\delta}{\sqrt{2}\,n} \text{ for each } i = 1, 2, \ldots, n\right\}$. As before, if $\boldsymbol{x} \in \mathcal{W}_3$, then $\|\boldsymbol{x} - \boldsymbol{p}\| < \delta$ and by the choice of δ, then $|D_j f^i(\boldsymbol{x}) - D_j f^i(\boldsymbol{p})| < \frac{1}{2n^2}$ for each $i, j \in \{1, 2, \ldots, n\}$. We have chosen a closed rectangle $\mathcal{W}_3$ with $\boldsymbol{p} \in \text{Int}(\mathcal{W}_3)$ so that for each $\boldsymbol{x} \in \mathcal{W}_3$ $|D_j f^i(\boldsymbol{x}) - D_j f^i(\boldsymbol{p})| < \frac{1}{2n^2}$ for each $i, j \in \{1, 2, \ldots, n\}$.

Let $\mathcal{W} = \mathcal{W}_1 \cap \mathcal{W}_2 \cap \mathcal{W}_3$. We write $\mathcal{W} = [w_{1\,l}, w_{1\,r}] \times \ldots \times [w_{n\,l}, w_{n\,r}]$ is a closed rectangle.

Next, if each of $_1$ and $\boldsymbol{x}_2$ is an element of $\mathcal{W}$, then we will show that $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| \leq 2\,\|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)\|$.

Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{W}$. By the simplification made at the start of this proof, we have λ as the identity map. If for each $\boldsymbol{x} \in \mathcal{W}$, we define $g(\boldsymbol{x}) = f(\boldsymbol{x}) - \boldsymbol{x}$, then g is continuously differentiable on $\mathcal{W}$, and $|D_j g^i(\boldsymbol{x})| = |D_j f^i(\boldsymbol{x}) - \delta_j^i| = |D_j f^i(\boldsymbol{x}) - D_j f^i(\boldsymbol{p})|$. Were δ_j^i is the Kronecker Delta. If $\boldsymbol{x} \in \mathcal{W}$ then $\in \mathcal{W}_3$ so $|D_j f^i(\boldsymbol{x}) - D_j f^i(\boldsymbol{p})| < \frac{1}{2n^2}$. Thus if $\boldsymbol{x} \in \mathcal{W}$, $|D_j g^i(\boldsymbol{x})| < \frac{1}{2n^2}$.

By Theorem 2.15, we conclude that $\|g(\boldsymbol{x}_2) - g(\boldsymbol{x}_2)\| \leq n^2 \frac{1}{2n^2} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$ which simplifies to $\|g(\boldsymbol{x}_2) - g(\boldsymbol{x}_2)\| \leq \frac{1}{2} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$. Now by a property of the norm we have $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| - \|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)\| \leq \|f(\boldsymbol{x}_1) - \boldsymbol{x}_1 - (f(\boldsymbol{x}_2) - \boldsymbol{x}_2)\|$. Since $g(\boldsymbol{x}) = f(\boldsymbol{x}) - \boldsymbol{x}$, then $\|f(\boldsymbol{x}_1) - \boldsymbol{x}_1 - (f(\boldsymbol{x}_2) - \boldsymbol{x}_2)\| = \|g(\boldsymbol{x}_1) - g(\boldsymbol{x}_2)\|$. Altogether then $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| - \|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)\| \leq \frac{1}{2} \|\boldsymbol{x}_2 - \boldsymbol{x}_1\|$, which transforms with algebra to

$\|x_1 - x_2\| \leq 2\,\|f(x_1) - f(x_2)\|$.

We have shown for each of x_1 and x_2 an element of $\mathcal{W}$, then
$\|x_1 - x_2\| \leq 2\,\|f(x_1) - f(x_2)\|$

We now pick our image set $\mathcal{V}$. Since $\mathcal{W}$ is a rectangle then $\mathcal{W}$ is compact and since the boundary of $\mathcal{W}$, call it $B(\mathcal{W})$, is a closed subset in $\mathcal{W}$ (since it is the complement of the Int($\mathcal{W}$) which is the largest open subset of $\mathcal{W}$) then $B(\mathcal{W})$ is compact. Since is continuous, then $f(B(\mathcal{W}))$ is compact. Furthermore since $f(B(\mathcal{W}))$ is compact, then it is closed. Since $p \in \text{Int}(\mathcal{W})$, then $p \notin B(\mathcal{W})$, so $f(p) \notin f(B(\mathcal{W}))$. Finally since $f(B(\mathcal{W}))$ is compact and $f(p) \notin f(B(\mathcal{W}))$, then there is a number $\delta > 0$ such that $\delta < \inf\{\|f(p) - f(x)\| \mid x \in B(\mathcal{W})\}$. We pick $\mathcal{V} = \{y \mid \|y - f(p)\| < \frac{\delta}{2}\}$.

Next we show if $y \in \mathcal{V}$ and $x \in B(\mathcal{W})$, then $\|y - f(p)\| < \|y - f(x)\|$.

Let $y \in \mathcal{V}$ and $x \in B(\mathcal{W})$. Since $y \in \mathcal{V}$, then $\|y - f(p)\| < \frac{\delta}{2}$. Now since $\delta < \inf\{\|f(p) - f(x)\| \mid x \in B(\mathcal{W})\}$, then $\delta < \|f(x) - f(p)\|$ and $\|f(x) - f(p)\| - \|y - f(p)\| > \delta - \frac{\delta}{2} = \frac{\delta}{2}$. By properties of the usual norm, $\|f(x) - f(p)\| - \|f(p) - y\| \leq \|f(x) - y\|$. Together then $\|y - f(p)\| < \|y - f(x)\|$.

Now we show if $y \in \mathcal{V}$, there is a unique $x \in \mathcal{W}$ such that $f(x) = y$.

Let $y \in \mathcal{V}$. Define $g : \mathcal{W} \longrightarrow \mathbb{R}$ by $g(x) = \|y - f(x)\|^2 = \sum_{i=1}^{n} (y^i - f^i(x))^2$. g is continuous since is and therefore g is bounded on $\mathcal{W}$ and moreover there exists $x \in \mathcal{W}$ so that $g(x) = \inf\{\text{range of } g\}$.

Assume $x \in B(\mathcal{W})$, then since $g(p) = \|y - f(p)\|^2$, $g(x) = \|y - f(x)\|^2$, then by above we have $g(p) < g(x)$. This contradicts $g(x)$ being the minimum. The only alternative then is that $x \in \text{Int}(\mathcal{W})$.

Since $\mathcal{W} \subset \mathbb{R}^n$, the $\inf\{\text{range of } g\} = g(x)$ occurs at $x \in \text{Int}(\mathcal{W})$, and $D_j\, g(x)$ exists for each $j \in \{1, 2, \ldots, n\}$, then by Theorem 2.11 $D_j\, g(x) = 0$. By Theorem 2.3, it follows that $_j\, g(x) = \sum_{i=1}^{n} 2\,(y^i - f^i(x))\, D_j f^i(x) = 0$ for each $j \in \{1, 2, \ldots, n\}$. We know from that $\mathcal{W}$ was chosen so that $\det(D_j f^i(x)) \neq 0$; thus, $(y^i - f^i(x)) = 0$ for each $i \in \{1, 2, \ldots, n\}$. Since $y^i = f^i(x)$ for each $i \in \{1, 2, \ldots, n\}$, then $y = f(x)$. If $x_1, x_2 \in \mathcal{W}$

such that $y = f(x_1) = f(x_2)$, then from above $\|x_1 - x_2\| \leq 2\,\|f(x_1) - f(x_2)\| = 0$. By property of norm, it follows that $x_1 - x_2 = 0$ so $x_1 = x_2$ and we have established the claim.

We are now ready to choose $\mathcal{U}$. Since for each $y \in \mathcal{V}$, there is a unique $x \in \mathcal{W}$, then we can construct a well-define a function $\ ^{-1} : \mathcal{V} \longrightarrow \mathcal{W}$. Pick $\mathcal{U} = \{x \in \operatorname{Int}(\mathcal{W}) \mid f^{-1}(y) = x \text{ for some } y \in \mathcal{V}\}$. We can show the first three conclusions of our theorem.

First, $p \in \mathcal{U}$ since $p \in \operatorname{Int}(\mathcal{W})$ and $f(p) \in \mathcal{V}$.

Second, if $x \in \mathcal{U}$, then $x = f^{-1}(y)$ for some $y \in \mathcal{V}$ so $f(\mathcal{U}) \subseteq \mathcal{V}$. Now if $y \in \mathcal{V}$, then $y = f(x)$ for some $x \in \mathcal{U}$, so $\mathcal{V} \subseteq f(\mathcal{U})$. The dual inclusions gives $\ (\mathcal{U}) = \mathcal{V}$.

Third, if $y \in \mathcal{V}$, then $f^{-1}(y) = x$ for some $x \in \mathcal{U}$, so $f^{-1}(\mathcal{V}) \subseteq \mathcal{U}$. If $x \in \mathcal{U}$ then $f(x) = y$ for some $y \in \mathcal{V}$. Since $f^{-1}(y) \in f^{-1}(\mathcal{V})$ and $x = f^{-1}(f(x)) = f^{-1}(y)$, then $x \in f^{-1}(\mathcal{V})$ so we have $f^{-1}(\mathcal{V}) = \mathcal{U}$.

Finally we show if $t \in \mathcal{V}$, then $(f^{-1})'(t)$ exists and is $[f'(f^{-1}(t))]^{-1}$.

Let $t \in \mathcal{V}$. Pick $s \in \mathcal{U}$ such that $f(s) = t$. Since is continuously differentiable on $\mathcal{U}$ and $s \in \mathcal{U}$, then call $\mu = f'(s)$. Since $s \in \mathcal{U} \subseteq \mathcal{W}_2$, then $\det[f'(s)] \neq 0$, so we also have the inverse linear transformation μ^{-1}. Since $f'(s)$ exists, we have $\lim\limits_{x \to s} \frac{\|\ (x) -\ (s) - \mu(x-s)\|}{\|s-x\|} = 0$.

Define $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by $\varphi(h) = f(s+h) - f(s) - \mu(h)$. For each x, $x - s \in \mathbb{R}^n$ so $\varphi(x-s) = f(x) - f(s) - \mu(x-s)$. We apply μ^{-1} to arrive at $\mu^{-1}(\varphi(x-s)) = \mu^{-1}(f(x) - f(s)) - (x-s)$. Since we have $f(x) = y$ for each $x \in \mathcal{U}$ and $f^{-1}(y) = x$ for each $y \in \mathcal{V}$, then $\mu^{-1}(\varphi(f^{-1}(y) - f^{-1}(t)))$
$= \mu^{-1}(y-t) - (f^{-1}(y) - f^{-1}(t))$.

To establish our claim, we need to show that $\lim\limits_{y \to t} \frac{\|\mu^{-1}(y-t) - (f^{-1}(y) - f^{-1}(t))\|}{\|y-t\|} = 0$. We can do this by showing $\lim\limits_{y \to t} \frac{\|\mu^{-1}(\varphi(f^{-1}(y) - f^{-1}(t)))\|}{\|y-t\|} = 0$ since $\mu^{-1}(\varphi(f^{-1}(y) - f^{-1}(t))) = \mu^{-1}(y-t) - (f^{-1}(y) - f^{-1}(t))$. Since μ^{-1} is a linear transformation, then by the chain rule, we show $\lim\limits_{y \to t} \frac{\|\varphi(f^{-1}(y) - f^{-1}(t))\|}{\|y-t\|} = 0$.

Now $\frac{\|\varphi(f^{-1}(y) - f^{-1}(t))\|}{\|y-t\|} = \frac{\|\varphi(f^{-1}(y) - f^{-1}(t))\|}{\|f^{-1}(y) - f^{-1}(t)\|} \cdot \frac{\|f^{-1}(y) - f^{-1}(t)\|}{\|y-t\|}$ for each $y \in \mathcal{V}$. Since

$\lim\limits_{x \to s} \frac{\|\varphi(x-s)\|}{\|x-s\|} = 0$ and since f^{-1} is continuous, then $\lim\limits_{y \to t} \frac{\|\varphi(f^{-1}(y)-f^{-1}(t))\|}{\|y-t\|} = 0$. By conclusions above, $\|x - s\| \leq 2\,\|f(x) - f(s)\|$ for each $x \in \mathcal{U}$, thus $\frac{\|x-s\|}{\|f(x)-f(s)\|} \leq 2$ for each $x \in \mathcal{U}$ except s. Since f is continuous and for each $y \in \mathcal{V}$, $\frac{\|f^{-1}(y)-f^{-1}(t)\|}{\|y-t\|} \leq 2$, then $\lim\limits_{y \to t} \frac{\|f^{-1}(y)-f^{-1}(t)\|}{\|y-t\|}$ exists call it γ. Now since each of $\lim\limits_{y \to t} \frac{\|\varphi(f^{-1}(y)-f^{-1}(t))\|}{\|y-t\|}$ and $\lim\limits_{y \to t} \frac{\|f^{-1}(y)-f^{-1}(t)\|}{\|y-t\|}$ exists, then $\lim\limits_{y \to t} \frac{\|\varphi(f^{-1}(y)-f^{-1}(t))\|}{\|f^{-1}(y)-f^{-1}(t)\|} \cdot \frac{\|f^{-1}(y)-f^{-1}(t)\|}{\|y-t\|}$ exists and is $0 \cdot \gamma = 0$. This shows $\lim\limits_{y \to t} \frac{\|\mu^{-1}(y-t)-(f^{-1}(y)-f^{-1}(t))\|}{\|y-t\|} = 0$ and we have established the final claim. ∎

Integration

We begin our discussion of integration with functions of real variables. We press forward to establish the Fundamental Theorem of Calculus since it is the analogy of Stokes' Theorem in the setting of real numbers and since it is used in the proof of Stokes' Theorem. Unlike differentiation, integration on rectangles in $\mathbb{R}^n$ is quite similar to $\mathbb{R}$. We give only the definitions and prove Fubini's Theorem.

Definition 2.5 A *subdivision* of the interval $[a, b]$ is a finite subset of $[a, b]$ denoted by D with the property that each of a and b belong to D. We usually denote D by $\{x_i\}_{i=0}^{n}$, where $x_0 = a$, $x_n = b$, and for each index i between 1 and n inclusive, $x_{i-1} \leq x_i$.

Definition 2.6 A *refinement* K of a subdivision of the interval $[a, b]$ is a subdivision of $[a, b]$ where $\subseteq K$.

Theorem 2.17 If D is a subdivision of $[a, b]$, then D is a refinement of D.

Theorem 2.18 If K is a refinement of H and H is a refinement of the subdivision D of $[a, b]$, then K is a refinement of .

Theorem 2.19 If each of D_1 and D_2 is a subdivision of $[a, b]$, then $D_1 \bigcup D_2$ is a subdivision of $[a, b]$ and $D_1 \bigcup D_2$ is a refinement of D_1.

Definition 2.7 If $D = \{x_i\}_{i=0}^{n}$ is a subdivision of $[a, b]$, $\{t_i\}_{i=1}^{n}$ is an interpolation sequence of D if for each $i \in \{1, 2, \ldots, n\}$, $x_{i-1} \leq t_i \leq x_i$.

Definition 2.8 For a function $f: [a, b] \longrightarrow \mathbb{R}$, if there exists a number A such that for each $\varepsilon > 0$ there is a subdivision D of $[a, b]$ such that if $H = \{x_i\}_{i=0}^{n}$ is a refinement of and $\{t_i\}_{i=1}^{n}$ is an interpolation sequence of H, $\left|\sum_{i=1}^{n} f(t_i) \triangle x_i - A\right| < \varepsilon$, then we say f is *A-integrable* or simply *integrable* on $[a, b]$ and we denote the number A by $\int_a^b f\, dx$.

Theorem 2.20 If f is a function defined on $[a, b]$ such that $\int_a^b f\, dj$ exists, then f is bounded on $[a, b]$.

Theorem 2.21 If each of f and g is a function defined on $[a, b]$ such that $\int_a^b f\, dj$ exists and $\int_a^b g\, dj$ exists, then $\int_a^b f + g\, dj$ exists.

Theorem 2.22 If f is a function defined on $[a, b]$, c is a number such that $a < c < b$, and each $\int_a^c f$ exists and $\int_c^b f$ exists, then $\int_a^b f$ exists.

Definition 2.9 Suppose f is a function defined on $[a, b]$ such that $\int_a^b f$ exists. We define $\int_b^a f = -\int_a^b f$.

Theorem 2.23 If $f: [a, b] \longrightarrow \mathbb{R}$ is a continuous function and k is a number such that $f(a) \leq k \leq f(b)$, then there is a number $p \in [a, b]$ such that $f(p) = k$.

Theorem 2.24 Suppose f is a function defined on $[a, b]$. The following statements are equivalent:

(a) $\int_a^b f\, dj$

(b) If $\varepsilon > 0$ then there exists a subdivision $D = \{x_i\}_{i=0}^{n}$ of $[a, b]$ such that if $H = \{y_i\}_{i=0}^{m}$ is a refinement of D and $\{t_i\}_{i=1}^{n}$ and $\{s_i\}_{i=1}^{m}$ are interpolating sequences for D and H, respectively, then $|\sum_{i=1}^{n} f(t_i) \triangle x_i - \sum_{i=1}^{m} f(s_i) \triangle y_i| < \varepsilon$.

(c) If $\varepsilon > 0$ then there exists a subdivision of $[a, b]$ such that if each of $H = \{x_i\}_{i=0}^{n}$ and $K = \{y_i\}_{i=0}^{m}$ is a refinement of and $\{t_i\}_{i=1}^{n}$ and $\{s_i\}_{i=1}^{m}$ are interpolating sequences for H and K, respectively, then $|\sum_{i=1}^{n} f(t_i) \triangle x_i - \sum_{i=1}^{m} f(s_i) \triangle y_i| < \varepsilon$.

Theorem 2.25 Suppose f is a function defined on $[a, c]$ such that $\int_a^c f\, dj$ exists. If b is a number such that $a < b < c$, then $\int_a^b f\, dj$ exists.

Theorem 2.26 Suppose f is a function defined on $[a, b]$ such that $\int_a^b f\, dj$ exists. If k is a number such that for each $x \in [a, b]$ $g(x) = k \cdot f(x)$, then g is a function defined on $[a, b]$ such that $\int_a^b g\, dj = k \cdot \int_a^b f\, dj$.

Theorem 2.27 Suppose $f : [a, b] \to \mathbb{R}$. If f is continuous, then $\int_a^b f$ exists.

Theorem 2.28 Suppose $f : [a, b] \to \mathbb{R}$. If f is continuous, then there is an $x \in [a, b]$ such that $f(x)\,(b - a) = \int_a^b f$.

Theorem 2.29 Suppose $f : [a, b] \to \mathbb{R}$. If f is continuous, $a < p < b$, and for each $x \in (a, b)$, $g(x) = \int_p^x f\, dt$, then for each $x \in (a, b)$, $g'(x)$ exists and is $f(x)$.

Proof. Let f be a continuous function. Since f is continuous on $[a, b]$, then by Theorem 2.27 $\int_a^b f\, dt$ exists. Let p be a number such that $a < p < b$. Since for each $x \in (a, p)$, f is continuous on $[x, p]$, then by Theorem 2.27, for each $x \in (a, p)$, $-\int_x^p f\, dt$ exists and by Definition 2.9 is $\int_p^x f\, dt$. If $x = p$, then $\int_p^x f\, dt = 0$ and thus exists. Since for each $x \in (p, b)$, f is continuous on $[p, x]$, then for each $x \in (p, x)$, $\int_p^x f\, dt$ exists. Regardless of the relative position of p, for each $x \in (a, b)$, $\int_p^x f\, dt$ exists. So we can define a function g on (a, b) such that for each $x \in (a, b)$, $g(x) = \int_p^x f\, dt$.

Let $q \in (a, b)$ and pick $f(q)$ as candidate for $g'(q)$.

Let ε be a positive number.

Since f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$. Since f is uniformly continuous on $[a, b]$ and $\varepsilon > 0$, pick $\delta > 0$ such that if each of $x, y \in [a, b]$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$.

Let $x \in [a, b]$ such that $0 < |x - q| < \delta$. Now we should at this point consider a number of cases based off the relative positions of p, q, and x on the number line. The essential difficulty is the same regardless of the case, so we examine each case up to this difficulty and then draw the cases together in a one-for-all conclusion. In each case, we use Definition 2.9 and Theorem 2.25 without additional mention.

Case I $p = q$.

If $q < x$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{\int_q^x f - 0}{x-q} = \frac{\int_q^x f}{x-q}$. If $x < q$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{-\int_x^q f - 0}{-(q-x)} = \frac{\int_x^q f}{q-x} = \frac{\int_q^x f}{x-q}$.

Case II $p < q$.

If $q < x$, then $p < x$ so $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{\int_p^x f + \int_q^p f}{x-q} = \frac{\int_q^x f}{x-q}$. If $x < q$ and $x < p$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{-\int_x^p f - \int_p^q f}{x-q} = \frac{-\int_x^q f}{x-q} = \frac{\int_q^x f}{x-q}$. If $x < q$and $x = p$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{0 - \int_x^q f}{x-q} = \frac{\int_q^x f}{x-q}$. If $x < q$ and $x > p$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{\int_p^x f + \int_q^p f}{x-q} = \frac{\int_q^x f}{x-q}$.

Case III $p > q$.

If $x < q$, then $x < p$ so $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{-\int_x^p f - \int_p^q f}{x-q} = \frac{-\int_x^q f}{x-q} = \frac{\int_q^x f}{x-q}$. If $x > q$ and $x < p$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{-\int_x^p f - \int_p^q f}{x-q} = \frac{-\int_x^q f}{x-q} = \frac{\int_q^x f}{x-q}$. If $x > q$and $x = p$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{0 - \int_x^q f}{x-q} = \frac{\int_q^x f}{x-q}$. If $x > q$ and $x > p$, then $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{\int_p^x f + \int_q^p f}{x-q} = \frac{\int_q^x f}{x-q}$.

No matter the case, if we allow $\int_q^x f$ to be meaningful when $x < q$ as we have done in Definition 2.9, then the result is $\frac{\int_p^x f - \int_p^q f}{x-q} = \frac{\int_q^x f}{x-q}$. We continue with showing $\lim_{x \to q} \left(\frac{g(x) - g(q)}{x-q} \right)$ exists.

$\left| \frac{g(x)-g(q)}{x-q} - f(q) \right| = \left| \frac{\int_p^x f - \int_p^q f}{x-q} - f(q) \right|$. From our case consideration this can be written as $\left| \frac{\int_q^x f}{x-q} - f(q) \right|$. Since f is continuous, by Theorem 2.28 we choose t between x and q such that $f(t) = \frac{\int_q^x f}{x-q}$. Since $f(t) = \frac{\int_q^x f}{x-q}$, then $\left| \frac{\int_q^x f}{x-q} - f(q) \right| = |f(t) - f(q)|$. Since $x < t < q$ or $q < t < x$ and $|x - q| < \delta$, then $|t - q| < \delta$. Since $|t - q| < \delta$, then by our choice of δ, $|f(t) - f(q)| < \varepsilon$. Together $\left| \frac{g(x)-g(q)}{x-q} - f(q) \right| < \varepsilon$.

We have shown that for $\varepsilon > 0$, we chose a $\delta > 0$ so that for $x \in [a, b]$ with $0 < |x - q| < \delta$, then $\left| \frac{g(x)-g(q)}{x-q} - f(q) \right| < \varepsilon$. This is to say $\lim_{x \to q} \left(\frac{g(x)-g(q)}{x-q} \right)$ exists and is $f(q)$. Since $\lim_{x \to q} \left(\frac{g(x)-g(q)}{x-q} \right)$ exists, $g(q)$ exists as f is continuous on $[a, b]$, and since $q \in (a, b)$, we say g is differentiable at q. We write $g'(q) = f(q)$. Since q was arbitrary in (a, b), then for all $x \in (a, b)$, g is differentiable and $g'(x) = f(x)$. ∎

Theorem 2.30 Suppose $f : [a, b] \to \mathbb{R}$. If f is continuous on $[a, b]$ and for each $x \in (a, b)$, $T'(x) = f(x)$, then $\int_a^b f\,dx = T(b) - T(a)$.

Proof Let f be a continuous function on $[a, b]$. By Theorem 2.29 there exists a function $T : [a, b] \to \mathbb{R}$ so that for each $x \in (a, b)$, $T'(x) = f(x)$.

Let $p \in (a, b)$ and for each $x \in (a, b)$ define $g(x) = \int_p^x f\,dt$. By Theorem 2.29, for each $x \in (a, b)$, $g'(x) = f(x)$.

Since for each $x \in (a, b)$, $T'(x) = g'(x)$, then there is a number k such that for each $x \in [a, b]$, $T(x) = g(x) + k$. Since $b \in [a, b]$ and $g(b) = \int_p^b f\,dt$, then $T(b) = \int_p^b f\,dt + k$. Since $a \in [a, b]$ and $g(a) = \int_p^a f\,dt$, then $T(a) = \int_p^a f\,dt + k$. Now since $T(b) = \int_p^b f\,dt + k$ and $T(a) = \int_p^a f\,dt + k$, then $T(b) - T(a) = \int_p^b f\,dt - \int_p^a f\,dt$. Since by Definition 2.9, $-\int_p^a f\,dt = \int_a^p f\,dt$ and by Theorem 2.25 $\int_p^b f\,dt + \int_a^p f\,dt = \int_a^b f\,dt$, then $\int_p^b f\,dt - \int_p^a f\,dt = \int_a^b f\,dt$. The results of these last two statements together give $T(b) - T(a) = \int_a^b f\,dt$. ∎

Theorem 2.31 Suppose $f : [a, b] \to \mathbb{R}$ is a function such that $\int_a^b f\,dx$ exists. If each of and M is a number such that for each $x \in [a, b]$, $L \le f(x) \le M$, then $L(b - a) \le \int_a^b f\,dx \le M(b - a)$.

Theorem 2.32 Suppose $f : [a, b] \to \mathbb{R}$ is a function. If $\int_a^b f\,dj$ exists and for each $x \in [a, b]$, $T(x) = \int_a^x f\,dj$, then T is continuous.

Proof Suppose $\int_a^b f\,dj$ exists and for each $x \in [a, b]$, $T(x) = \int_a^x f\,dj$. Since $\int_a^b f\,dj$ exists and for each $x \in [a.b]$, $a \le x \le b$, then by Theorem 2.25 for each $x \in [a.b]$, $\int_a^x f\,dj$ exists. Thus $T : [a, b] \to \mathbb{R}$ is well defined.

Let $p \in [a, b]$ and $\varepsilon > 0$.

Since $\int_a^b f\,dj$ exists, then by Theorem 2.20 there is a number $M > 0$ such that if $q \in [a, b]$, then $|f(q)| < M$.

Pick $0 < \delta < \frac{\varepsilon}{M}$. Let $q \in [a, b]$ such that $0 < |q - p| < \delta$.

Since for each $x \in [a, b]$, $T(x) = \int_a^x f\,dj$ and since $p, q \in [a, b]$, then $|T(q) - T(p)| = \left|\int_a^q f\,dj - \int_a^p f\,dj\right|$. Since by Definition 2.22 $-\int_a^p f\,dj = \int_p^a f\,dj$, then

$\left|\int_a^q f\,dj - \int_a^p f\,dj\right| = \left|\int_a^q f\,dj + \int_p^a f\,dj\right|$. Since each of $\int_a^q f\,dj$ and $\int_p^a f\,dj$ exists, then by Theorem 2.25 $\int_a^q f\,dj + \int_p^a f\,dj = \int_p^q f\,dj$. Since $\int_a^q f\,dj + \int_p^a f\,dj = \int_p^q f\,dj$, then $\left|\int_a^q f\,dj + \int_p^a f\,dj\right| = \left|\int_p^q f\,dj\right|$. Since for each $x \in [a, b]$, $|f(x)| < M$ and $p, q \in [a, b]$, then for each x between p and q inclusive, $|f(x)| < M$. Since for each x between p and q inclusive, $-M < f(x) < M$ and $\int_p^q f\,dj$ exists, then by Theorem 2.31, $-M(q - p) < \int_p^q f\,dj < M(q - p)$. Now since $0 < |q - p| < \delta$ and $-M(q - p) < \int_p^q f\,dj < M(q - p)$, then $-M\delta < \int_p^q f\,dj < M\delta$ and $\left|\int_p^q f\,dj\right| < M\delta$. Since $0 < \delta < \frac{\varepsilon}{M}$, then $M\delta < M\,\frac{\varepsilon}{M} = \varepsilon$. Altogether $|T(q) - T(p)| < \varepsilon$.

We have shown that for $\varepsilon > 0$, we could chose a $\delta > 0$ so that for arbitrarily established $q \in [a, b]$ where $0 < |q - p| < \delta$, then $|T(q) - T(p)| < \varepsilon$. We conclude that T is continuous at p. Since p was arbitrary in $[a, b]$, we conclude T is continuous on $[a, b]$. ∎

Definition 2.10 A *partition* of the *rectangle* $R = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ is a finite subset $P = D_1 \times D_2 \times \cdots \times D_n$ where D_i is a subdivision of $[a_i, b_i]$ for each $i \in \{1, 2, \ldots, n\}$.

If there are $N_1 + 1$ elements in D_1, $N_2 + 1$ elements in D_2, and $N_n + 1$ elements in D_n, then there are $(N_1 + 1)(N_2 + 2) \cdots (N_n + 1)$ elements in P and the partition creates $N_1 \cdot N_2 \cdot \cdots \cdot N_n = m$ sub-rectangles in R. We choose to define the integral of a real valued function over a rectangle in Euclidean n-space building on the previous work.

Definition 2.11 A *refinement* $H = H_1 \times H_2 \times \cdots \times H_n$ of a partition $P = D_1 \times D_2 \times \cdots \times D_n$ of the rectangle $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ satisfies ${}_i$ is a refinement of subdivision ${}_i$ of $[a_i, b_i]$ for each $i \in \{1, 2, \ldots, n\}$or simply H is a partition of R where $P \subseteq H$.

For a partition of a rectangle and a refinement of a partition we have theorems analogous to Theorems 2.17, 2.18, and 2.19, e.g., suppose ${}_1 = \{x^{i_1}\}_{i_1=1}^{k_1}$, $D_2 = \{x^{i_2}\}_{i_2=1}^{k_2}$, $\ldots$, and $D_n = \{x^{i_n}\}_{i_n=1}^{k_n}$, and $P = D_1 \times \cdots \times D_n$, then P is a refinement of P, et cetera.

Definition 2.12 For a partition P of a rectangle $R = [a_1, b_1]\times[a_2, b_2]\times\cdots\times[a_n, b_n]$ and a refinement $= H_1 \times\cdots\times H_n$ of P with $H_j = \{x^j_{i_j}\}_{i_j=0}^{k_j}$ for each $j \in \{1, 2, \ldots, n\}$, then for interpolation sequences $\{t^1_{i_1}\}_{i_1=1}^{k_1}$ of H_1, $\{t^2_{i_2}\}_{i_2=1}^{k_2}$ of H_2, …, and $\{t^n_{i_n}\}_{i_n=1}^{k_n}$ of H_n we define $\{t^{i_1}\}_{i_1=1}^{k_1}\times\{t^{i_2}\}_{i_2=1}^{k_2}\times\cdots\times\{t^{i_n}\}_{i_n=1}^{k_n} = \left\{\cdots\left\{\{(t^1_{i_1}, t^2_{i_2}, \ldots, t^n_{i_n})\}_{i_1=1}^{k_1}\right\}_{i_2=1}^{k_2}\cdots\right\}_{i_n=1}^{k_n}$ to be an *interpolating sequence* of H.

Definition 2.13 For a rectangle $R = [a_1, b_1]\times[a_2, b_2]\times\cdots\times[a_n, b_n]$ and a function $f : R \longrightarrow \mathbb{R}$, if there exists a number A such that for each $\varepsilon > 0$ there is a partition P of R such that if $H = \{x^1_{i_1}\}_{i_1=0}^{k_1}\times\cdots\times\{x^n_{i_n}\}_{i_n=0}^{k_n}$ is a refinement of P and $\{t_{i_1 i_2 \ldots i_n}\}$ is an interpolating sequence of ,
$\left|\sum_{i_1=1}^{k_1}\left(\sum_{i_2=1}^{k_2}\left(\cdots\left(\sum_{i_n=1}^{k_n} f(t_{i_1 i_2 \quad i_n})\,\Delta(x^{i_1})\,\Delta(x^{i_2})\cdots\Delta(x^{i_n})\right)\cdots\right)\right) - A\right| < \varepsilon$, then we say f is integrable on R and we denote the number A by $\int_R f\, dR$.

We could also produce alternate definitions for integrable on a rectangle, then state and prove many of the previous theorems regarding the integral of a real valued function on a closed interval in the setting of a real valued function on a closed rectangle. We will not since the proofs are similar albeit more cumbersome. We will assume them and refer to the simpler and one new result for rectangles. What we need is a way to calculate $\int_R f\, dR$ in the manner that Theorem 2.30 allowed us to calculate $\int_a^b f(x)\, dx$.

Fortunately we can use the same theorem to carry out the calculation of $\int_R f\, dR$ without any new theory. For example, if $R = [a_1, b_1]\times[a_2, b_2]$ and $f : R \longrightarrow \mathbb{R}$ is a continuous function, then we can define $A_1 : [a_1, b_1] \longrightarrow \mathbb{R}$ by $A_1(x^1) = \int_{a_2}^{b_2} f(x^1, x^2)\, dx^2$ which is well-defined by Theorem 2.27 since if $x^1 \in [a_1, b_1]$, then $f(x^1, x^2)$ is a continuous real valued function on $[a_2, b_2]$. A_1 is continuous since the integral operator is continuous by Theorem 2.32, therefore, by Theorem 2.27 $\int_{a_1}^{b_1} A_1(x^1)\, dx^1$ exists and we write $\int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x^1, x^2)\, dx^2\right) dx^1$. The question that remains is, "Is $\int_R f\, dR$ the same as $\int_{a_1}^{b_1}\left(\int_{a_2}^{b_2} f(x^1, x^2)\, dx^2\right) dx^1$?", and the answer to this question and a corollary of it, will be our last theorem in this section on integration of real valued functions. The theorem is a special case of a more general theorem known as Fubini's Theorem.

Theorem 2.33 Suppose $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ and $f : R \longrightarrow \mathbb{R}$ is continuous, then $\int\limits_R f\, dR$ is equal to $\int_{a_n}^{b_n} \big(\cdots \int_{a_2}^{b_2} \big(\int_{a_1}^{b_1} f(x^1, x^2, \ldots, x^n)\, dx^1 \big)\, dx^2 \big) \cdots dx^n$.

Proof. We prove this by induction on n. For $n = 1$, there is nothing to show. We make our base case for $n = 2$.

Define $A_1 : [a_1, b_1] \longrightarrow \mathbb{R}$ by $A_1(x^1) = \int_{a_2}^{b_2} f(x^1, x^2)\, dx^2$. Note $A_1(x^1)$ is well-defined for each $x^1 \in [a_1, b_1]$ since if $x^1 \in [a_1, b_1]$, then $f(x^1, \cdot)$ is a continuous real valued function on $[a_2, b_2]$ by Theorem 2.27. A_1 is continuous since the integral operator is continuous by Theorem 2.32, therefore, by Theorem 2.27, $\int_{a_1}^{b_1} A_1(x^1)\, dx^1$ exists and is $\int_{a_1}^{b_1} \big(\int_{a_2}^{b_2} f(x^1, x^2)\, dx^2 \big)\, dx^1$ call it Z_2. Also by the assumption $\int\limits_R f\, dR$ exists call it Z_1.

Let $\varepsilon > 0$

Since $\int_R f\, dR$ exists and $\frac{\varepsilon}{3} > 0$, then we pick a partition $P = F_1 \times F_2$ of R so that if each of $H_1 = \{x^1_{i_1}\}_{i_1=0}^{k_1}$ and $H_2 = \{x^2_{i_2}\}_{i_1=0}^{k_2}$ is a refinement of F_1 and F_2, respectively, and $\Big\{\{(t^1_{i_1}, t^2_{i_2})\}_{i_1=1}^{k_1}\Big\}_{i_2=1}^{k_2}$ is an interpolating sequence of refinement $H_1 \times H_2$, then $\left| \sum\limits_{i_1=1}^{k_1} \sum\limits_{i_2=1}^{k_2} f(t^1_{i_1}, t^2_{i_2})\, \Delta x^1_{i_1}\, \Delta x^2_{i_2} - Z_1 \right| < \frac{\varepsilon}{3}$.

Since $\int\limits_{a_1}^{b_1} A_1(x^1)\, dx^1$ exists and $\frac{\varepsilon}{3} > 0$, pick subdivision D_1 of $[a_1, b_1]$ such that if $H = \{x^1_{i_1}\}_{i_1=0}^{k_1}$ is a refinement of D_1 and $\{t^1_{i_1}\}_{i_1=1}^{k_1}$ is an interpolating sequence of H, then $\left| \sum\limits_{i_1=1}^{k_1} A_1(t^1_{i_1})\, dx^{i_1} - Z_2 \right| < \frac{\varepsilon}{3}$. Let $H = \{x^1_{i_1}\}_{i_1=0}^{k_1}$ be a refinement of $D_1 \cup F_1$ and let $\{t^1_{i_1}\}_{i_1=1}^{k_1}$ be an interpolation sequence of H. Therefore,

$$\left| \sum_{i_1=1}^{k_1} A_1(t^1_{i_1})\, dx^{i_1} - Z_2 \right| < \frac{\varepsilon}{3}. \tag{1}$$

Since for each $i_1 \in \{1, 2, \ldots, k_1\}$, $\int\limits_{a_2}^{b_2} f(t^1_{i_1}, x^2_{i_2})\, dx^{i_2}$ exists and $\frac{\varepsilon}{3(b_1 - a_1)} > 0$, then pick subdivision K_{i_1} of $[a_1, b_1]$ such that if $H_{i_1} = \{x^2_{i_2}\}_{i_2=0}^{k_2}$ is a refinement of K_{i_1} and

$\{t_{i_2}^2\}_{i_2=1}^{k_2}$ is an interpolation sequence of H_{i_1}, then $\left|\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2 - A_1(t_{i_1}^1)\right| < \frac{\varepsilon}{3(b_1-a_1)}$.

Pick $D_2 = \bigcup_{i_1=1}^{k_2} K_{i_1}$. D_2 is a refinement of K_{i_1} for each $i_1 \in \{1, 2, \ldots, k_1\}$ by Theorem 2.19. Let $H_2 = \{x_{i_2}^2\}_{i_2=1}^{k_2}$ be a refinement of $D_2 \cup F_2$ and $\{t_{i_2}^2\}_{i_2=1}^{k_2}$ be an interpolation sequence of H_2.

First, since H_1 is a refinement of F_1 and H_2 is a refinement of F_2, then $H_1 \times H_2$ is a refinement of P, and $\left\{\{(t_{i_1}^1, t_{i_2}^2)\}_{i_1=1}^{k_1}\right\}_{i_2=1}^{k_2}$ is an interpolation sequence of $H_1 \times H_2$ so

$$\left|\sum_{i_1=1}^{k_1}\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_1}^1\,\Delta x_{i_2}^2 - Z_1\right| < \frac{\varepsilon}{3}. \tag{2}$$

Second, since H_2 is a refinement of K_{i_1} for each $i_1 \in \{1, 2, \ldots, k_1\}$ and $\{t_{i_2}^2\}_{i_2=1}^{k_2}$, then for each $i_1 \in \{1, 2, \ldots, k_1\}$, $\left|\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2 - A_1(t_{i_1}^1)\right| < \frac{\varepsilon}{3(b_1-a_1)}$ and in particular $A_1(t_{i_1}^1) - \frac{\varepsilon}{3(b_1-a_1)} < \sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2 < A_1(t_{i_1}^1) + \frac{\varepsilon}{3(b_1-a_1)}$ Since for each $i_1 \in \{1, 2, \ldots, k_1\}$, $\Delta x_{i_1}^1 > 0$ it follows

$$\sum_{i_1=1}^{k_1}\left(\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2\right)\Delta x_{i_1}^1 < \sum_{i_1=1}^{k_1}\left(A_1(t_{i_1}^1) + \frac{\varepsilon}{3(b_1-a_1)}\right)\Delta x_{i_1}^1 = \sum_{i_1=1}^{k_1} A_1(t_{i_1}^1)\,\Delta x_{i_1}^1 + \frac{\varepsilon}{3}$$

and

$$\sum_{i_1=1}^{k_1}\left(\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2\right)\Delta x_{i_1}^1 < \sum_{i_1=1}^{k_1}\left(A_1(t_{i_1}^1) - \frac{\varepsilon}{3(b_1-a_1)}\right)\Delta x_{i_1}^1 = \sum_{i_1=1}^{k_1} A_1(t_{i_1}^1)\,\Delta x_{i_1}^1 - \frac{\varepsilon}{3}.$$

Therefore,

$$\left|\sum_{i_1=1}^{k_1}\left(\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2\right)\Delta x_{i_1}^1 - \sum_{i_1=1}^{k_1} A_1(t_{i_1}^1)\,\Delta x_{i_1}^1\right| < \frac{\varepsilon}{3}. \tag{3}$$

From the triangle inequality we have the first inequality

$$\begin{aligned}|Z_1 - Z_2| \le &\left|Z_1 - \sum_{i_1=1}^{k_1}\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_1}^1\,\Delta x_{i_2}^2\right| \\ &+\left|\sum_{i_1=1}^{k_1}\left(\sum_{i_2=1}^{k_2} f(t_{i_1}^1, t_{i_2}^2)\,\Delta x_{i_2}^2\right)\Delta x_{i_1}^1 - \sum_{i_1=1}^{k_1} A_1(t_{i_1}^1)\,\Delta x_{i_1}^1\right| \\ &+\left|\sum_{i_1=1}^{k_1} A_1(t_{i_1}^1)\,dx^{i_1} - Z_2\right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}\end{aligned}$$

the final inequality from (1), (2), and (3); thus, $|Z_1 - Z_2| < \varepsilon$. We conclude that $Z_1 = Z_2$. We have established the result for the base case $n = 2$.

We have now the inductive hypothesis:

$$\int_{a_n}^{b_n}\left(\cdots\int_{a_2}^{b_2}\left(\int_{a_n}^{b_n} f(x^1, x^2, \ldots, x^n)\,dx^1\right)dx^2\right)\cdots dx^n$$

$$= \int_{a_n}^{b_n} \cdots \int_{a_2}^{b_2} \int_{a_n}^{b_n} f(x^1, x^2, \ldots, x^n)\, dx^1\, dx^2 \cdots dx^n.$$

By the supposition, $\int_{a_{n+1}}^{b_{n+1}} \cdots \int_{a_2}^{b_2} \int_{a_n}^{b_n} f(x^1, x^2, \ldots, x^n)\, dx^1\, dx^2 \cdots dx^{n+1}$ exists call it Z_1. Define $A_{n+1}(x^{n+1}) = \int_{a_n}^{b_n} \cdots \int_{a_2}^{b_2} \int_{a_n}^{b_n} f(x^1, x^2, \ldots, x^n, x^{n+1})\, dx^1\, dx^2 \cdots dx^n$, for each $x^{n+1} \in [a_{n+1}, b_{n+1}]$, which is well defined by our inductive hypothesis and even continuous by consideration of Theorem 2.32. Since A_{n+1} is continuous on $[a_{n+1}, b_{n+1}]$, then $\int_{a_{n+1}}^{b_{n+1}} A_{n+1}(x^{n+1})\, dx^{n+1}$ exists, call it Z_2.

Let $\varepsilon > 0$.

Since Z_1 exists and $\frac{\varepsilon}{3} > 0$, we pick a partition $P_1 = F_1 \times F_2 \times \cdots \times F_{n+1}$ of $R_{n+1} = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_{n+1}, b_{n+1}]$ such that if $H = \{x_{\iota_1}^1\}_{\iota_1=0}^{k_1} \times \cdots \times \{x_{\iota_{n+1}}^{n+1}\}_{\iota_{n+1}=1}^{k_{n+1}}$ is a refinement of P_1 and $\left\{\cdots \{(t_{\iota_1}^1, \ldots, t_{\iota_{n+1}}^{n+1})\}_{\iota_1=1}^{k_1} \cdots\right\}_{\iota_{n+1}=1}^{k_{n+1}}$ is an interpolating sequence of H, then $\left|\sum_{\iota_{n+1}=1}^{k_{n+1}} \cdots \sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_{n+1}}^{n+1})\, \Delta x_{\iota_1}^1 \cdots \Delta x_{\iota_{n+1}}^{n+1} - Z_1\right| < \frac{\varepsilon}{3}$.

Since Z_2 exists and $\frac{\varepsilon}{3} > 0$, pick a subdivision $_{n+1}$ of $[a_{n+1}, b_{n+1}]$ such that if $H_{n+1} = \{x_{\iota_{n+1}}^{n+1}\}_{\iota_{n+1}=1}^{k_{n+1}}$ is a refinement of $_{n+1}$ and $\{t_{\iota_{n+1}}^{n+1}\}_{\iota_{n+1}=1}^{k_{n+1}}$ is an interpolation sequence of H_{n+1}, then $\left|\sum_{\iota_{n+1}=1}^{k_{n+1}} A_{n+1}(x^{n+1})\, \Delta x_{\iota_{n+1}}^{n+1} - Z_2\right| < \frac{\varepsilon}{3}$. Let $H_{n+1} = \{x_{\iota_{n+1}}^{n+1}\}_{\iota_{n+1}=1}^{k_{n+1}}$ be a refinement of $D_{n+1} \bigcup F_{n+1}$ and let $\{t_{\iota_{n+1}}^{n+1}\}_{\iota_{n+1}=1}^{k_{n+1}}$ be an interpolation sequence of H_{n+1}. We thus have

$$\left|\sum_{\iota_{n+1}=1}^{k_{n+1}} A_{n+1}(x^{n+1})\, \Delta x_{\iota_{n+1}}^{n+1} - Z_2\right| < \frac{\varepsilon}{3}. \tag{4}$$

For each $\iota_{n+1} \in \{1, 2, \ldots, k_{n+1}\}$,

$\int_{a_n}^{b_n} \cdots \int_{a_2}^{b_2} \int_{a_n}^{b_n} f(x^1, x^2, \ldots, x^n, t_{\iota_{n+1}}^{n+1})\, dx^1\, dx^2 \cdots dx^n$ exists and since $\frac{\varepsilon}{3(b_{n+1} - a_{n+1})} > 0$, then there exists a partition $P_{\iota_{n+1}}$ of $R_n = [a_1, b_1] \times \cdots \times [a_n, b_n]$, such that if $H = \{x_{\iota_1}^1\}_{\iota_1=0}^{k_1} \times \cdots \times \{x_{\iota_n}^n\}_{\iota_n=0}^{k_n}$ is a refinement of $P_{\iota_{n+1}}$ and $\left\{\cdots \{(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\}_{\iota_1=1}^{k_1} \cdots\right\}_{\iota_n=1}^{k_n}$ is an interpolating sequence of H, then

$$\left|\sum_{\iota_n=1}^{k_n} \cdots \sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n, t_{\iota_{n+1}}^{n+1})\, \Delta x_{\iota_1}^1 \cdots \Delta x_{\iota_n}^n - A_{n+1}(t_{\iota_{n+1}}^{n+1})\right| < \frac{\varepsilon}{3(b_{n+1} - a_{n+1})}.$$

Let $P_n = \left(\bigcup_{\iota_{n+1}=1}^{k_{n+1}} P_{\iota_{n+1}}\right)$. Thus P_n is a refinement of $P_{\iota_{n+1}}$ for each $i_{n+1} \in \{1, 2, \ldots, k_{n+1}\}$. Let $H_n = \{x_{\iota_1}^1\}_{\iota_1=0}^{k_1} \times \cdots \times \{x_{\iota_{n+1}}^{n+1}\}_{\iota_{n+1}=0}^{k_{n+1}}$ be a refinement of $P_n \bigcup (F_1 \times \cdots \times F_n)$ and let $\left\{\cdots \{(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\}_{\iota_1=1}^{k_1} \cdots\right\}_{\iota_n=1}^{k_n}$ be an interpolating sequence of H_n.

First we have $= H_n \times H_{n+1}$ is a refinement of $F_1 \times \cdots \times F_n \times F_{n+1}$ and $\left\{ \cdots \left\{ (t^1_{i_1}, \ldots, t^{n+1}_{i_{n+1}}) \right\}_{i_1=1}^{k_1} \cdots \right\}_{i_{n+1}=1}^{k_{n+1}}$ is an interpolation sequence of H; therefore,

$$\left| \sum_{i_{n+1}=1}^{k_{n+1}} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^{n+1}_{i_{n+1}} - Z_1 \right| < \tfrac{\varepsilon}{3}. \tag{5}$$

Second, since H_n is a refinement of $P_{i_{n+1}}$ for each $i_{n+1} \in \{1, 2, \ldots, k_{n+1}\}$ and $\left\{ \cdots \left\{ (t^1_{i_1}, \ldots, t^n_{i_n}) \right\}_{i_1=1}^{k_1} \cdots \right\}_{i_n=1}^{k_n}$ is an interpolating sequence of H_n, then

$$\left| \sum_{i_n=1}^{k_n} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^n_{i_n} - A_{n+1}(t^{n+1}_{i_{n+1}}) \right| < \frac{\varepsilon}{3\,(b_{n+1} - a_{n+1})},$$

in particular

$$A_{n+1}(t^{n+1}_{i_{n+1}}) + \frac{\varepsilon}{3\,(b_{n+1} - a_{n+1})} < \sum_{i_n=1}^{k_n} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x_{i_n} < A_{n+1}(t^{n+1}_{i_{n+1}}) + \frac{\varepsilon}{3\,(b_{n+1} - a_{n+1})}.$$

$$\begin{aligned} \sum_{i_{n+1}=1}^{k_{n+1}} \left(\sum_{i_n=1}^{k_n} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^n_{i_n} \right) \Delta x^{n+1}_{i_{n+1}} &< \sum_{i_{n+1}=1}^{k_{n+1}} \left(A_{n+1}(t^{n+1}_{i_{n+1}}) + \frac{\varepsilon}{3\,(b_{n+1} - a_{n+1})} \right) \Delta x^{n+1}_{i_{n+1}} \\ &= \sum_{i_{n+1}=1}^{k_{n+1}} A_{n+1}(t^{n+1}_{i_{n+1}}) \, \Delta x^{n+1}_{i_{n+1}} + \frac{\varepsilon}{3} \end{aligned}$$

and

$$\begin{aligned} \sum_{i_{n+1}=1}^{k_{n+1}} \left(\sum_{i_n=1}^{k_n} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^n_{i_n} \right) \Delta x^{n+1}_{i_{n+1}} &> \sum_{i_{n+1}=1}^{k_{n+1}} \left(A_{n+1}(t^{n+1}_{i_{n+1}}) - \frac{\varepsilon}{3\,(b_{n+1} - a_{n+1})} \right) \Delta x^{n+1}_{i_{n+1}} \\ &= \sum_{i_{n+1}=1}^{k_{n+1}} A_{n+1}(t^{n+1}_{i_{n+1}}) \, \Delta x^{n+1}_{i_{n+1}} - \frac{\varepsilon}{3} \end{aligned}$$

thus

$$\left| \sum_{i_{n+1}=1}^{k_{n+1}} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^{n+1}_{i_{n+1}} - \sum_{i_{n+1}=1}^{k_{n+1}} A_{n+1}(t^{n+1}_{i_{n+1}}) \, \Delta x^{n+1}_{i_{n+1}} \right| < \tfrac{\varepsilon}{3}. \tag{6}$$

We have

$$\begin{aligned} |Z_1 - Z_2| \le{} & \left| Z_1 - \sum_{i_{n+1}=1}^{k_{n+1}} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^{n+1}_{i_{n+1}} \right| + \\ & \left| \sum_{i_{n+1}=1}^{k_{n+1}} \cdots \sum_{i_1=1}^{k_1} f(t^1_{i_1}, \ldots, t^n_{i_n}, t^{n+1}_{i_{n+1}}) \, \Delta x^1_{i_1} \cdots \Delta x^{n+1}_{i_{n+1}} - \sum_{i_{n+1}=1}^{k_{n+1}} A_{n+1}(t^{n+1}_{i_{n+1}}) \, \Delta x^{n+1}_{i_{n+1}} \right| + \\ & \left| \sum_{i_{n+1}=1}^{k_{n+1}} A_{n+1}(t^{n+1}_{i_{n+1}}) \, \Delta x^{n+1}_{i_{n+1}} - Z_2 \right| \end{aligned}$$

which, from (4), (5), and (6), is less than $\frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}$. That is $|Z_1 - Z_2| < \varepsilon$ for arbitrary $\varepsilon > 0$ so $Z_1 = Z_2$. By the principle of proof by induction we have our result. ∎

Theorem 2.34 Suppose $R = [a_1, b_1]\times[a_2, b_2]\times\cdots\times[a_n, b_n]$ and $f: R \longrightarrow \mathbb{R}$ is continuous, then $\int_{a_n}^{b_n}\cdots\int_{a_t}^{b_t}\cdots\int_{a_s}^{b_s}\cdots\int_{a_1}^{b_1} f(x^1, x^2, \ldots, x^n)\,dx^1\cdots dx^s\cdots dx^t\cdots dx^n$ is equal to $\int_{a_n}^{b_n}\cdots\int_{a_s}^{b_s}\cdots\int_{a_t}^{b_t}\cdots\int_{a_1}^{b_1} f(x^1, x^2, \ldots, x^n)\,dx^1\cdots dx^t\cdots dx^s\cdots dx^n$.

Proof. We let Z_1 be the first integral and Z_2 be the second. Let $\varepsilon > 0$. Since Z_1 exists and $\frac{\varepsilon}{2} > 0$, then there exists a partition P_1 of R such that if $H = \{x_{\iota_1}^1\}_{\iota_1=0}^{k_1}\times\cdots\times\{x_{\iota_n}^n\}_{\iota_n=1}^{k_n}$ is a refinement of P_1 and $\left\{\cdots\{(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\}_{\iota_1=1}^{k_1}\cdots\right\}_{\iota_n=1}^{k_n}$ is an interpolation sequence of , then

$$\left|\sum_{\iota_n=1}^{k_n}\cdots\sum_{\iota_t=1}^{k_t}\cdots\sum_{\iota_s=1}^{k_s}\cdots\sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_n}^n - Z_1\right| < \frac{\varepsilon}{2}.$$

Since Z_2 exists and $\frac{\varepsilon}{2} > 0$, then there exists a partition P_2 of R such that if $H = \{x_{\iota_1}^1\}_{\iota_1=0}^{k_1}\times\cdots\times\{x_{\iota_n}^n\}_{\iota_n=1}^{k_n}$ is a refinement of P_2 and $\left\{\cdots\{(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\}_{\iota_1=1}^{k_1}\cdots\right\}_{\iota_n=1}^{k_n}$ is an interpolating sequence of , then

$$\left|\sum_{\iota_n=1}^{k_n}\cdots\sum_{\iota_s=1}^{k_s}\cdots\sum_{\iota_t=1}^{k_t}\cdots\sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_n}^n - Z_2\right| < \frac{\varepsilon}{2}.$$

Let $H = \{x_{\iota_1}^1\}_{\iota_1=0}^{k_1}\times\cdots\times\{x_{\iota_n}^n\}_{\iota_n=1}^{k_n}$ be a refinement of $P_1 \cup P_2$ and let $\left\{\cdots\{(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\}_{\iota_1=1}^{k_1}\cdots\right\}_{\iota_n=1}^{k_n}$ be an interpolating sequence of .

We observe

$$\left|\sum_{\iota_n=1}^{k_n}\cdots\sum_{\iota_t=1}^{k_t}\cdots\sum_{\iota_s=1}^{k_s}\cdots\sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_n}^n - \sum_{\iota_n=1}^{k_n}\cdots\sum_{\iota_s=1}^{k_s}\cdots\sum_{\iota_t=1}^{k_t}\cdots\sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_n}^n - Z_1\right| = 0$$

since for each term $f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_n}^n$ in the first summation, there is the term $-f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_n}^n$ in the second summation, but by commutativity of multiplication

$$-f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_n}^n = -f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_n}^n.$$

The following observation will yield the desired result:

$$|Z_1 - Z_2| < \left|Z_1 - \sum_{\iota_n=1}^{k_n}\cdots\sum_{\iota_t=1}^{k_t}\cdots\sum_{\iota_s=1}^{k_s}\cdots\sum_{\iota_1=1}^{k_1} f(t_{\iota_1}^1, \ldots, t_{\iota_n}^n)\,\Delta x_{\iota_1}^1\cdots\Delta x_{\iota_s}^s\cdots\Delta x_{\iota_t}^t\cdots\Delta x_{\iota_n}^n\right| +$$

$$\left| \sum_{i_n=1}^{k_n} \cdots \sum_{i_t=1}^{k_t} \cdots \sum_{i_s=1}^{k_s} \cdots \sum_{i_1=1}^{k_1} f(t_{i_1}^1, \ldots, t_{i_n}^n) \triangle x_{i_1}^1 \cdots \triangle x_{i_s}^s \cdots \triangle x_{i_t}^t \cdots \triangle x_{i_n}^n - \right.$$
$$\left. \sum_{i_n=1}^{k_n} \cdots \sum_{i_s=1}^{k_s} \cdots \sum_{i_t=1}^{k_t} \cdots \sum_{i_1=1}^{k_1} f(t_{i_1}^1, \ldots, t_{i_n}^n) \triangle x_{i_1}^1 \cdots \triangle x_{i_t}^t \cdots \triangle x_{i_s}^s \cdots \triangle x_{i_n}^n - Z_1 \right| +$$
$$\left| \sum_{i_n=1}^{k_n} \cdots \sum_{i_s=1}^{k_s} \cdots \sum_{i_t=1}^{k_t} \cdots \sum_{i_1=1}^{k_1} f(t_{i_1}^1, \ldots, t_{i_n}^n) \triangle x_{i_1}^1 \cdots \triangle x_{i_t}^t \cdots \triangle x_{i_s}^s \cdots \triangle x_{i_n}^n - Z_2 \right| < \frac{\varepsilon}{2} + 0 + \frac{\varepsilon}{2}.$$

Therefore $|Z_1 - Z_2| < \varepsilon$ for abritrary $\varepsilon > 0$ and we conclude $Z_1 = Z_2$. ∎

CHAPTER III

DIFFERENTIAL FORMS

We start by defining tensors. After some work, we specialize to alternating tensors. Next, we develop the wedge product to work with alternating tensors. Finally, we introduce fields and put these together with alternating tensors to build the construct called differential forms. At the last of this chapter we define the differential operator to build new forms from old ones.

k-tensors

Linear functions carried individual vectors from one vector space to another in a particular manner that made the transformation linear. Multilinear functions will carry a finite number of vectors in one vectors space to a finite number of vectors in another, and they will carry these vectors in a manner that resembles a linear transformation for each one.

We will denote the k-fold product $\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ by $(\mathbb{R}^n)^k$. To elaborate, if $v \in (\mathbb{R}^n)^k$, then v is comprised of k ordered vectors chosen from $\mathbb{R}^n$. For example, $\left(\begin{pmatrix} 6 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 8 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}\right) \in (\mathbb{R}^2)^3$. A k-tensor can be defined as a multilinear function from $(\mathbb{R}^n)^k$ to $\mathbb{R}$.

Definition 3.1 A function $T : (\mathbb{R}^n)^k \longrightarrow \mathbb{R}$ is a k-tensor if for each position $i \in \{1, 2, \ldots, k\}$, $T(\boldsymbol{x}_1, \ldots, \boldsymbol{a}_i + \boldsymbol{b}_i, \ldots, \boldsymbol{x}_k) = T(\boldsymbol{x}_1, \ldots, \boldsymbol{a}_i, \ldots, \boldsymbol{x}_k) + T(\boldsymbol{x}_1, \ldots, \boldsymbol{b}_i, \ldots, \boldsymbol{x}_k)$ and $T(\boldsymbol{x}_1, \ldots, \alpha \cdot \boldsymbol{x}_i, \ldots, \boldsymbol{x}_k) = \alpha T(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_k)$ where $\alpha \in \mathbb{R}$.

A well-known example of a k-tensor is the determinant, which acts on n columns and n rows (easily thought of as n vectors each with n components) and returns a real number. We will give another example. Define $T: (\mathbb{R}^2)^2 \longrightarrow \mathbb{R}$ by $T(\boldsymbol{x}, \boldsymbol{y}) = (x^1 + x^2)(y^1 + y^2)$. We claim T is a 2-tensor. Following the definition of T, we calculate $T\left(\begin{pmatrix} a^1 + b^1 \\ a^2 + b^2 \end{pmatrix}, \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}\right) = ((a^1 + b^1) + (a^2 + b^2))(y^1 + y^2)$
$= (a^1 + b^1)(y^1 + y^2) + (a^2 + b^2)(y^1 + y^2) = a^1(y^1 + y^2) + a^2(y^1 + y^2)$
$+ b^1(y^1 + y^2) + b^2(y^1 + y^2) = T\left(\begin{pmatrix} a^1 \\ a^2 \end{pmatrix}, \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}\right) + T\left(\begin{pmatrix} b^1 \\ b^2 \end{pmatrix}, \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}\right)$ and similarly
$T\left(\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \begin{pmatrix} a^1 + b^1 \\ a^2 + b^2 \end{pmatrix}\right) = T\left(\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \begin{pmatrix} a^1 \\ a^2 \end{pmatrix}\right) + T\left(\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \begin{pmatrix} b^1 \\ b^2 \end{pmatrix}\right)$, which satisfies the first condition of being multilinear. Also, $T\left(\alpha \cdot \begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}\right) = (\alpha x^1 + \alpha x^2)(y^1 + y^2)$
$= \alpha(x^1 + x^2)(y^1 + y^2) = \alpha T\left(\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}, \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}\right)$ and similarly for α in the second position. T satisfies the conditions of Definition 3.1, so T is a 2-tensor. One can see from this simple example how unruly notation with tensors might become when answering n-dimensional questions for large n.

Definition 3.2 Let $\mathcal{T}^k(\mathbb{R}^n)$ be the set of all k-tensors under "+" and "·" defined for $S, T \in \mathcal{T}^k(\mathbb{R}^n)$, $_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$by
$(S + T)(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) = S(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) + T(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)$ and
$(\alpha \cdot S)(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) = \alpha S(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)$.

From the definition, adding two k-tensors together amounts to adding the real number images together and multiplying a k-tensor by a real number α turns out simply to be the usual real number multiplication of real number image of the k-tensor with α. We define a new operation between k-tensors and l-tensors.

Definition 3.3 For $S \in \mathcal{T}^k(\mathbb{R}^n)$, $T \in \mathcal{T}^l(\mathbb{R}^n)$, and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{k+l} \in \mathbb{R}^n$, define the *tensor product* "$\otimes$" by $S \otimes T(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{k+l}) = S(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)\, T(\boldsymbol{x}_{k+1}, \boldsymbol{x}_{k+2}, \ldots, \boldsymbol{x}_{k+l})$.

We build some familiarity with k-tensors, their grouping in vector spaces, and the tensor product between those vector spaces. As is the usual convention applied to our situation, multiplication takes precedence over addition and so in consideration of order in expressions with multiple operations we apply "⊗" before we apply "+".

Theorem 3.1 Suppose $S, S_1, S_2 \in \mathcal{T}^k(\mathbb{R}^n)$, $T, T_1, T_2 \in \mathcal{T}^l(\mathbb{R}^n)$, and $U \in \mathcal{T}^m(\mathbb{R}^n)$. The following are properties:

1. $(S_1 + S_2) \otimes T = S_1 \otimes T + S_2 \otimes T$
2. $S \otimes (T_1 + T_2) = S \otimes T_1 + S \otimes T_2$
3. $(S \otimes T) \otimes U = S \otimes (T \otimes U)$
4. $(\alpha \cdot S) \otimes T = S \otimes (\alpha T) = \alpha(S \otimes T)$

Proof. Let $_1, x_2, \ldots, x_{k+l+m} \in \mathbb{R}^n$. The following manipulations are easy to follow using Definitions 3.1, 3.2, 3.3, and properties of the real number system.

$$
\begin{aligned}
1. \quad (S_1 + S_2) \otimes T(x_1, \ldots, x_{k+l}) &= (S_1 + S_2)(x_1, \ldots, x_k)\, T(x_{k+1}, \ldots, x_{k+l}) \\
&= [S_1(x_1, \ldots, x_k) + S_2(x_1, \ldots, x_k)]\, T(x_{k+1}, \ldots, x_{k+l}) \\
&= S_1(x_1, \ldots, x_k)\, T(x_{k+1}, \ldots, x_{k+l}) + \\
&\qquad S_2(x_1, \ldots, x_k)\, T(x_{k+1}, \ldots, x_{k+l}) \\
&= S_1 \otimes T(x_1, \ldots, x_{k+l}) + S_2 \otimes T(x_1, \ldots, x_{k+l}) \\
&= (S_1 \otimes T + S_2 \otimes T)(x_1, \ldots, x_{k+l})
\end{aligned}
$$

$$
\begin{aligned}
2. \quad S \otimes (T_1 + T_2)(x_1, \ldots, x_{k+l}) &= S(x_1, \ldots, x_k)(T_1 + T_2)(x_{k+1}, \ldots, x_{k+l}) \\
&= S(x_1, \ldots, x_k)[T_1(x_{k+1}, \ldots, x_{k+l}) + T_2(x_{k+1}, \ldots, x_{k+l}) \\
&= S(x_1, \ldots, x_k)\, T_1(x_{k+1}, \ldots, x_{k+l}) + \\
&\qquad S(x_1, \ldots, x_k)\, T_2(x_{k+1}, \ldots, x_{k+l}) \\
&= S \otimes T_1(x_1, \ldots, x_{k+l}) + S \otimes T_2(x_1, \ldots, x_{k+l}) \\
&= (S \otimes T_1 + S \otimes T_2)(x_1, \ldots, x_{k+l})
\end{aligned}
$$

$$
\begin{aligned}
3.\ (S\otimes T)\otimes U(x_1,\dots,x_{k+l+m}) &= (S\otimes T)(x_1,\dots,x_{k+l})\,U(x_{k+l+1},\dots,x_{k+l+m}) \\
&= S(x_1,\dots,x_k)\,T(x_{k+1},\dots,x_{k+l})\,U(x_{k+l+1},\dots,x_{k+l+m} \\
&= S(x_1,\dots,x_k)\,(T\otimes U)(x_{k+1},\dots,x_{k+l+m}) \\
&= S\otimes(T\otimes U)(x_1,\dots,x_{k+l+m})
\end{aligned}
$$

$$
\begin{aligned}
4.\ ((\alpha\cdot S)\otimes T)(x_1,\dots,x_{k+l}) &= (\alpha\cdot S)(x_1,\dots,x_k)\,T(x_{k+1},\dots,x_{k+l}) \\
&= \alpha S(x_1,\dots,x_k)\,T(x_{k+1},\dots,x_{k+l}) \\
&= \alpha(S\otimes T)(x_1,\dots,x_{k+l}) \\
&= S(x_1,\dots,x_k)\,\alpha T(x_{k+1},\dots,x_{k+l}) \\
&= S(x_1,\dots,x_k)\,(\alpha\cdot T)(x_{k+1},\dots,x_{k+l}) \\
&= (S\otimes(\alpha\cdot T))(x_1,\dots,x_{k+l})
\end{aligned}
$$

We interject a concept referred to as the dual of a vector space as it will be helpful in establishing a basis for $\mathcal{T}^k(\mathbb{R}^n)$.

Definition 3.4 For $\mathbb{R}^n$ the *dual*, denoted $(\mathbb{R}^n)^*$, is the set of linear functions from $\mathbb{R}^n$ to $\mathbb{R}$. As elements of $\mathbb{R}^n$ are called *vectors*, elements of $(\mathbb{R}^n)^*$ are called *linear functionals*.

Our first step after creating this dual vector space is to establish a basis.

Definition 3.5 A *projection function* $\pi^i : \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined for each $\boldsymbol{x} \in \mathbb{R}^n$ by $\pi^i(\boldsymbol{x}) = x^i$ for $i \in \{1, 2, \dots, n\}$.

The projection functions output the magnitude and direction of the ith basis element. For example in $\mathbb{R}^2$, the vector $\omega = (-2, 3)$ has the expansion in terms of the standard basis for $\mathbb{R}^2$ of $\omega = -2\cdot\boldsymbol{e}_1 + 3\cdot\boldsymbol{e}_2$, and $\pi^1(\omega) = -2$ while $\pi^2(\omega) = 3$.

Theorem 3.2 The set of projection functions form a basis for $(\mathbb{R}^n)^*$.

Suppose $a_1, a_2, \dots, a_n \in \mathbb{R}$ and $\left(\sum_{i=1}^{n} a_i\,\pi^i\right)(\boldsymbol{x}) = 0$ for each $\boldsymbol{x} \in \mathbb{R}$. We cleverly pick a particular $\boldsymbol{x} \in \mathbb{R}$, namely $\boldsymbol{a}$ such that $a^i = a_i$ for each $i \in \{1, 2, \dots, n\}$. Since $\left(\sum_{i=1}^{n} a_i\cdot\pi^i\right)(\boldsymbol{a}) = 0$ and $\left(\sum_{i=1}^{n} a_i\cdot\pi^i\right)(\boldsymbol{a}) = \sum_{i=1}^{n} a_i\cdot\pi^i(\boldsymbol{a}) = \sum_{i=1}^{n} a_i\,a^i = \sum_{i=1}^{n} a_i\,a_i$, then $\sum_{i=1}^{n} a_i{}^2 = 0$. Since $\sum_{i=1}^{n} a_i{}^2 = 0$, then $a_i = 0$ for each $i \in \{1, 2, \dots, n\}$. We have established the set

$\{\pi^1, \pi^2, \ldots, \pi^n\}$ is linearly independent.

Next we show $\{\pi^1, \pi^2, \ldots, \pi^n\}$ spans $(\mathbb{R}^n)^*$. Let $\varphi \in (\mathbb{R}^n)^*$ and $\quad \in \mathbb{R}^n$. Using the usual basis for $\mathbb{R}^n$ and the linear properties of φ, we write

$\varphi(x) = \varphi(a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n) = a_1 \cdot \varphi(e_1) + a_2 \cdot \varphi(e_2) + \cdots + a_n \cdot \varphi(e_n)$. Now if $x^i = 0$ for $i \in \{1, 2, \ldots, n\}$, then pick $b_i = 1$ otherwise we pick $b_i \in \mathbb{R}$ such that $b_i = \frac{\varphi\ e_i)}{\pi^i(x)}$ for each $i \in \{1, 2, \ldots, n\}$, then $\varphi(x) = a_1\, b_1 \cdot \pi^1(x) + a_2\, b_2 \cdot \pi^2(x) + \cdots + a_n\, b_n \cdot \pi^n(x)$. We can chose $c_i \in \mathbb{R}$, by $c_i = a_i\, b_i$ for each $i \in \{1, 2, \ldots, n\}$, and we have

$\varphi(x) = c_1 \cdot \pi^1(x) + c_2 \cdot \pi^2(x) + \cdots + c_n \cdot \pi^n(x)$. ∎

Sometimes questions about a vector space can be more easily answered by working in the dual and applying conclusions to the vector space. The reason why applying conclusions is valid comes from the surprising relationship between a finite dimension vector space and its dual.

Theorem 3.3 There is a bijective linear function from $\mathbb{R}^n$ to $(\mathbb{R}^n)^*$.

Proof. For each $x \in \mathbb{R}^n$ define $\varphi_x : \mathbb{R}^n \longrightarrow \mathbb{R}$ by $\varphi_x(y) = \langle x, y\rangle$ for each $y \in \mathbb{R}^n$.

We first show $\varphi_x \in (\mathbb{R}^n)^*$ for each $x \in \mathbb{R}^n$. Let $x, y, z \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$, $\varphi_x(\alpha \cdot y + \beta \cdot z) = \langle x, \alpha \cdot y + \beta \cdot z\rangle$. From properties of the inner product $\langle x, \alpha \cdot y + \beta \cdot z\rangle = \alpha \langle x, y\rangle + \beta \langle x, z\rangle = \alpha\, \varphi_x(y) + \beta\, \varphi\ \ (z)$. We have shown for each $x \in \mathbb{R}^n$, φ_x is a linear function from $\mathbb{R}^n$ to $\mathbb{R}$.

Define $T : \mathbb{R}^n \longrightarrow (\mathbb{R}^n)^*$ for each $x \in \mathbb{R}^n$ by $T(x) = \varphi\ $. Next we will show that T is injective.

Suppose $T(x) = T(y)$ for $x, y \in \mathbb{R}^n$. By definition of T, $\varphi_x(a) = \varphi_y(a)$ for all $a \in \mathbb{R}^n$. In particular, $\varphi_x(x) = \varphi_y(x)$ and $\varphi_x(y) = \varphi_y(y)$. Since $\varphi_x(x) = \varphi_y(x)$, then $\langle x, x\rangle = \langle y, x\rangle$, similarly, $\langle x, y\rangle = \langle y, y\rangle$. Since $\langle x, x\rangle = \langle y, x\rangle$, then $0 = \sum_{i=1}^{n} x^i(x^i - y^i)$, and since $\langle x, y\rangle = \langle y, y\rangle$, then $0 = \sum_{i=1}^{n} y^i(x^i - y^i)$. The difference gives $0 = \sum_{i=1}^{n} x^i(x^i - y^i) - y^i(x^i - y^i) = \sum_{i=1}^{n} (x^i - y^i)^2$. We have here the sum of non-negative terms equal to zero, therefore each term must be zero. Hence $x^i = y^i$ for each

$i \in \{1, 2, \ldots, n\}$. This is to say that $\boldsymbol{x} = \boldsymbol{y}$, and we conclude that T is injective.

Now assume $\varphi \in (\mathbb{R}^n)^*$. Using the set of projection functions $\{\pi_1, \pi_2, \ldots, \pi_n\}$ as the basis, we write φ as the unique linear combination of basis elements, i.e., $\varphi = x^1\,\pi_1 + x^2\,\pi_2 + \cdots + x^n\,\pi_n$ where $x^i \in \mathbb{R}$ for each $i \in \{1, 2, \ldots, n\}$. These claims can all be substantiated by common knowledge of linear algebra. We apply φ and its equivalent linear combination of basis elements to an arbitrary $\boldsymbol{y} \in \mathbb{R}^n$ and arrive at $\varphi(\boldsymbol{y}) = (x^1 \cdot \pi^1 + x^2 \cdot \pi^2 + \cdots + x^n \cdot \pi^n)(\boldsymbol{y})$. Using the standard definition of point-wise addition for functions, we write $(x^1 \cdot \pi^1 + x^2 \cdot \pi^2 + \cdots + x^n \cdot \pi^n)(\boldsymbol{y})$ as $x^1\,\pi^1\,(\boldsymbol{y}) + x^2\,\pi^2\,(\boldsymbol{y}) + \cdots + x^n\,\pi^n\,(\boldsymbol{y})$. Since $\pi^i\,(\boldsymbol{y}) = y^i$ for each $\boldsymbol{y} \in \mathbb{R}^n$ and $i \in \{1, 2, \ldots, n\}$, then $x^1\,\pi^1\,(\boldsymbol{y}) + x^2\,\pi^2\,(\boldsymbol{y}) + \cdots + x^n\,\pi^n\,(\boldsymbol{y}) = x^1\,y^1 + x^2\,y^2 + \cdots + x^n\,y^n$. From the meaning of the usual inner product, we know $x^1\,y^1 + x^2\,y^2 + \cdots + x^n\,y^n = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Since $\boldsymbol{x}$ is unique and $\boldsymbol{y}$ is arbitrary, then $\varphi = \varphi_{\boldsymbol{x}} = T(\boldsymbol{x})$. We have shown for each element φ of $(\mathbb{R}^n)^*$, there is an element $\boldsymbol{x}$ of $\mathbb{R}^n$ so that $T(\boldsymbol{x}) = \varphi$. We conclude T is surjective.

Since $T : \mathbb{R}^n \longrightarrow (\mathbb{R}^n)^*$ is linear, injective, and surjective, we conclude that T is a bijective linear function from $\mathbb{R}^n$ to $(\mathbb{R}^n)^*$. ∎

We recall that a k-tensor is a function that behaves multi-linearly and pairs k vectors from $\mathbb{R}^n$ with a real number. In light of our definition for the dual of $\mathbb{R}^n$, we see that a 1-tensor is merely an element of $(\mathbb{R}^n)^*$. Moreover the collection of all 1-tensors, $\mathcal{T}^1(\mathbb{R}^n)$, is the same as $(\mathbb{R}^n)^*$. At first this does not seems like a moving realization, but then we use our definition for the tensor product "$\otimes$" and with a great deal of complexity, we find that we can build any k-tensor space out of elements from the dual of $\mathbb{R}^n$.

We start by considering the simplest k-tensor space that exemplifies the complexities of the general case but keeps the number of tensor terms to a manageable size. Consider $\omega \in \mathcal{T}^2(\mathbb{R}^3)$ and $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^3$. We use the usual basis of $\mathbb{R}^3$ to redescribe our two arbitrary vectors as a unique linear combination of the basis elements: $\boldsymbol{x}_1 = a_{1,1} \cdot \boldsymbol{e}_1 + a_{1,2} \cdot \boldsymbol{e}_2 + a_{1,3} \cdot \boldsymbol{e}_3$ and $\boldsymbol{x}_2 = a_{2,1} \cdot \boldsymbol{e}_1 + a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3$ for $a_{i,j} \in \mathbb{R}$,

$i \in \{1, 2\}$, $j \in \{1, 2, 3\}$. We apply ω to the linear expansion of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ and carefully apply the multilinear property of ω.

$$
\begin{aligned}
\omega(\boldsymbol{x}_1, \boldsymbol{x}_2) &= \omega(a_{1,1} \cdot \boldsymbol{e}_1 + a_{1,2} \cdot \boldsymbol{e}_2 + a_{1,3} \cdot \boldsymbol{e}_3,\ a_{2,1} \cdot \boldsymbol{e}_1 + a_{2,2} \cdot \boldsymbol{e}_2 - a_{2,3} \cdot \boldsymbol{e}_3) \\
&= a_{1,1}\, \omega(\boldsymbol{e}_1,\ a_{2,1} \cdot \boldsymbol{e}_1 + a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) + \\
&\qquad \omega(a_{1,2} \cdot \boldsymbol{e}_2 + a_{1,3} \cdot \boldsymbol{e}_3,\ a_{2,1} \cdot \boldsymbol{e}_1 + a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) \\
&= a_{1,1}\, a_{2,1}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_1) + a_{1,1}\, \omega(\boldsymbol{e}_1,\ a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) + \\
&\qquad a_{1,2}\, \omega(\boldsymbol{e}_2,\ a_{2,1} \cdot \boldsymbol{e}_1 + a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) + \\
&\qquad\qquad a_{1,2}\, \omega(\boldsymbol{e}_3,\ a_{2,1} \cdot \boldsymbol{e}_1 + a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) \\
&= a_{1,1}\, a_{2,1}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_1) + a_{1,1}\, a_{2,2}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_2) + a_{1,1}\, a_{2,3}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_3) + \\
&\qquad a_{1,2}\, a_{2,1}\, \omega(\boldsymbol{e}_2, \boldsymbol{e}_1) + a_{1,2}\, \omega(\boldsymbol{e}_2,\ a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) - \\
&\qquad a_{1,2}\, a_{2,1}\, \omega(\boldsymbol{e}_3, \boldsymbol{e}_1) + a_{1,2}\, \omega(\boldsymbol{e}_3,\ a_{2,2} \cdot \boldsymbol{e}_2 + a_{2,3} \cdot \boldsymbol{e}_3) \\
&= a_{1,1}\, a_{2,1}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_1) + a_{1,1}\, a_{2,2}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_2) + a_{1,1}\, a_{2,3}\, \omega(\boldsymbol{e}_1, \boldsymbol{e}_3) + \\
&\qquad a_{1,2}\, a_{2,1}\, \omega(\boldsymbol{e}_2, \boldsymbol{e}_1) + a_{1,2}\, a_{2,2}\, \omega(\boldsymbol{e}_2, \boldsymbol{e}_2) - a_{1,2}\, a_{2,3}\, \omega(\boldsymbol{e}_2, \boldsymbol{e}_3) + \\
&\qquad a_{1,2}\, a_{2,1}\, \omega(\boldsymbol{e}_3, \boldsymbol{e}_1) + a_{1,2}\, a_{2,2}\, \omega(\boldsymbol{e}_3, \boldsymbol{e}_2) + a_{1,2}\, a_{2,3}\, \omega(\boldsymbol{e}_3, \boldsymbol{e}_3) \\
&= \sum_{i_1=1}^{3} \left(\sum_{i_2=1}^{3} a_{1,i_1}\, a_{2,i_2}\, \omega(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}) \right)
\end{aligned}
$$

Next we experiment by applying $\pi^i \otimes \pi^j$ to ${}_1$ and $\boldsymbol{x}_2$ for $i, j \in \{1, 2, 3\}$.

$\pi^1 \otimes \pi^1 \ (x_1, x_2)$

$$= \pi^1 \otimes \pi^1 (a_{1,1} \cdot e_1 + a_{1,2} \cdot e_2 + a_{1,3} \cdot e_3, a_{2,1} \cdot e_1 + a_{2,2} \cdot e_2 + a_{2,3} \cdot e_3)$$

$$= \pi^1 (a_{1,1} \cdot e_1 + a_{1,2} \cdot e_2 + a_{1,3} \cdot e_3) \pi^1 (a_{2,1} \cdot e_1 + a_{2,2} \cdot e_2 + a_{2,3} \cdot e_3)$$

$$= (a_{1,1} \, \pi^1(e_1) + a_{1,2} \, \pi^1(e_2) + a_{1,3} \, \pi^1(e_3)) \cdot$$

$$(a_{2,1} \, \pi^1(e_1) + a_{2,2} \, \pi^1(e_2) + a_{2,3} \, \pi^1(e_3))$$

$$= (a_{1,1} \cdot 1 + a_{1,2} \cdot 0 + a_{1,2} \cdot 0) \, (a_{2,1} \cdot 1 + a_{2,1} \cdot 0 + a_{2,3} \cdot 0)$$

$$= a_{1,1} \, a_{2,1}$$

$$\pi^1 \otimes \pi^2 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,1} \, a_{2,2}$$

$$\pi^1 \otimes \pi^3 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,1} \, a_{2,3}$$

$$\pi^2 \otimes \pi^1 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,2} \, a_{2,1}$$

$$\pi^2 \otimes \pi^2 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,2} \, a_{2,2}$$

$$\pi^2 \otimes \pi^3 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,2} \, a_{2,3}$$

$$\pi^3 \otimes \pi^1 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,3} \, a_{2,1}$$

$$\pi^3 \otimes \pi^2 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,3} \, a_{2,2}$$

$$\pi^3 \otimes \pi^3 \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \ = a_{1,3} \, a_{2,3}$$

We combine these results with $\omega(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{\iota_1=1}^{3} \left(\sum_{\iota_2=1}^{3} a_{1,\iota_1} \, a_{2,\iota_2} \, \omega(\boldsymbol{e}_{\iota_1}, \boldsymbol{e}_{\iota_2}) \right)$ to show $\omega(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{\iota_1=1}^{3} \left(\sum_{\iota_2=1}^{3} \pi^{\iota_1} \otimes \pi^{\iota_2} \ (\boldsymbol{x}_1, \boldsymbol{x}_2) \, \omega(\boldsymbol{e}_{\iota_1}, \boldsymbol{e}_{\iota_2}) \right)$, which is a linear combination of the elements from the $\{\pi^i \otimes \pi^j \mid 1 \le i, \, j \le 3\}$. We conclude that the $\{\pi^i \otimes \pi^j \mid 1 \le i, \, j \le 3\}$ span $\mathcal{T}^2(\mathbb{R}^3)$. Next we show that this set is linearly independent.

Suppose $0 = \sum_{\iota_1=1}^{3} \left(\sum_{\iota_2=1}^{3} a_{\iota_1,\iota_2} \, \pi^{\iota_1} \otimes \pi^{\iota_2} \right)$, where each $a_{\iota_1,\iota_2} \in \mathbb{R}$ and 0 in this context is the 2-tensor that takes pairs of elements of $\mathbb{R}^3$ to the number 0. We must show each $a_{\iota_1,\iota_2} = 0$. We apply $\sum_{\iota_1=1}^{3} \left(\sum_{\iota_2=1}^{3} a_{\iota_1,\iota_2} \, \pi^{\iota_1} \otimes \pi^{\iota_2} \right)$ to $\boldsymbol{e}_1$ and $\boldsymbol{e}_1$.

$$
\begin{aligned}
0 &= \left(\sum_{i_1=1}^{3}\left(\sum_{i_2=1}^{3} a_{i_1,i_2}\, \pi^{i_1} \otimes \pi^{i_2}\right)\right)(e_1, e_1) \\
&= \sum_{i_1=1}^{3}\left(\sum_{i_2=1}^{3} a_{i_1,i_2}\, [(\pi^{i_1} \otimes \pi^{i_2})(e_1, e_1)]\right) \\
&= \sum_{i_1=1}^{3}\left(\sum_{i_2=1}^{3} a_{i_1,i_2}\, \pi^{i_1}(e_1)\pi^{i_2}(e_1)\right) \\
&= a_{1,1}\, \pi^1(e_1)\pi^1(e_1) + a_{1,2}\, \pi^1(e_1)\pi^2(e_1) + a_{1,3}\, \pi^1(e_1)\pi^3(e_1) + \\
&\quad a_{2,1}\, \pi^2(e_1)\pi^1(e_1) + a_{2,2}\, \pi^2(e_1)\pi^2(e_1) + a_{2,3}\, \pi^2(e_1)\pi^3(e_1) + \\
&\quad a_{3,1}\, \pi^3(e_1)\pi^1(e_1) + a_{3,2}\, \pi^3(e_1)\pi^2(e_1) + a_{3,3}\, \pi^3(e_1)\pi^3(e_1) \\
&= a_{1,1} \cdot 1 \cdot 1 + a_{1,2}\, 1 \cdot 0 + a_{1,3}\, 1 \cdot 0 + \\
&\quad a_{2,1}\, 0 \cdot 1 + a_{2,2}\, 0 \cdot 0 + a_{2,3}\, 0 \cdot 0 + \\
&\quad a_{3,1}\, 0 \cdot 1 + a_{3,2}\, 0 \cdot 0 + a_{3,3}\, 0 \cdot 0 \\
&= a_{1,1}
\end{aligned}
$$

We have shown that $a_{1,1} = 0$. We can apply the same 2-tensor to e_1 and e_2 to show $a_{1,2} = 0$. Following the pattern, we can apply the same 2-tensor to e_{i_1} and e_{i_2} to show $a_{i_1,i_2} = 0$ for $i_1, i_2 \in \{1, 2, 3\}$. Since each $a_{i_1,i_2} = 0$, we conclude that the $\{\pi^i \otimes \pi^j \mid 1 \le i, j \le 3\}$ is a linearly independent set.

We have shown that in the case of $\mathcal{T}^2(\mathbb{R}^3)$, the basis is $\{\pi \otimes \pi \mid 1 \le i, j \le 3\}$, so $\mathcal{T}^2(\mathbb{R}^3)$ has dimension 3^2 or 9. This example is good to work through before trying to understands the general case. We will show $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \le i_1, \ldots, i_k \le n\}$ is a basis for $\mathcal{T}^k(\mathbb{R}^n)$, which will therefore have dimension n^k From this last comment on the dimension of $\mathcal{T}^k(\mathbb{R}^n)$, we can see how quickly the number of terms in a linear combination of the basis elements will grow to an unmanageable size. We must therefore gather our wits in the use summation notation and carefully follow our example of $\mathcal{T}^2(\mathbb{R}^3)$ to prove the following theorem.

Theorem 3.4 Suppose $\mathcal{T}^k(\mathbb{R}^n)$ is the k-tensor space of $\mathbb{R}^n$, and $\{e_1, e_2, \ldots, e_n\}$ is the usual basis for $\mathbb{R}^n$. If $\{\pi^1, \pi^2, \ldots, \pi^n\}$ is the usual basis for $(\mathbb{R}^n)^*$, then the set of k-fold tensor products $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \le i_1, \ldots, i_k \le n\}$ is a basis for $\mathcal{T}^k(\mathbb{R}^n)$.

Proof. Let $\{\pi^1, \pi^2, \ldots, \pi^n\}$ be the usual basis for $(\mathbb{R}^n)^*$, which means if $\in \mathbb{R}^n$, then $\pi^i(x) = x^i$ for $i \in \{1, 2, \ldots, n\}$. Let ω be a k-tensor in $\mathcal{T}^k(\mathbb{R}^n)$, and let $x_1, x_2, \ldots, x_k$ be k vectors in $\mathbb{R}^n$. If for each $i \in \{1, 2, \ldots, k\}$, we write x_i as a linear combination of elements from $\{e_1, e_2, \ldots, e_n\}$, then $_i = a_{i,1} e_1 + a_{i,2} e_2 + \cdots + a_{i,n} e_n$ with $a_{i,j} \in \mathbb{R}$ for each $i \in \{1, 2, \ldots, k\}$ and $j \in \{1, 2, \ldots, n\}$. We apply ω to our k vectors from $\mathbb{R}^n$ and following lessons learned from the example $\mathcal{T}^2(\mathbb{R}^3)$, we find

$\omega(x_1, x_2, \ldots, x_k) = \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n} \cdots \left(\sum_{i_k=1}^{n} a_{1,i_1} a_{2,i_2} \cdots a_{k,i_k}\, \omega(e_{i_1}, e_{i_2}, \ldots, e_{i_k})\right) \cdots\right)$. Based on our experience from the example $\mathcal{T}^2(\mathbb{R}^3)$, it is not difficult to inductively see that

$\pi^{i_1} \otimes \pi^{i_2} \otimes \cdots \otimes \pi^{i_k}(x_1, x_2, \ldots, x_k) = a_{1,i_1} a_{2,i_2} \cdots a_{k,i_k}$ for each $i_1, i_2, \ldots i_k \in \{1, 2, \ldots, n\}$, so we make this substitution and have

$\omega(x_1, \ldots, x_k) = \sum_{i_1=1}^{n}\left(\cdots\left(\sum_{i_k=1}^{n} \pi^{i_1} \otimes \cdots \otimes \pi^{i_k}(x_1, \ldots, x_k)\, \omega(e_{i_1}, \ldots, e_{i_k})\right)\cdots\right)$. Since the vectors $x_1, x_2, \ldots, x_k$ were arbitrary then

$\omega = \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n} \cdots\left(\sum_{i_k=1}^{n} \omega(e_{i_1}, e_{i_2}, \ldots, e_{i_k}) \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \cdots \otimes \pi^{i_k}\right)\cdots\right)$. Since ω was arbitrary in $\mathcal{T}^k(\mathbb{R}^n)$, then we have shown the $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \le i_1, \ldots, i_k \le n\}$ span $\mathcal{T}^k(\mathbb{R}^n)$.

Next we show the $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \le i_1, \ldots, i_k \le n\}$ is linearly independent.

Suppose $0 = \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n} \cdots\left(\sum_{i_k=1}^{n} a_{i_1,i_2,\ldots,i_k} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \cdots \otimes \pi^{i_k}\right)\cdots\right)$ where the zero here is the k-tensor that takes elements of $\mathbb{R}^n$ to the number 0. We apply this k-tensor to k elements chosen (with repetition allowed) from $\{e_1, e_2, \ldots, e_n\} \subset \mathbb{R}^n$.

$$\begin{aligned}
0 &= \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n} \cdots\left(\sum_{i_k=1}^{n} a_{i_1,i_2,\ldots,i_k} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \cdots \otimes \pi^{i_k}\right)\cdots\right)(e_{j_1}, e_{j_2}, \ldots, e_{j_k}) \\
&= \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n} \cdots\left(\sum_{i_k=1}^{n} a_{i_1,i_2,\ldots,i_k} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \cdots \otimes \pi^{i_k}(e_{j_1}, e_{j_2}, \ldots, e_{j_k})\right)\cdots\right) \\
&= \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n} \cdots\left(\sum_{i_k=1}^{n} a_{i_1,i_2,\ldots,i_k} \cdot \pi^{i_1}(e_{j_1})\, \pi^{i_2}(e_{j_2}) \cdots \pi^{i_k}(e_{j_k})\right)\cdots\right) \\
&= a_{j_1,j_2,\ldots,j_k}
\end{aligned}$$

From the example, every term in this complex sum turns out to be zero except the one where $i_l = j_l$ for every $l \in \{1, 2, \ldots, k\}$; hence the sum collapses to the single term $a_{j_1, j_2, \ldots, j_k}$. We have shown that $0 = a_{j_1, j_2, \ldots, j_k}$ for each $j_1, j_2, \ldots, j_k \in \{1, 2, \ldots, n\}$ since our choice of $e_{j_1}, e_{j_2}, \ldots, e_{j_k} \in \{e_1, e_2, \ldots, e_n\} \subset \mathbb{R}^n$ was arbitrary. We conclude the $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \leq i_1, \ldots, i_k \leq n\}$ is linearly independent.

Since $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \leq i_1, \ldots, i_k \leq n\}$ is a linearly independent set of order n^k that spans $\mathcal{T}^k(\mathbb{R}^n)$, we conclude that $\{\pi^{i_1} \otimes \cdots \otimes \pi^{i_k} \mid 1 \leq i_1, \ldots, i_k \leq n\}$ is a basis for $\mathcal{T}^k(\mathbb{R}^n)$, which has dimension n^k. ∎

In the example and theorem, we have brought an importance of the dual space to light. We can build any k-tensor space or subspace from elements of $(\mathbb{R}^n)^*$, which from Theorem 3.3, is structurally the same as $\mathbb{R}^n$.

Alternating k-tensors

In this section we take the next step in increasing the complexity of our work Ironically, we do this by investigating a subset of the k-tensor space we just developed in the last section. Specializing will bring k-tensors together with the concept of differential forms, a key idea in the generalized Stokes' Theorem.

Definition 3.6 A k-tensor $\omega \in \mathcal{T}^k(\mathbb{R}^n)$ is *alternating* if for $_1, x_2, \ldots, x_k \in \mathbb{R}^n$, $\omega(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_k) = -\omega(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k)$ for $i \neq j$ with $i, j \in \{1, 2, \ldots, k\}$.

The next natural step after defining a special type of k-tensor is to collect all the special k-tensors of this type together. We designate this collection $\Lambda^k(\mathbb{R}^n)$. If $\omega, \phi \in \Lambda^k(\mathbb{R}^n)$ and $a \in \mathbb{R}$ then

$$(\omega + \phi)(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_k)$$

$$= -1 \cdot \omega(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k) +$$

$$-1 \cdot \phi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k)$$

$$= -1 \cdot (\omega + \phi)(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_k)$$

so $\omega + \phi = -(\omega + \phi)$ and

$$
\begin{aligned}
(a \cdot \omega)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_k) & \\
&= a\omega(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_k) \\
&= a \cdot -\omega(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_k) \\
&= -1 \cdot a\omega(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_k) \\
&= -(a \cdot \omega)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_k)
\end{aligned}
$$

so $a \cdot \omega = -(a \cdot \omega)$.

By the elementary theorems of linear algebra $\Lambda^k(\mathbb{R}^n)$ is then a subspace of $\mathcal{T}^k(\mathbb{R}^n)$ or vector space in its own right.

We will be interested in finding a basis, but surprisingly will find the quest even more challenging than the one to find a basis for $\mathcal{T}^k(\mathbb{R}^n)$. We cannot build the basis elements with $\otimes$ out of the projection function $\pi^i . \mathbb{R}^n \longrightarrow \mathbb{R}$ because if each of a k and l-tensor is alternating, their tensor product is not necessarily alternating We start the quest by defining a function that pairs a k-tensor with an alternating k-tensor To make the definition we need to be reminded of some background in Group Theory.

Recall S_n is usually reserved in Group Theory to represent the symmetric group of permutations on n symbols If 123456 is a sequence of six symbols, then 153426 is a transposition, as well as a permutation, because exactly two symbols are interchanged. The sequence 632154 is another permutation that can be achieved from 123456 by the following sequence of transpositions:

1) transposing the symbols in the first and fourth positions, 423156;

2) transposing the symbols in the first and sixth positions, 623154;

3) transposing the symbols in the second and third positions, 632154.

While one can see the order of this sequence is not unique, from group theory we know every permutation can be decomposed into a minimum number of transpositions. The permutation 632154 required 3 transpositions, and we call it odd since the number of transpositions is an odd number We call some other permutation even, if it can be decomposed into an even number of transpositions. Now for our last note on group theory before we make our definition, if $\sigma \in S_n$ then $\text{sgn}(\sigma) = 1$ if σ is even and $\text{sgn}(\sigma) = -1$ if σ is odd.

Definition 3.7 For $\omega \in \mathcal{T}^k(\mathbb{R}^n)$ we define $\text{Alt}(\omega) \in \Lambda^k(\mathbb{R}^n)$ for $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$ by $\text{Alt}(\omega)(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \cdot \omega(\boldsymbol{x}_{\sigma(1)}, \boldsymbol{x}_{\sigma(2)}, \ldots, \boldsymbol{x}_{\sigma(k)})$.

Now we must show that the definition for $\text{Alt}(\cdot)$ does indeed produce and alternating k-tensor. This essentially turns into an observation of the fact, but we make it a Theorem.

Theorem 3.5 If $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, then $\text{Alt}(\omega) \in \Lambda^k(\mathbb{R}^n)$.

Proof. Let $\omega \in \mathcal{T}^k(\mathbb{R}^n)$ and $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$. Since $\text{sgn}(\sigma)$ is determined by the number of transpositions σ can be decomposed into, then composing σ with an additional transposition will simply change the sign of $\text{sgn}(\sigma)$. Therefore, if $\sigma \in S_k$ and (i, j) is the transposition that interchanges i and j for $i, j \in \{1, 2, \ldots, k\}$, then $\text{sgn}(\sigma \cdot (i, j)) = -\text{sgn}(\sigma)$.

$$\begin{aligned}
&\text{Alt}(\omega)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_k) \\
&\quad = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma)\, \omega(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(j)}, \ldots, \boldsymbol{x}_{\sigma(i)}, \ldots, \boldsymbol{x}_{\sigma(k)}) \\
&\quad = \frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma \cdot (i, j))\, \omega(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(i)}, \ldots, \boldsymbol{x}_{\sigma(j)}, \ldots, \boldsymbol{x}_{\sigma(k)}) \\
&\quad = \frac{1}{k!} \sum_{\sigma \in S_k} - \text{sgn}(\sigma)\, \omega(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(i)}, \ldots, \boldsymbol{x}_{\sigma(j)}, \ldots, \boldsymbol{x}_{\sigma(k)}) \\
&\quad = -\frac{1}{k!} \sum_{\sigma \in S_k} \text{sgn}(\sigma)\, \omega(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(i)}, \ldots, \boldsymbol{x}_{\sigma(j)}, \ldots, \boldsymbol{x}_{\sigma(k)}) \\
&\quad = -\text{Alt}(\omega)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_j, \ldots, \boldsymbol{x}_k)
\end{aligned}$$

We have shown ω is alternating which allows the conclusion $\text{Alt}(\omega) \in \Lambda^k(\mathbb{R}^n)$.

Now what effect does $\text{Alt}(\cdot)$ have on a k-tensor that is already alternating? We will investigate this in an example and then generalize it in a theorem. Suppose $\omega \in \Lambda^3(\mathbb{R}^n)$.

$\mathrm{Alt}(\omega)\,(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3)$

$$= \tfrac{1}{3!} \sum_{\sigma \in S_3} \mathrm{sgn}(\sigma)\, \omega(\boldsymbol{x}_{\sigma(1)}, \boldsymbol{x}_{\sigma(2)}, \boldsymbol{x}_{\sigma(3)}) \qquad \textit{Def 3.7}$$

$$= \tfrac{1}{6}\,(\omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) - \omega(\boldsymbol{x}_1, \boldsymbol{x}_3, \boldsymbol{x}_2) + \omega(\boldsymbol{x}_3, \boldsymbol{x}_1, \boldsymbol{x}_2) - \omega(\boldsymbol{x}_3, \boldsymbol{x}_2, \boldsymbol{x}_1) + \omega(\boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_1) - \omega(\boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{x}_3)) \qquad \text{expanding the sum}$$

$$= \tfrac{1}{6}\,(\omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) - \omega(\boldsymbol{x}_1, \boldsymbol{x}_3, \boldsymbol{x}_2) + \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) - \omega(\boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{x}_3)) \qquad \textit{Def 3.6}$$

$$= \tfrac{1}{6}\,(\omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) + \omega(\boldsymbol{x}_2, \boldsymbol{x}_1, \boldsymbol{x}_3)) \qquad \textit{Def 3.6}$$

$$= \tfrac{1}{6}\,(6\,\omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3))$$

$$= \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \qquad \blacksquare$$

Through the process of making the order of the sequence of $\boldsymbol{x}_i$, $\boldsymbol{x}_j$, $\boldsymbol{x}_k$ identical for each of the 3! terms all the -1 coefficients become $+1$. In words, if it takes an odd number of transpositions to obtain a particular permutation then it takes odd number of transpositions to undo the permutation so $-1 \cdot -1 = +1$. Similarly for an even permutation, $1 \cdot 1 = +1$. In the language of group theory, a permutation and its inverse have the same sign, i.e., $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$. This is the basis for the following theorem.

Theorem 3.6 If $\omega \in \Lambda^k(\mathbb{R}^n)$, then $\mathrm{Alt}(\omega) = \omega$.

Proof. Suppose $\omega \in \Lambda^k(\mathbb{R}^n)$. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k \in \mathbb{R}^n$. For each $\sigma \in S_k$, $\mathrm{sgn}(\sigma)\,\mathrm{sgn}(\sigma^{-1}) = 1$ and $\omega(\boldsymbol{x}_{(\sigma\,\sigma^{-1})(1)}, \boldsymbol{x}_{(\sigma\,\sigma^{-1})(2)}, \ldots, \boldsymbol{x}_{(\sigma\,\sigma^{-1})(k)}) = \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)$.

$$\begin{aligned}
&\mathrm{Alt}(\omega)\,(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) \\
&\quad = \tfrac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\, \omega(\boldsymbol{x}_{\sigma(1)}, \boldsymbol{x}_{\sigma(2)}, \ldots, \boldsymbol{x}_{\sigma(k)}) \\
&\quad = \tfrac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\sigma^{-1})\, \omega(\boldsymbol{x}_{(\sigma\sigma^{-1})(1)}, \boldsymbol{x}_{(\sigma\sigma^{-1})(2)}, \ldots, \boldsymbol{x}_{(\sigma\sigma^{-1})(k)}) \\
&\quad = \tfrac{1}{k!}\,(k! \cdot \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)) \\
&\quad = \omega(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k)
\end{aligned}$$

where the factor $k!$ comes from the fact that S_k has $k!$ elements and hence $k!$ terms in the sum over the elements in S_k. We conclude that Alt(ω) = ω. ∎

We see in this theorem where the factor $\frac{1}{k!}$ comes from in the definition of Alt($\cdot$). The factor is not necessary to make a k-tensor alternating as can be seen from Theorem 3.5, but to make Theorem 3.6 true, it is necessary. We use both the previous theorems for our next theorem.

Theorem 3.7 If $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, then Alt(Alt($\omega$)) = Alt($\omega$).

By Theorem 3.5, Alt(ω) $\in \Lambda^k(\mathbb{R}^n)$. Since Alt($\omega$) $\in \Lambda^k(\mathbb{R}^n)$, then by Theorem 3.6 Alt(Alt(ω)) = Alt(ω). ∎

Continuing on our quest to write a basis for $\Lambda^k(\mathbb{R}^n)$, we must take care of the original problem in that for $\omega \in \Lambda^k(\mathbb{R}^n)$ and $\nu \in \Lambda^l(\mathbb{R}^n)$, $\omega \otimes \nu$ is not necessarily part of $\Lambda^{k+l}(\mathbb{R}^n)$. We therefore use our definition of Alt($\cdot$) together with $\otimes$ to write a new tensor product between alternating tensors called the wedge product.

Definition 3.8 For $\omega \in \Lambda^k(\mathbb{R}^n)$ and $\nu \in \Lambda^l(\mathbb{R}^n)$, we define the wedge product $\omega \wedge \nu$ as $\frac{(k+l)!}{k!\,l!}$ Alt($\omega \otimes \nu$).

To gain experience with the wedge product, we work out several properties of this new binary operation. The proofs are all trivial in theory since they involve mostly definitions to follow the reasoning, however, there is a great complexity in the meaning of the notation.

Theorem 3.8 Suppose ω, ω_1, $\omega_2 \in \Lambda^k(\mathbb{R}^n)$, η, η_1, $\eta_2 \in \Lambda^l(\mathbb{R}^n)$, and $\alpha \in \mathbb{R}$. The following are properties of the wedge product:

1.) $(\omega_1 + \omega_2) \wedge \eta = \omega_1 + \eta \wedge \omega_2 + \eta$,

2.) $\omega \wedge (\eta_1 + \eta_2) = \omega \wedge \eta_1 + \omega \wedge \eta_2$,

3.) $a\omega \wedge \eta = \omega \wedge a\eta = a(\omega \wedge \eta)$,

4.) $\omega \wedge \eta = (-1)^{kl}\, \eta \wedge \omega$.

Proof. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{k+l} \in \mathbb{R}^n$.

1) $(\omega_1 + \omega_2) \wedge \eta(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$

$$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}((\omega_1 + \omega_2) \otimes \eta)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Def 3.8}$$

$$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega_1 \otimes \eta + \omega_2 \otimes \eta)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Thrm 3.1}$$

$$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\,(\omega_1 \otimes \eta + \omega_2 \otimes \eta)\,(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}) \qquad \textit{Def 3.7}$$

$$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\,(\omega_1 \otimes \eta(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}) \qquad \textit{Def 3.2}$$

$$+ \omega_2 \otimes \eta(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}))$$

$$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\,\omega_1 \otimes \eta(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$$

$$+ \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\,\omega_2 \otimes \eta(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$$

$$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega_1 \otimes \eta)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) + \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega_2 \otimes \eta)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Def 3.7}$$

$$= \omega_1 \wedge \eta(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) + \omega_2 \wedge \eta(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Def 3.8}$$

$$= (\omega_1 \wedge \eta + \omega_2 \wedge \eta)\,(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Def 3.2}$$

2) $\omega \wedge (\eta_1 + \eta_2)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$

$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes (\eta_1 + \eta_2))(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.8*

$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta_1 + \omega \otimes \eta_2)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Thrm 3.1*

$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)(\omega \otimes \eta_1 + \omega \otimes \eta_2)(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$ *Def 3.7*

$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)(\omega \otimes \eta_1 (\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}) +$ *Def 3.2*

$\omega \otimes \eta_2 (\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}))$

$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, \omega \otimes \eta_1 (\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}) +$

$\frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \omega \otimes \eta_2 (\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$

$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta_1)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) + \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta_2)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.7*

$= \omega \wedge \eta_1 (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) + \omega \wedge \eta_2 (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.8*

$= (\omega \wedge \eta_1 + \omega \wedge \eta_2)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.2*

3) $a(\omega \wedge \eta)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$

$= a\, \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.8*

$= a\, \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, \omega \otimes \eta(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$ *Def 3.7*

$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, a(\omega \otimes \eta)(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$

$= \frac{(k+l)!}{k!\,l!} \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)((a \cdot \omega) \otimes \eta)(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})$ *Thrm 3.1*

$= \frac{(k+l)!}{k!\,l!} \operatorname{Alt}((a \cdot \omega) \otimes \eta)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.7*

$= (a \cdot \omega) \wedge \eta(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$ *Def 3.8*

4) $\omega \wedge \eta(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l})$

$$= \tfrac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta)(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Def 3.8}$$

$$= \tfrac{(k+l)!}{k!\,l!} \tfrac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, \omega \otimes \eta(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k)}, \boldsymbol{x}_{\sigma(k+1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}) \qquad \textit{Def 3.7}$$

$$= \tfrac{(k+l)!}{k!\,l!} \tfrac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, \omega(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k)})\, \eta(\boldsymbol{x}_{\sigma(k+1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}) \qquad \textit{Def 3.3}$$

$$= \tfrac{(k+l)!}{k!\,l!} \tfrac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, \eta(\boldsymbol{x}_{\sigma(k+1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)})\, \omega(\boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k)})$$

$$= \tfrac{(k+l)!}{k!\,l!} \tfrac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)\, \eta \otimes \omega(\boldsymbol{x}_{\sigma(k+1)}, \ldots, \boldsymbol{x}_{\sigma(k+l)}, \boldsymbol{x}_{\sigma(1)}, \ldots, \boldsymbol{x}_{\sigma(k)}) \qquad \textit{Def 3.3}$$

$$= \tfrac{(k+l)!}{k!\,l!} \operatorname{Alt}(\eta \otimes \omega)(\boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_{k+l}, \boldsymbol{x}_{\sigma 1}, \ldots, \boldsymbol{x}_k) \qquad \textit{Def 3.7}$$

$$= \tfrac{(k+l)!}{k!\,l!} (-1)^{l} \operatorname{Alt}(\eta \otimes \omega)(\boldsymbol{x}_1, \boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_{k+l}, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k) \qquad \textit{Thrm 3.5}$$

$$= \tfrac{(k+l)!}{k!\,l!} (-1)^{2l} \operatorname{Alt}(\eta \otimes \omega)(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_{k+l}, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_k) \qquad \textit{Thrm 3.5}$$

$$= \tfrac{(k+l)!}{k!\,l!} (-1)^{(k-2)l} \operatorname{Alt}(\eta \otimes \omega)(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{k-2}, \boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_{k+l}, \boldsymbol{x}_{k-1}, \boldsymbol{x}_k) \qquad \textit{Thrm 3.5}$$

$$= \tfrac{(k+l)!}{k!\,l!} (-1)^{kl} \operatorname{Alt}(\eta \otimes \omega)(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k, \boldsymbol{x}_{k+1}, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Thrm 3.5}$$

$$= (-1)^{kl}\, \eta \wedge \omega(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+l}) \qquad \textit{Def 3.8}$$

The properties above will be essential to progress on our quest for a basis for $\Lambda^k(\mathbb{R}^n)$. We have one last property of the wedge product that turns out to be no triviality. We would like the wedge product to be associative to make it a useful tool in building our basis.

Theorem 3.9 Suppose $\omega \in \mathcal{T}^k(\mathbb{R}^n)$ and $\eta \in \mathcal{T}^l(\mathbb{R}^n)$. If $\operatorname{Alt}(\omega) = 0$, then $\operatorname{Alt}(\omega \otimes \eta) = 0$ and $\operatorname{Alt}(\eta \otimes \omega) = 0$.

Proof. Suppose $\operatorname{Alt}(\omega) = 0$.

Let $G = \{\sigma \in S_{k+l} \mid \sigma(k+i) = k+i \text{ for } i = 1, 2, \ldots, l\}$. Suppose $\sigma, \tau \in G$. Since $\tau \in G$, then for $i \in \{1, 2, \ldots, l\}$, $\tau(k+i) = k+i$. Since S_{k+l} is a group then $\tau^{-1} \in S_{k+l}$ and $\tau^{-1}\,\tau(k+i) = \tau^{-1}(k+i)$. Now $\tau^{-1}\,\tau(k+i) = \iota(k+i) = k+i$ so $k+i = \tau^{-1}(k+i)$. We have shown $\tau^{-1} \in$. As the litmus test of a subset being a subgroup we must show $\sigma\tau^{-1} \in G$. This is obvious by applying $\sigma\tau^{-1}$ to $k + i$ for $i \in \{1, 2, \ldots, l\}$. Therefore G is a subgroup of S_{k+l}. Since G is in essence the permutation group on k symbols, then $|G| = k!$.

We recall from group theory that for some $\sigma \in S_{k+l}$ a right coset is $G\sigma = \{g\sigma \mid g \in G\}$, and the right cosets partition S_{k+l}. Since the definition of $\mathrm{Alt}(\cdot)$ gives us a sum over S_{k+l}, we will group this sum by right cosets and make a generalizable conclusion for each grouped sum. To do this we pick a subset S of S_{k+l} where each element of S is in a distinct coset from any other element in S and each right coset has a representative element in S. We have then $|S|$ is the same as the number of right cosets. We write then $\mathrm{Alt}(\omega \otimes \eta) = \frac{1}{(k+l)!} \sum_{\tau \in S} \left(\sum_{\sigma \in G\tau} \mathrm{sgn}(\sigma)\, \omega \otimes \eta \right)$.

Let $\tau \in S$. $G\tau$ is a right coset of S_{k+l}. Let $x_1, x_2, \ldots, x_{k+l} \in \mathbb{R}^n$.

By Definition 3.3, $\frac{1}{(k+l)!} \sum_{\sigma \in G\tau} \mathrm{sgn}(\sigma)\, \omega \otimes \eta(x_{\sigma(1)}, \ldots, x_{\sigma(k+l)}) =$ $\frac{1}{(k+l)!} \sum_{\sigma \in G\tau} \mathrm{sgn}(\sigma)\, \omega(x_{\sigma(1)}, \ldots, x_{\sigma(k)})\, \eta(x_{\sigma(k+1)}, \ldots, x_{\sigma(k+l)})$. Since $|G| = k!$ then each $\sigma \in G\tau$ is $g_i \tau$ for $i \in \{1, 2, \ldots, k!\}$ and $g_i \in G$. We rename each vector by applying τ, so for each $i \in \{1, 2, \ldots, k!\}$ and $j \in \{1, 2, \ldots, k + l\}$, $x_{(g_i \tau)(j)} = w_{g_i(j)}$. We use this to rewrite our sum. $\frac{1}{(k+l)!} \sum_{\sigma \in G\tau} \mathrm{sgn}(\sigma)\, \omega(x_{\sigma(1)}, \ldots, x_{\sigma(k)})\, \eta(x_{\sigma(k+1)}, \ldots, x_{\sigma(k+l)}) =$ $\frac{1}{(k+l)!} \sum_{i=1}^{k!} \mathrm{sgn}(g_i \tau)\, \omega(w_{g_i(1)}, \ldots, w_{g_i(k)})\, \eta(w_{g_i(k+1)}, \ldots, w_{g_i(k+l)})$. Note that by an argument that follows the sum of odd and odd or even and even is even while the sum of odd and even is odd, we write $\mathrm{sgn}(g_i \tau) = \mathrm{sgn}(g_i)\,\mathrm{sgn}(\tau)$ for each $i \in \{1, 2, \ldots, k!\}$. Also since for each $i \in \{1, 2, \ldots, k!\}$, $g_i(k + j) = k + j$ for $j \in \{k + 1, \ldots, k + l\}$, then

$$\frac{1}{(k+l)!} \sum_{i=1}^{k!} \mathrm{sgn}(g_i \tau)\, \omega(w_{g_i(1)}, \ldots, w_{g_i(k)})\, \eta(w_{g_i(k+1)}, \ldots, w_{g_i(k+l)})$$
$$= \frac{1}{(k+l)!} \sum_{i=1}^{k!} \mathrm{sgn}(g_i)\,\mathrm{sgn}(\tau)\, \omega(w_{g_i(1)}, \ldots, w_{g_i(k)})\, \eta(w_{k+1}, \ldots, w_{k+l})$$
$$= \mathrm{sgn}(\tau)\, \eta(w_{k+1}, \ldots, w_{k+l}) \frac{1}{(k+l)!} \sum_{i=1}^{k!} \mathrm{sgn}(g_i)\, \omega(w_{g_i(1)}, \ldots, w_{g_i(k)})$$

$= \mathrm{sgn}(\tau)\, \eta(w_{k+1}, \ldots, w_{k+l}) \frac{k!}{(k+l)!} \mathrm{Alt}(\omega)\,(w_1, \ldots, w_k)$. Now by the supposition $\mathrm{Alt}(\omega) = 0$, so $= \mathrm{sgn}(\tau)\, \eta(w_{k+1}, \ldots, w_{k+l}) \frac{k!}{(k+l)!} \mathrm{Alt}(\omega)\,(w_1, \ldots, w_k) = 0$.

Since we were considering an arbitrary $\tau \in S$, then $\sum_{\sigma \in G\tau} \operatorname{sgn}(\sigma)\, \omega \otimes \eta = 0$ for each $\tau \in S$, hence $\operatorname{Alt}(\omega \otimes \eta) = \frac{1}{(k+l)!} \sum_{\tau \in S} \left(\sum_{\sigma \in G\tau} \operatorname{sgn}(\sigma)\, \omega \otimes \eta \right) = \frac{1}{(k+l)!} \sum_{\tau \in S} (0) = 0$. We conclude that $\operatorname{Alt}(\omega \otimes \eta) = 0$. It is a similar argument to show that if we used the subgroup of permutation elements that fix $1, \ldots, k$ we could show $\operatorname{Alt}(\eta \otimes \omega) = 0$ too! ∎

Theorem 3.10 If $\omega, \eta \in \mathcal{T}^k(\mathbb{R}^n)$, then $\operatorname{Alt}(\omega - \eta) = \operatorname{Alt}(\omega) - \operatorname{Alt}(\eta)$ and $\operatorname{Alt}(\omega + \eta) = \operatorname{Alt}(\omega) + \operatorname{Alt}(\eta)$.

Proof. Let $\omega, \eta \in \mathcal{T}^k(\mathbb{R}^n)$ and $x_1, x_2, \ldots, x_k \in \mathbb{R}^n$. By Definition 3.7, $\operatorname{Alt}(\omega - \eta)(x_1, \ldots, x_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, (\omega - \eta)(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$. By Definition 3.2
$\frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, (\omega - \eta)(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$
$= \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, (\omega(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) - \eta(x_{\sigma(1)}, \ldots, x_{\sigma(k)}))$. We regroup the sum and distribute $\frac{1}{k!}$ and have $\frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, (\omega(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) - \eta(x_{\sigma(1)}, \ldots, x_{\sigma(k)}))$
$= \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \omega(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) - \frac{1}{k!} \sum_{\sigma \in S_k} \operatorname{sgn}(\sigma)\, \eta(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$, which using Definition 3.7 again gives $\operatorname{Alt}(\omega) - \operatorname{Alt}(\eta)$. Similarly for "+". ∎

Theorem 3.11 If $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, $\eta \in \mathcal{T}^l(\mathbb{R}^n)$, and $\theta \in \mathcal{T}^m(\mathbb{R}^n)$, then $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) \otimes \theta) = \operatorname{Alt}(\omega \otimes \eta \otimes \theta)$ and $\operatorname{Alt}(\omega \otimes \operatorname{Alt}(\eta \otimes \theta)) = \operatorname{Alt}(\omega \otimes \eta \otimes \theta)$

Proof. Let $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, $\eta \in \mathcal{T}^l(\mathbb{R}^n)$, and $\theta \in \mathcal{T}^m(\mathbb{R}^n)$. Since $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta)) = \operatorname{Alt}(\omega \otimes \eta)$ by Theorem 3.7, then $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta)) - \operatorname{Alt}(\omega \otimes \eta) = 0$. Since $\operatorname{Alt}(\omega - \eta) = \operatorname{Alt}(\omega) - \operatorname{Alt}(\eta)$ by Theorem 3.10, then $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta)) - \operatorname{Alt}(\omega \otimes \eta) = \operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) - \omega \otimes \eta)$, and since $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta)) - \operatorname{Alt}(\omega \otimes \eta) = 0$, then $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) - \omega \otimes \eta) = 0$. Since $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) - \omega \otimes \eta) = 0$, then by Theorem 3.9, $\operatorname{Alt}((\operatorname{Alt}(\omega \otimes \eta) - \omega \otimes \eta) \otimes \theta) = 0$. Now by the first of the four tensor product properties of Theorem 3.1, we can distribute the tensor product on the right so the last equation is equivalent to $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) \otimes \theta - \omega \otimes \eta \otimes \theta) = 0$. We reapply Theorem 3.10 to have $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) \otimes \theta) - \operatorname{Alt}(\omega \otimes \eta \otimes \theta) = 0$ or $\operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) \otimes \theta) = \operatorname{Alt}(\omega \otimes \eta \otimes \theta)$. We have our first conclusion. The argument for the second conclusion is similar. ∎

We are now ready to show the wedge product is associative.

Theorem 3.12 If $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, $\eta \in \mathcal{T}^l(\mathbb{R}^n)$, and $\theta \in \mathcal{T}^m(\mathbb{R}^n)$, then $(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta)$.

Proof. Let $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, $\eta \in \mathcal{T}^l(\mathbb{R}^n)$, and $\theta \in \mathcal{T}^m(\mathbb{R}^n)$.

$$
\begin{aligned}
(\omega \wedge \eta) \wedge \theta &= \tfrac{(k+l+m)!}{(k+l)!\,m!} \operatorname{Alt}((\omega \wedge \eta) \otimes \theta) && \textit{Def 3.8} \\
&= \tfrac{(k+l+m)!}{(k+l)!\,m!} \operatorname{Alt}\left(\tfrac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta) \otimes \theta\right) && \textit{Def 3.8} \\
&= \tfrac{(k+l+m)!}{(k+l)!\,m!}\, \tfrac{(k+l)!}{k!\,l!} \operatorname{Alt}(\operatorname{Alt}(\omega \otimes \eta) \otimes \theta) && \textit{Def 3.7} \\
&= \tfrac{(k+l+m)!}{k!\,l!\,m!} \operatorname{Alt}(\omega \otimes \eta \otimes \theta) && \textit{Thrm 3.11} \\
&= \tfrac{(k+l+m)!}{k!\,(l+m)!}\, \tfrac{(l+m)!}{l!\,m!} \operatorname{Alt}(\omega \otimes \operatorname{Alt}(\eta \otimes \theta)) && \textit{Thrm 3.11} \\
&= \tfrac{(k+l+m)!}{k!\,(l+m)!} \operatorname{Alt}\left(\omega \otimes \tfrac{(l+m)!}{l!\,m!} \operatorname{Alt}(\eta \otimes \theta)\right) && \textit{Def 3.7} \\
&= \tfrac{(k+l+m)!}{k!\,(l+m)!} \operatorname{Alt}(\omega \otimes (\eta \wedge \theta)) && \textit{Def 3.8} \\
&= \omega \wedge (\eta \wedge \theta) && \textit{Def 3.8}
\end{aligned}
$$

This was the final hurdle on our quest to find a basis for $\Lambda^k(\mathbb{R}^n)$. As with finding a basis for $\mathcal{T}^k(\mathbb{R}^n)$, we will begin with an example. We shall consider $\Lambda^3(\mathbb{R}^5)$ with $\{e_1, e_2, e_3, e_4, e_5\}$ the standard basis for $\mathbb{R}^5$ and $\{\pi^1, \pi^2, \pi^3, \pi^4, \pi^5\}$ the standard basis for $(\mathbb{R}^5)^*$. We hope to use our wedge product to build a set of elements of the form $\pi^i \wedge \pi^j \wedge \pi^k$ with $i, j, k \in \{1, 2, 3, 4, 5\}$ that will span $\Lambda^3(\mathbb{R}^5)$.

First note if $i = j$, $j = k$, or $i = k$, then $\pi^i \wedge \pi^j \wedge \pi^k = 0$. For example, since by Theorem 3.8 part 4 $\pi^1 \wedge \pi^1 \wedge \pi^2 (x_1, x_2, x_3) = (-1)^{1\,1}\, \pi^1 \wedge \pi^1 \wedge \pi^2 (x_1, x_2, x_3)$ (interchanging the 1 in the first position and the 1 in the second position) and since for $a \in \mathbb{R}$ satisfying $a = -a$ implies $a = 0$ then it follows that $\pi^1 \wedge \pi^1 \wedge \pi^2 = 0$. This tell us that any basis element of the form $\pi^i \wedge \pi^j \wedge \pi^k$ with $i, j, k \in \{1, 2, 3, 4, 5\}$ will have $i \neq j$ and $j \neq k$ and $i \neq k$.

Second note $\pi^i \wedge \pi^j \wedge \pi^k = -\pi^j \wedge \pi^i \wedge \pi^k$ by Theorem 3.8 part 4. We see that any set containing $\pi^i \wedge \pi^j \wedge \pi^k$ and $\pi^j \wedge \pi^i \wedge \pi^k$ would not be linearly independent. This is to say that if each of $\pi^{i_1} \wedge \pi^{j_1} \wedge \pi^{k_1}$ and $\pi^{i_2} \wedge \pi^{j_2} \wedge \pi^{k_2}$ is an element in a basis for $\Lambda^3(\mathbb{R}^5)$, then i_1, j_1, k_1 cannot be a permutation of i_2, j_2, k_2.

Now that we know which elements of the form $\pi^i \wedge \pi^j \wedge \pi^k$ with $i, j, k \in \{1, 2, 3, 4, 5\}$ we cannot have in a basis for $\Lambda^3(\mathbb{R}^5)$, we will put together the remaining elements and determine if they form a basis. We notice that according to our criterion, to be an element of the basis, we are choosing i, j, k from $\{1, 2, 3, 4, 5\}$ unordered without repetition, which of course is a combination of 5 elements taken 3 at a time. So the order of such a basis would be 5 choose 3 or $\frac{5!}{3!(5-3)!} = 10$. We claim

$$\{\pi^1 \wedge \pi^2 \wedge \pi^3,\ \pi^1 \wedge \pi^2 \wedge \pi^4,\ \pi^1 \wedge \pi^2 \wedge \pi^5,\ \pi^1 \wedge \pi^3 \wedge \pi^4,\ \pi^1 \wedge \pi^3 \wedge \pi^5,$$
$$\pi^1 \wedge \pi^4 \wedge \pi^5,\ \pi^2 \wedge \pi^3 \wedge \pi^4,\ \pi^2 \wedge \pi^3 \wedge \pi^5,\ \pi^2 \wedge \pi^4 \wedge \pi^5,\ \pi^3 \wedge \pi^4 \wedge \pi^5\}$$

is a basis for $\Lambda^3(\mathbb{R}^5)$. We see the need for some notation so we write $\{\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \mid 1 \le i_1 < i_2 < i_3 \le 5\}$, and we say whenever $\pi^{i_1} \wedge \pi^{i_2} \wedge \cdots \wedge \pi^{i_k}$ satisfies $i_1 < i_2 < \cdots < i_k$, then $\pi^{i_1} \wedge \pi^{i_2} \wedge \cdots \wedge \pi^{i_k}$ is in standard form. We proceed to show $\{\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \mid 1 \le i_1 < i_2 < i_3 \le 5\}$ spans $\Lambda^3(\mathbb{R}^5)$ and is linearly independent. We don't have to start completely from scratch since we know already that $\Lambda^3(\mathbb{R}^5) \subset \mathcal{T}^3(\mathbb{R}^5)$.

Let $\omega \in \Lambda^3(\mathbb{R}^5)$ and $\ _1, x_2, x_3 \in \mathbb{R}^5$. Since $\Lambda^3(\mathbb{R}^5) \subset \mathcal{T}^3(\mathbb{R}^5)$, then $\omega \in \mathcal{T}^3(\mathbb{R}^5)$, and since we have a basis for $\mathcal{T}^3(\mathbb{R}^5)$, then $\omega = \sum_{i_1=1}^{5}\left(\sum_{i_2=1}^{5}\left(\sum_{i_3=1}^{5} b_{i_1,i_2,i_3} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \pi^{i_3}\right)\right)$, which could have as many as 125 terms so we hold on to a mild level of abstraction. Since $\omega \in \Lambda^3(\mathbb{R}^5)$, then by Theorem 3.6 $\omega = \operatorname{Alt}(\omega)$. From these last two conclusions, we have $\omega = \operatorname{Alt}\left(\sum_{i_1=1}^{5}\left(\sum_{i_2=1}^{5}\left(\sum_{i_3=1}^{5} b_{i_1,i_2,i_3} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \pi^{i_3}\right)\right)\right)$. An extensive use of Theorem 3.10, yields the conclusion $\omega = \sum_{i_1=1}^{5}\left(\sum_{i_2=1}^{5}\left(\sum_{i_3=1}^{5} \operatorname{Alt}(b_{i_1,i_2,i_3} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \pi^{i_3})\right)\right)$. We will consider an arbitrary one of the 125 terms. From Definition 3.7 it is easy to see that $\operatorname{Alt}(a \cdot \omega) = a \cdot \operatorname{Alt}(\omega)$ for $a \in \mathbb{R}$ and $\omega \in \mathcal{T}^k(\mathbb{R}^n)$, thus, $\operatorname{Alt}(b_{i_1,i_2,i_3} \cdot \pi^{i_1} \otimes \pi^{i_2} \otimes \pi^{i_3}) = b_{i_1\,i_2\,i_3} \operatorname{Alt}(\pi^{i_1} \otimes \pi^{i_2} \otimes \pi^{i_3})$ for each $i_1, i_2, i_3 \in \{1, 2, 3, 4, 5\}$. Now from Definition 3.8 and Theorem 3.12, for each $i_1, i_2, i_3 \in \{1, 2, 3, 4, 5\}$, $b_{i_1\,i_2\,i_3} \operatorname{Alt}(\pi^{i_1} \otimes \pi^{i_2} \otimes \pi^{i_3}) = b_{i_1\,i_2\,i_3} \frac{1!\,1!\,1!}{(1+1+1)!}(\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3})$, this being a scalar product of our arbitrary term. We have already shown that $b_{i_1 i_2 i_3} \frac{1!1!1!}{1+1+1)!} (\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3}) = 0$ whenever $i_1 = i_2$ or $i_1 = i_3$ or $i_2 = i_3$. So we have ω as the sum of only 60 potentially non-zero terms since we have 5 choices for i_1, then 4 choices for i_2, and only 3 choices left for i_3. For each combination of the three symbols chosen from $\{1, 2, 3, 4, 5\}$, we have 3! or six permutations. For example six of the terms from our 60 potentially non-zero terms are as follows: $b_{123} \frac{1}{3!} \pi^1 \wedge \pi^2 \wedge \pi^3 + b_{132} \frac{1}{3!} \pi^1 \wedge \pi^3 \wedge \pi^2 + b_{213} \frac{1}{3!} \pi^2 \wedge \pi^1 \wedge \pi^3 + b_{231} \frac{1}{3!} \pi^2 \wedge \pi^3 \wedge \pi^1 + b_{312} \frac{1}{3!} \pi^3 \wedge \pi^1 \wedge \pi^2 + b_{321} \frac{1}{3!} \pi^3 \wedge \pi^2 \wedge \pi^1$. We use Theorem 3.8 part 4 to put $\pi^{\sigma(1)} \wedge \pi^{\sigma(2)} \wedge \pi^{\sigma(3)}$ in standard form written as $\operatorname{sgn}(\sigma) \cdot \pi^1 \wedge \pi^2 \wedge \pi^3$ for $\sigma \in S_3$. So we let $a_{i_1 i_2 i_3} = \sum_{\sigma \in S_3} b_{\sigma(i_1)\sigma(i_2)\sigma(i_3)} \frac{1}{3!} \operatorname{sgn}(\sigma)$. We write all of our alternating tensors in standard form, take our 60 terms divide them into 10 groups each of 6 terms of the same alternating tensor $\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3}$ in standard form. We use the distributive property of real numbers and make the substitution of $a_{i_1 i_2 i_3} = \sum_{\sigma \in S_3} b_{\sigma(i_1)\sigma(i_2)\sigma(i_3)} \frac{1}{3!} \operatorname{sgn}(\sigma)$ so we have ten real numbers $a_{i_1 i_2 i_3}$ each scalar product with one of the ten corresponding $\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3}$ in standard form. Thus we have written our sum originally involving a theoretical 125 terms as a sum involving only 10 terms. That is $\omega = \sum_{i_1=1}^{3} \left(\sum_{i_2=i_1+1}^{4} \left(\sum_{i_3=i_2+1}^{5} a_{i_1 i_2 i_3} \cdot \pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \right)\right) = a_{123}\, \pi^1 \wedge \pi^2 \wedge \pi^3 + a_{124}\, \pi^1 \wedge \pi^2 \wedge \pi^4 + a_{125}\, \pi^1 \wedge \pi^2 \wedge \pi^5 + a_{134}\, \pi^1 \wedge \pi^3 \wedge \pi^4 + a_{135}\, \pi^1 \wedge \pi^3 \wedge \pi^5 + a_{145}\, \pi^1 \wedge \pi^4 \wedge \pi^5 + a_{234}\, \pi^2 \wedge \pi^3 \wedge \pi^4 + a_{235}\, \pi^2 \wedge \pi^3 \wedge \pi^5 + a_{245}\, \pi^2 \wedge \pi^4 \wedge \pi^5 + a_{345}\, \pi^3 \wedge \pi^4 \wedge \pi^5$, and consequently we have written ω as a linear combination of elements from $\{\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \mid 1 \le i_1 < i_2 < i_3 \le 5\}$.

It remains to show $\{\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \mid 1 \le i_1 < i_2 < i_3 \le 5\}$ is a linearly independent set. We write the sum of elements from the set $\{\pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \mid 1 \le i_1 < i_2 < i_3 \le 5\}$ in a rather elaborate but well defined manner. Suppose

$$\sum_{i_1=1}^{3} \left(\sum_{i_2=i_1+1}^{4} \left(\sum_{i_3=i_2+1}^{5} a_{i_1 i_2 i_3} \cdot \pi^{i_1} \wedge \pi^{i_2} \wedge \pi^{i_3} \right)\right) = 0.$$

We apply $\sum_{i_1=1}^{3}\left(\sum_{i_2=i_1+1}^{4}\left(\sum_{i_3=i_2+1}^{5} a_{i_1\, i_2\, i_3}\cdot\pi^{i_1}\wedge\pi^{i_2}\wedge\pi^{i_3}\right)\right)$ to $(e_{j_1}, e_{j_2}, e_{j_3})$ where $j_1, j_2, j_3 \in \{1, 2, 3, 4, 5\}$. A typical term in the sum is $a_{i_1\, i_2\, i_3}\cdot\pi^{i_1}\wedge\pi^{i_2}\wedge\pi^{i_3}\,(e_{j_1}, e_{j_2}, e_{j_3})$. From the definition of wedge product, Alt, and tensor product this term is the sum

$$a_{i_1\, i_2\, i_3}\,\frac{(1+1+1)}{1!\,1!\,1!}\,\frac{1}{(1+1+1)!}\sum_{\sigma\in S_3}\operatorname{sgn}(\sigma)\,\pi^{i_1}(e_{\sigma(j_1)})\,\pi^{i_2}(e_{\sigma(j_2)})\,\pi^{i_3}(e_{\sigma(j_3)}).$$

The 1-tensor factor $\pi^{i_1}(e_{\sigma(j_1)}) = 0$ unless $\sigma(j_1) = i_1$, in which case it is 1. The same is true for the other two factors; thus, since there is only one term in the original sum where j_1, j_2, j_3 is a permutation of i_1, i_2, i_3, nine of the terms are zero and for the remaining term expanded as the sum of six elements with the definition of wedge product, Alt, and tensor product as above, five of these terms are zero and the only quantity remaining is $a_{i_1\, i_2\, i_3}$. Since the 3-tensor was defined to map every ordered triplet of vectors from $\mathbb{R}^5$ to the number zero, then it must be the case that $a_{i_1\, i_2\, i_3} = 0$. We repeat this process 10 times for the other possible combinations of $(e_{j_1}, e_{j_2}, e_{j_3})$, and will show each of the coefficients must be zero.

Now we have sufficient understanding to prove the general theorem. We collect the set of elements $\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k}$ with $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$ in standard form, that is, with $i_1, i_2, \ldots, i_k$ satisfying $i_1 < i_2$ and $i_2 < i_3$ and … and $i_{k-1} < i_k$ and write this all as $\{\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k} \mid 1 \le i_1 < i_2 < \cdots < i_k \le n\}$.

Theorem 3.13 Let $\pi^j : \mathbb{R}^n \longrightarrow \mathbb{R}$ be the projection functions of Definition 3.5. A basis for $\Lambda^k(\mathbb{R}^n)$ is $\{\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k} \mid 1 \le i_1 < i_2 < \cdots < i_k \le n\}$.

Proof. We first show $\{\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k} \mid 1 \le i_1 < i_2 < \cdots < i_k \le n\}$ spans and then is linearly independent.

Let $\omega \in \Lambda^k(\mathbb{R}^n)$. From the ideas present in the above example, we see that $\omega = \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n}\cdots\left(\sum_{i_k=1}^{n} b_{i_1\, i_2\, \ldots\, i_k}\cdot\operatorname{Alt}(\pi^{i_1}\otimes\pi^{i_2}\otimes\cdots\otimes\pi^{i_k})\right)\cdots\right)$ with n^k terms in the sum before expounding on the meaning of $\operatorname{Alt}(\cdot)$. Furthermore,
$\omega = \sum_{i_1=1}^{n}\left(\sum_{i_2=1}^{n}\cdots\left(\sum_{i_k=1}^{n} b_{i_1\, i_2\, \ldots\, i_k}\,\frac{1}{k!}\cdot(\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k})\right)\cdots\right)$ at which point we see all but $n!/(n-k)!$ of the terms are zero. Now if we let $a_{i_1\, i_2\, \ldots\, i_k} = \sum_{\sigma\in S_k}\frac{1}{k!}\operatorname{sgn}(\sigma)\, b_{\sigma(1)\,\sigma(2)\,\cdots\,\sigma(k)}$

then we will have $n!/(k!(n-k)!)$ groups each of $k!$ terms and as we see from the patterns of the example above that $\omega = \sum_{i_1=1}^{(n-k)+1}\left(\sum_{i_2=i_1+1}^{(n-k)+2}\cdots\left(\sum_{i_3=i_{k-1}+1}^{n} a_{i_1 i_2 \quad i_k}\cdot\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k}\right)\cdots\right)$ with each $\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k}$ written in standard form. We have an arbitrary element $\omega\in\Lambda^k(\mathbb{R}^n)$ written as a linear combination of elements from $\{\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k} \mid 1\le i_1<i_2<\cdots<i_k\le n\}$, so we conclude that this set spans $\Lambda^k(\mathbb{R}^n)$.

Now we suppose that $\omega = \sum_{i_1=1}^{(n-k)+1}\left(\sum_{i_2=i_1+1}^{(n-k)+2}\cdots\left(\sum_{i_3=i_{k-1}+1}^{n} a_{i_1 i_2 \quad i_k}\cdot\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k}\right)\cdots\right)$ and $\omega(x)=0$ for all $x\in\mathbb{R}^n\times\cdots\mathbb{R}^n$ (k-times). We apply ω to $(e_{j_1}, e_{j_2}, \ldots, e_{j_k})$. This forces $a_{j_1 j_2 \quad j_3}=0$ as explained in the example above. Repeating this for each permutation of $j_1, j_2, \ldots, j_k$ satisfying $j_1<j_2<\cdots<j_k$, we force each $a_{i_1 i_2 i_3}=0$ for each $i_1, i_2, \ldots, i_k\in\{1, 2, \ldots, n\}$ satisfying $i_1<i_2<\cdots<i_k$. This verifies that $\{\pi^{i_1}\wedge\pi^{i_2}\wedge\cdots\wedge\pi^{i_k} \mid 1\le i_1<i_2<\cdots<i_k\le n\}$ is a linearly independent set. ∎

We have at last a basis for $\Lambda^k(\mathbb{R}^n)$. We see in the proof that the number of elements in the basis is $\frac{n!}{k!(n-k)!}=\binom{n}{k}$. Thus there are no alternating tensors on $\mathbb{R}^n$ if $k>n$. If $n=k$ we have a basis containing only one element.

Fields and Forms

In this section we begin by defining a field and then defining a form in terms of a field. The differential k-form is of particular interest in the proof of Stokes' Theorem.

Definition 3.9 The tangent space of $\mathbb{R}^n$ at $p\in\mathbb{R}^n$ is the collection of ordered paris $(p, v)\in\{p\}\times\mathbb{R}^n$ and denoted $\mathbb{R}^n$ with elements denoted v .

Definition 3.10 For $a\in\mathbb{R}$ and $v_p, w_p\in(\mathbb{R}^n)_p$ we define $v_p+w_p=(v+w)_p$ and $a\cdot v_p=(a\cdot v)_p$.

With these definition the claim that $(\mathbb{R}^n)_p$ is a vector space since it is not empty, and it is closed under addition and scalar multiplication. $(\mathbb{R}^n)_p$ has a usual basis $\{(e_1)_p, (e_2)_p, \ldots, (e_n)_p\}$ and many other constructs analogous to $\mathbb{R}^n$. Of particular note is the usual inner product $\langle , \rangle_p$ for $(\mathbb{R}^n)_p$ which is defined by $\langle v_p, w_p \rangle_p = \langle v, w \rangle$. The inner product is an essential construct since it is used in connecting a vector space with its dual as in Theorem 3.3. Without the inner product defined in this way, it is not clear how the dual $((\mathbb{R}^n)_p)^*$ should be defined. With this definition of the usual inner product for $(\mathbb{R}^n)_p$, if we define $\varphi_{x_p} : (\mathbb{R}^n)_p \longrightarrow \mathbb{R}$ for a unique $x_p \in (\mathbb{R}^n)_p$ by $\varphi_{x_p}(y_p) = \langle x_p, y_p \rangle$ for each $y_p \in (\mathbb{R}^n)_p$, then analogous to what we have shown in Theorem 3.3, $(\mathbb{R}^n)_p$ is isomorphic to $((\mathbb{R}^n)_p)^*$. With this clarification it is now clear that the basis for $((\mathbb{R}^n)_p)^*$ is the set of linear functionals $\pi^i(p)$ that project a vector v_p onto basis vector (e_i) for each $i \in \{1, 2, \ldots, n\}$. That is $(\pi^i(p))(v_p) = \langle (e_i)_p, v_p \rangle = \langle e_i, v \rangle = v^i$ for each $i \in \{1, 2, \ldots, n\}$.

In other words, $(\mathbb{R}^n)_p$ is essentially $\mathbb{R}^n$ with origin p, but rather than re-center the origin, Definition 3.9 allows us to overlay pieces of various tangent spaces consisting of only one element and aligned them at their common origin to produce what is called a vector field.

Definition 3.11 A *vector field* is a function F consisting of pairs $(p, F(v))$ for $p \in \mathbb{R}^n$ and $F(v) = v_p \in (\mathbb{R}^n)$.

For example, suppose F is a function defined for $p \in \mathbb{R}^2$ by $F(p) = (2, -3)$. F is a constant field of vectors.

We can also write $(2, -3)_p = 2 \cdot (e_1)_p - 3 \cdot (e_2)$, so we can choose two component functions $F^1 : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $F^2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ so that $F^1(p) = 2$ and $F^2(p) = -3$ for each $p \in \mathbb{R}^n$, and we can define the vector field F for each $p \in \mathbb{R}^n$ by $F(p) = F^1(p) \cdot (e_1)_p + F^2(p) \cdot (e_2)$. We would call this vector field constant since each F^i is constant. In general we describe a vector field $F(p) = \sum_{i=1}^{n} F^i(p) \cdot (e_i)_p$ for each $p \in \mathbb{R}^n$ with component functions $F^i : \mathbb{R}^n \longrightarrow \mathbb{R}$. We classify each in terms of the component functions. If each component function is constant, then F is called constant. To simplify the statement of theorems hereafter, to say a function is differentiable will mean that a function has continuous partial derivatives of all orders. A function with this degree of differentiability is referred to as C^∞.

Before defining a k-form, we should make clear the meaning of $\Lambda^k((\mathbb{R}^n)_p)$. Many authors assume that it is understood what is meant by $\Lambda^k((\mathbb{R}^n)_p)$ through analogy to the constructs of similar appearance and form. We do little more in the sense of rigor, but to appeal to our sense of the meaning, we elaborate on some results.

Definition 3.12 Let $\omega, \vartheta \in \Lambda^k((\mathbb{R}^n)_p)$, $\eta \in \Lambda^l((\mathbb{R}^n)_p)$, $(v_1)_p, (v_2)_p, \ldots, (v_{k+l})_p \in (\mathbb{R}^n)_p$, $(v_i)_p = (a_i + b_i)_p = (a_i)_p + (b_i)_p$ and $a \cdot (v_i)_p = (a \cdot v_i)_p$ for $i \in \{1, 2, \ldots, k\}$ and some $(a_i)_p, (b_i)_p \in (\mathbb{R}^n)_p$ and $a \in \mathbb{R}$, then we define the following:

1. $\omega((v_1)_p, \ldots, (v_i)_p, \ldots, (v_k)_p) = \omega((v_1)_p, \ldots, (a_i)_p, \ldots, (v_k)_p) + \omega((v_1)_p, \ldots, (b_i)_p, \ldots, (v_k)_p)$,
2. $a \cdot \omega((v_1)_p, \ldots, (v_i)_p, \ldots, (v_k)_p) = \omega((v_1)_p, \ldots, a \cdot (v_i)_p, \ldots, (v_k)_p) = \omega((v_1)_p, \ldots, (a \cdot v_i)_p, \ldots, (v_k)_p)$,
3. $(\omega + \vartheta)((v_1)_p, \ldots, (v_i)_p, \ldots, (v_k)_p) = \omega((v_1)_p, \ldots, (v_i)_p, \ldots, (v_k)_p) + \vartheta((v_1)_p, \ldots, (v_i)_p, \ldots, (v_k)_p)$,
4. $(\omega \otimes \eta)((v_1)_p, \ldots, (v_{k+l})_p) = \omega((v_1)_p, \ldots, (v_k)_p)\ \eta((v_{k+1})_p, \ldots, (v_{k+l})_p)$,
5. $\omega\big((v_1)_p, \ldots, (v_i)_p, \ldots, (v_j)_p, \ldots, (v_k)_p\big) = -\omega\big((v_1)_p, \ldots, (v_j)_p, \ldots, (v_i)_p, \ldots, (v_k)_p\big)$ (interchange i and j with $i \neq j$),
6. $\mathrm{Alt}(\omega)((v_1)_p, (v_2)_p, \ldots, (v_k)_p) = \frac{1}{k!} \sum_{\sigma \in S_k} \mathrm{sgn}(\sigma) \cdot \omega\big((v_{\sigma(1)})_p, \ldots, (v_{\sigma(k)})_p\big)$.

With some experience, these abstract definitions, in terms of the tangent space of $\mathbb{R}^n$ at p, all seem very natural. Still it gives us a point of reference as the notation continues to harbor increasingly complex meaning. We did not mention the wedge product since it is defined in terms of $\mathrm{Alt}(\cdot)$ and "$\otimes$", which are now well defined in (4.) and (6.) above. What remains somewhat unclear is what is meant by a function $\omega : ((\mathbb{R}^n)_p)^k \longrightarrow \mathbb{R}$. We clarify this in our suggestion of a basis for $\Lambda^k((\mathbb{R}^n)_p)$. In the same manner that we constructed our usual basis for $\Lambda^k(\mathbb{R}^n)$ using "$\wedge$" and the usual basis for $(\mathbb{R}^n)^*$, we can construct a basis for $\Lambda^k((\mathbb{R}^n)_p)$ using "$\wedge$" and the usual basis for $((\mathbb{R}^n)_p)^*$. We have the $\{\pi^{i_1}(p) \wedge \pi^{i_2}(p) \wedge \cdots \wedge \pi^{i_k}(p) \mid 1 \leq i_1 < i_2 < \cdots < i_k \leq n\}$ as the basis for

$\Lambda^k((\mathbb{R}^n)_p)$, which is reasonable following the pattern of extending our other tensor constructs from $\mathbb{R}^n$ to $(\mathbb{R}^n)$. Now the definition of a k-form follows easily.

Definition 3.14 A *differential form* or *k-form* is a function ω consisting of pairs $(\boldsymbol{p}, \omega(\boldsymbol{p}))$ where $\boldsymbol{p} \in \mathbb{R}^n$ and $\omega(\boldsymbol{p}) \in \Lambda^k((\mathbb{R}^n)_p)$.

In other words, a differential form is a mapping for each point $\boldsymbol{p}$ in a Euclidean vector space $\mathbb{R}^n$ to alternating k-tensors of the tangent space of $\mathbb{R}^n$ at the same point $\boldsymbol{p}$. From Definition 3.14 we se an emerging connection between vector fields and differential forms. With vector fields, we associated to each point in $\mathbb{R}^n$, a vector, now with differential forms, we associated to each point in $\mathbb{R}^n$, an alternating k-tensor. Just as we could write each vector field F as the sum of certain component functions, one for each basis element of $(\mathbb{R}^n)_p$, we can also write each differential form ω as the sum of certain component functions, one for each basis element of $\Lambda^k((\mathbb{R}^n)_p)$. If ω is a differential form on $\mathbb{R}^n$, then for each $\boldsymbol{p} \in \mathbb{R}^n$, there are component functions $\omega_{i_1 i_2 \ldots i_k} : \mathbb{R}^n \longrightarrow \mathbb{R}$, $\{1 \le i_1 < i_2 < \cdots < i_k < n\}$ such that

$$\omega(\boldsymbol{p}) = \sum_{i_1=1}^{(n-k)+1} \left(\sum_{i_2=i_1+1}^{(n-k)+2} \cdots \left(\sum_{i_k=i_{k-1}+1}^{n} \omega_{i_1 i_2 \ldots i_k}(\boldsymbol{p}) \cdot \pi^{i_1}(\boldsymbol{p}) \wedge \pi^{i_2}(\boldsymbol{p}) \wedge \cdots \wedge \pi^{i_k}(\boldsymbol{p}) \right) \cdots \right).$$

Just as with F, we have the same considerations in describing ω as continuous, differentiable, et cetera depending on the component functions $\omega_{i_1 i_2 \ldots i_3}$.

Definition 3.15 If ω and ϑ are k-forms on $\mathbb{R}^n$, η is an l-form on $\mathbb{R}^n$, and $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, then for each $\boldsymbol{p} \in \mathbb{R}^n$,

1. $\omega + \vartheta$ is defined by $(\omega + \vartheta)(\boldsymbol{p}) = \omega(\boldsymbol{p}) + \vartheta(\boldsymbol{p})$,
2. $f \cdot \omega$ is defined by $(f \cdot \omega)(\boldsymbol{p}) = f(\boldsymbol{p}) \cdot \omega(\boldsymbol{p})$, and
3. $\omega \wedge \eta$ is defined by $(\omega \wedge \eta)(\boldsymbol{p}) = \omega(\boldsymbol{p}) \wedge \eta(\boldsymbol{p})$.

Now that we have made the definitions, we consider the consequences.

Suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable, then for each $\boldsymbol{p} \in \mathbb{R}^n$, $Df(\boldsymbol{p}) \in \Lambda^1(\mathbb{R}^n)$. Notice $Df(\boldsymbol{p}) : \mathbb{R}^n \longrightarrow \mathbb{R}$ is linear by Definition 2.1, and the alternating criterion is automatic since $k = 1$, that is there are not two positions to interchange. Of course since the basis for $\Lambda^1(\mathbb{R}^n)$ is the basis for the dual of $\mathbb{R}^n$ which we are writing as

$\{\pi^1, \pi^2, \ldots, \pi^n\}$, then $Df(p) = \sum_{i=1}^{n} a_i \pi^i$ for some $a_i \in \mathbb{R}$ and we say $Df(p)$ is an alternating 1-tensor. For a differentiable function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, we define $d f$ for each $p \in \mathbb{R}^n$ by $d f(p) = Df(p)$, and furthermore, if $v_p \in (\mathbb{R}^n)$, we define $d f(p)(v_p) = Df(p)(v)$, thus, $d f$ is a 1-form as it takes vectors of $\mathbb{R}^n$ to alternating tensors of $\Lambda^1((\mathbb{R}^n)_p)$ that in turn take elements of $(\mathbb{R}^n)_p$ to numbers. This has immediate consequence for our notation. Since for each $i \in \{1, 2, \ldots, n\}$, π^i is a function from $\mathbb{R}^n$ to $\mathbb{R}$, then $d\pi^i(p)(v_p) = D\pi^i(p)(v)$. For each $i \in \{1, 2, \ldots, n\}$, $D\pi^i$ is the same at every point in $\mathbb{R}^n$ and has the Jacobian matrix $1 \times n$ with zeros for all entries except the ith entry is 1. Therefore for each $i \in \{1, 2, \ldots, n\}$, $D\pi^i(p)(v) = \langle e_i, v\rangle = v^i = \pi^i(v)$, in particular $d\pi^i(p) = \pi^i(p)$ for each $p \in (\mathbb{R}^n)$. If in the classical tradition we rename the function π^i as x^i, then $\{dx^1(p),\ dx^2(p), \ldots, dx^n(p)\}$ is just the usual basis for $((\mathbb{R}^n)_p)^*$, the dual of $(\mathbb{R}^n)$. Furthermore, we now write each k-form ω as

$$\omega = \sum_{i_1=1}^{(n-k)+1} \left(\sum_{i_2=i_1+1}^{(n-k)+2} \cdots \left(\sum_{i_k=i_{k-1}+1}^{n} \omega_{i_1 i_2 \quad i_k} \cdot dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} \right) \cdots \right).$$

The following theorem summarizes the notation and is used often in what follows.

Theorem 3.14 If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable, then $d f = D_1 f \cdot dx^1 + D_2 f \cdot dx^2 + \cdots + D_n f \cdot dx^n$.

Proof. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ be differentiable and $p, v \in \mathbb{R}^n$. We have defined $d f(p)(v\ \ ) = Df(p)(v)$ and since f is differentiable (C^∞) and we have the usual basis for $\mathbb{R}^n$ and $\mathbb{R}$, then $Df(p) = (\, D_1 f(p) \ \cdots \ D_n f(p)\,)$. We restate $D f(p)(v)$ as $(\, D_1 f(p), \ \ldots, \ D_n f(p)\,) \cdot v = \sum_{i=1}^{n} D_i f(p) v^i$. Since we have shown in previous discussion that $dx^i(p)(v_p) = v^i$, then $\sum_{i=1}^{n} D_i f(p) v^i = \sum_{i=1}^{n} D_i f(p) dx^i(p)(v_p)$. Since p and v were arbitrary then we have $d f = \sum_{i=1}^{n} D_i f \cdot dx^i$. ∎

We back up for a moment to the simple setting of k-tensors to introduce a new notation that will carry our current notationally complex constructs to the next level.

Definition 3.16 For a linear transformation $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $T \in \mathcal{T}^k(\mathbb{R}^m)$, we define $f^* : \mathcal{T}^k(\mathbb{R}^n) \longrightarrow \mathcal{T}^k(\mathbb{R}^m)$ by $f^*\, T(v_1, v_2, \ldots, v_k) = T(f(v_1), f(v_2), \ldots, f(v_k))$.

This is perfectly reasonable and meaningful, but to further convince ourselves, we show the following.

Theorem 3.15 If $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, $S \in \mathcal{T}^k(\mathbb{R}^m)$ and $T \in \mathcal{T}^l(\mathbb{R}^m)$, then $f^*(S \otimes T) = f^* S \otimes f^* T$.

Proof. Let $v_1, \ldots, v_{k+l} \in \mathbb{R}^n$. $f^*(S \otimes T)(v_1, \ldots, v_{k+l}) = (S \otimes T)(f(v_1), \ldots, f(v_{k+l}))$ by Definition 3.16. Then $(S \otimes T)(f(v_1), \ldots, f(v_{k+l})) =$ $S(f(v_1), \ldots, f(v_k))\ T(f(v_{k+1}), \ldots, f(v_{k+l}))$ by Definition 3.3. Finally $S(f(v_1), \ldots, f(v_k))\ T(f(v_{k+1}), \ldots, f(v_{k+l})) = f^* S(v_1, \ldots, v_k)\, f^* T(v_{k+1}, \ldots, v_{k+l})$ by Definition 3.16, and $f^* S(v_1, \ldots, v_k)\, f^* T(v_{k+1}, \ldots, v_{k+l}) = f^* S \otimes f^* T(v_1, \ldots, v_{k+l})$ by Definition 3.3. ∎

Theorem 3.16 If $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, $\omega \in \Lambda^k(\mathbb{R}^m)$ and $\eta \in \Lambda^l(\mathbb{R}^m)$, then $f^*(\omega \wedge \eta) = f^* \omega \wedge f^* \eta$.

Proof. Let $v_1, v_2, \ldots, v_{k+l} \in \mathbb{R}^n$. $f^*(\omega \wedge \eta)(v_1, \ldots, v_{k+l}) =$ $(\omega \wedge \eta)(f(v_1), \ldots, f(v_{k+l}))$ by Definition 3.16. Next $(\omega \wedge \eta)(f(v_1), \ldots, f(v_{k+l})) =$ $\frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta)(f(v_1), \ldots, f(v_{k+l}))$ by Definition 3.8. Then $\frac{(k+l)!}{k!\,l!} \operatorname{Alt}(\omega \otimes \eta)(f(v_1), \ldots, f(v_{k+l})) = \frac{(k+l)!}{k!\,l!} \operatorname{Alt}(f^* \omega \otimes f^* \eta)(v_1, \ldots, v_{k+l})$ by Definition 3.7 and Theorem 3.15. Pulling back out with Definition 3.8 $\frac{(k+l)!}{k!\,l!} \operatorname{Alt}(f^* \omega \otimes f^* \eta)(v_1, \ldots, v_{k+l}) = f^* \omega \wedge f^* \eta((v_1, \ldots, v_{k+l}))$. ∎

Recall for a differentiable function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $Df(p): \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation by Theorem 2.1.

Definition 3.17 For a differentiable function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$, we define $f_*: (\mathbb{R}^n)\ \longrightarrow (\mathbb{R}^m)_{f(\ )}$ by $f_*(v_p) = (Df(p)(v))_{f(p)}$ for $v \in \mathbb{R}^n$.

We need to reiterate Definition 3.16 for $\Lambda^k((\mathbb{R}^n)_p)$ and we make it a new definition so we can refer to it easily in the context of k-forms without pondering its extension from k-tensors as is done in Definition 3.16.

Definition 3.18 For a differentiable function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and a k-form ω on $\mathbb{R}^m$ and $p \in \mathbb{R}^n$, we define $f^* : \Lambda^k((\mathbb{R}^m)_{f(\ )}) \longrightarrow \Lambda^k((\mathbb{R}^n)_p)$ by $(f^* \omega)(p)$ for $(v_1)_p, \ldots, (v_k)_p \in (\mathbb{R}^n)$ by

$(f^* \omega)(p)((v_1)_p, \ldots, (v_k)_p) = \omega(f(p))(f_*((v_1)_p), \ldots, f_*((v_k)_p))$.

Theorem 3.17 Suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable, $\omega, \omega_1, \omega_2$ are k-forms on $\mathbb{R}^m$, η is an l-form on $\mathbb{R}^m$, and $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ then:

1. $f^*(dx^i) = \sum_{j=1}^{n} D_j f^i \cdot dx^j$,
2. $f^*(\omega_1 + \omega_2) = f^*(\omega_1) + f^*(\omega_2)$,
3. $f^*(g \cdot \omega) = (g \circ f) \cdot f^* \omega$, and
4. $f^*(\omega \wedge \eta) = f^* \omega \wedge f^* \eta$.

Proof. Let $p \in \mathbb{R}^n$ and $v_p, (v_1)_p, (v_2)_p, \ldots, (v_{k+l})_p \in (\mathbb{R}^n)_p$.

1. Let $i \in \{1, 2, \ldots, m\}$. Since $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable and dx^i is a 1-form on $\mathbb{R}^m$, then by Definition 3.18, $f^* dx^i(p)(v_p) = dx^i(f(p))(f_*(v_p))$. Considering the supposition of f, by Definition 3.17, $f_*(v_p) = (Df(p)(v))_{f(p)}$ and by Theorem 2.12,

$$Df(p)(v) = \begin{pmatrix} D_1 f^1(p) & D_2 f^1(p) & \cdots & D_n f^1(p) \\ D_1 f^2(p) & D_2 f^2(p) & \cdots & D_n f^2(p) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f^m(p) & D_2 f^m(p) & \cdots & D_n f^m(p) \end{pmatrix} \cdot \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}.$$

Computing the dot product and placing that vector at $f(p)$ yields

$f_*(v_p) = \left(\sum_{j=1}^{n} D_j f^1(p) v^j, \sum_{j=1}^{n} D_j f^2(p) v^j, \ldots, \sum_{j=1}^{n} D_j f^m(p) v^j \right)_{f(p)}$. Since $dx^i f(p)$ is by definition the projection function that for $w_{f(p)} \in (\mathbb{R}^m)_{f(p)}$ gives its projection onto $(e_i)_{f(p)}$, then given that $f_*(v_p) \in (\mathbb{R}^m)_{f(p)}$ as given above,

$dx^i(f(p))(f_*(v_p)) = \sum_{j=1}^{n} D_j f^i(p) v^j$, the ith entry in $f_*(v_p)$. Furthermore, for $j \in \{1, 2, \ldots, n\}$, $dx^j(p)(v\ ) = v^j$, then we make this substitution for v^j and arrive at our result: $dx^i(f(p))(f_*(v_p)) = \sum_{j=1}^{n} D_j f^i(p)\, dx^j(p)(v_p)$.

2. Since $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable and $\omega_1 + \omega_2$ is a k-form on $\mathbb{R}^m$, then by Definition 3.18, $f^*(\omega_1 + \omega_2)(p)((v_1)\ , \ldots, (v_k)\ )$

$= (\omega_1 + \omega_2)\,(f(p))\,(f_*((v_1)_p), \ldots, f_*((v_k)_p))$. In general,the right hand side is a rather involved sum of $\binom{m}{k}$ terms, but for an arbitrary term we have $(\omega_1 + \omega_2)_{\iota_1 \ \ \iota_k}\,(f(p)) \cdot dx^{\iota_1}(f(p)) \wedge \cdots \wedge dx^{\iota_k}\,(f(p))$ for component function $(\omega_1 + \omega_2)_{\iota_1 \ \ \iota_k} : \mathbb{R}^m \longrightarrow \mathbb{R}$. Now $(\omega_1 + \omega_2)_{\iota_1 \,.\, \iota_k}\,(f(p))$ can be evaluated point-wise for component functions $(\omega_1)_{\iota_1 \ \ \iota_k}, (\omega_2)_{\iota_1 \ \ \iota_k} : \mathbb{R}^m \longrightarrow \mathbb{R}$ which exist since each of ω_1 and ω_2 is a k-form on $\mathbb{R}^m$. Therefore $(\omega_1 + \omega_2)_{\iota_1 \ \ \iota_k} \cdot dx^{\iota_1} \wedge \cdots \wedge dx^{\iota_k} =$ $((\omega_1)_{\iota_1 \ \ \iota_k} + (\omega_2)_{\iota_1 \ \ \iota_k}) \cdot dx^{\iota_1} \wedge \cdots \wedge dx^{\iota_k}$. Now we take the right hand side and apply it to a point $f(p) \in \mathbb{R}^m$ which gives us a k-tensor in $\Lambda^k\,((\mathbb{R}^m)_{f(p)})$, which we can apply to an element $\tau \in ((\mathbb{R}^m)_{f(p)})^k$ and get a real number. Observe $((\omega_1)_{\iota_1 \,.\, \iota_k}\, f(p) + (\omega_2)_{\iota_1 \ .\iota_k}\, f(p)) \cdot dx^{\iota_1}\, f(p) \wedge \cdots \wedge dx^{\iota_k}\, f(p) \in \Lambda^k\,((\mathbb{R}^m)_{f(p)})$ and since $(dx^{\iota_1}\, f(p) \wedge \cdots \wedge dx^{\iota_k}\, f(p))\,(\tau) \in \mathbb{R}$, then $(dx^{\iota_1}\, f(p) \wedge \cdots \wedge dx^{\iota_k}\, f(p))\,(\tau)$ distributes over the sum $((\omega_1)_{\iota_1 \ \ \iota_k}\, f(p) + (\omega_2)_{\iota_1 \ \ \iota_k}\, f(p))$ and we have shown the arbitrary term $(\omega_1 + \omega_2)_{\iota_1 \ .\ \iota_k} \cdot dx^{\iota_1} \wedge \cdots \wedge dx^{\iota_k} = (\omega_1)_{\iota_1 \ .\ \iota_k} \cdot dx^{\iota_1} \wedge \cdots \wedge dx^{\iota_k} +$ $(\omega_2)_{\iota_1 \ .\ \iota_k} \cdot dx^{\iota_1} \wedge \cdots \wedge dx^{\iota_k}$. Thus our rather involved sum of $\binom{m}{k}$ terms can be expanded as the sum of two sums of $\binom{m}{k}$ terms, and grouped by component functions for ω_1 and ω_2. We have supplied as basis for the identity $(\omega_1 + \omega_2)\,(f(p))\,(f_*((v_1)_p), \ldots, f_*((v_k)_p)) =$ $(\omega_1)\,(f(p))\,(f_*((v_1)_p), \ldots, f_*((v_k)_p)) + (\omega_2)\,(f(p))\,(f_*((v_1)_p), \ldots, f_*((v_k)_p))$. Applying Definition 3.18 in the opposite manner as at the start and in consideration of the point-wise definition for the sum of k-forms, then the right hand side of the previous equation becomes $(f^*(\omega_1) + f^*(\omega_2))\,(p)\,((v_1)_p, \ldots, (v_k)_p)$.

3. Since $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable and $g \cdot \omega$ is a k-form on $\mathbb{R}^m$, then by Definition 3.18, $f^*(g \cdot \omega)\,(p)\,((v_1)_p, \ldots, (v_k)_p) = (g \cdot \omega)\,(f(p))\,(f_*((v_1)_p), \ldots, f_*((v_k)_p))$. Since $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ and ω is a k-form on $\mathbb{R}^m$, then by Definition 3.15 Part 2, $(g \cdot \omega)\,(f(p)) = g(f(p)) \cdot \omega(f(p))$, thus, $(g \cdot \omega)\,(f(p))\,(f_*((v_1)_p), \ldots, f_*((v_k)_p)) =$ $(g \circ f)\,(p) \cdot f^*\,\omega(f(p))\,((v_1)_p, \ldots, (v_k)_p)$.

4. Since $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable and each of ω and η is a k-form and l-form,

respectively on $\mathbb{R}^m$, then by Definition 3.18, $f^*(\omega \wedge \eta)(\boldsymbol{p})((v_1)_p, \ldots, (v_{k+l})_p) =$ $(\omega \wedge \eta)(f(\boldsymbol{p}))(f_*((v_1)\ ), \ldots, f_*((v_{k+l})\ ))$. Since each of ω and η is a k-form and l-form, respectively on $\mathbb{R}^m$ then by Definition 3.15 Part 3, $(\omega \wedge \eta)(f(\boldsymbol{p})) = \omega(f(\boldsymbol{p})) \wedge \eta(f(\boldsymbol{p}))$. From the previous identity, we have $(\omega \wedge \eta)(f(\boldsymbol{p}))(f_*((v_1)_p), \ldots, f_*((v_{k+l})_p)) =$ $(\omega(f(\boldsymbol{p})) \wedge \eta(f(\boldsymbol{p})))(f_*((v_1)\ ), \ldots, f_*((v_{k+l})\ ))$. The right had side of the last equality can be written as $\frac{(k+l)!}{k!\,l!} \cdot \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \operatorname{sgn}(\sigma)(\omega(f(\boldsymbol{p})) \otimes \eta(f(\boldsymbol{p})))(f_*((v_1)_p), \ldots, f_*((v_{k+l})_p))$ using Definitions 3.8 and 3.12 Part 6. Each term of the sum becomes $\operatorname{sgn}(\sigma)\,\omega(f(\boldsymbol{p})(f_*((v_1)_p), \ldots, f_*((v_k)_p)))\,\eta(f(\boldsymbol{p}))(f_*((v_{k+1})_p), \ldots, f_*((v_{k+l})_p))$ by Definition 3.12 Part 4. Reapplying Definition 3.18, each term becomes $\operatorname{sgn}(\sigma)\, f^*\,\omega(\boldsymbol{p})((v_1)_p, \ldots, (v_k)_p)\, f^*\,\eta(\boldsymbol{p})((v_{k+1})_p, \ldots, (v_{k+l})_p)$. We retrace or steps: by Definition 3.12 Part 4, each term is $\operatorname{sgn}(\sigma)(f^*\,\omega \otimes f^*\,\eta)(\boldsymbol{p})((v_1)_p, \ldots, (v_{k+l})_p)$, by Definition 3.12 Part 6, the sum becomes $\frac{(k+l)!}{k!\,l!} \operatorname{Alt}(f^*\,\omega \otimes f^*\,\eta)((v_1)_p, \ldots, (v_{k+l})_p)$, which by Definition 3.8 is $(f^*\,\omega \wedge f^*\,\eta)((v_1)_p, \ldots, (v_{k+l})_p)$. ∎

Recall that for a differentiable function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, we defined $d f$ for each $\boldsymbol{p} \in \mathbb{R}^n$ by $d f(\boldsymbol{p}) = Df(\boldsymbol{p})$. A k-form is differentiable if each of the component functions is differentiable. We extend the idea of d to differentiable k-forms.

Definition 3.19 For a differentiable k-form ω the *differential* of ω, $d\omega$ is defined by $\sum_{i_1=1}^{(n-k)+1}\left(\sum_{i_2=i_1+1}^{(n-k)+2} \cdots \left(\sum_{i_k=i_{k-1}+1}^{n} d\omega_{i_1 i_2 \ \ i_k} \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k}\right) \cdots\right)$.

Since for $1 \le i_1 < i_2 < \cdots < i_k \le n$ each $\omega_{i_1 i_2 \ \ i_k} : \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable, then by Theorem 3.14, $d\omega_{i_1 i_2 \ \ i_k} = \sum_{\alpha=1}^{n} D_\alpha\, \omega_{i_1 i_2 \ \ i_k} \cdot dx^\alpha$ We write out the terms in an example to make the differential operator a bit less formidable. In the example, let $n = 4$ and $k = 3$, then we have $\binom{4}{3} = 4$ terms in the expansion of ω:

$$d\omega = d\omega_{123} \cdot dx^1 \wedge dx^2 \wedge dx^3 + d\omega_{124} \cdot dx^1 \wedge dx^2 \wedge dx^4 + d\omega_{134} \cdot dx^1 \wedge dx^3 \wedge dx^4$$
$$+ d\omega_{234} \cdot dx^2 \wedge dx^3 \wedge dx^4.$$

Next we use Theorem 3.14 and expand each $d\omega_{i_1 i_2 i_3}$.

$$d\omega =$$
$$(D_1\,\omega_{123}\cdot dx^1 + D_2\,\omega_{123}\cdot dx^2 + D_3\,\omega_{123}\cdot dx^3 + D_4\,\omega_{123}\cdot dx^4)\wedge dx^1\wedge dx^2\wedge dx^3 +$$
$$(D_1\,\omega_{124}\cdot dx^1 + D_2\,\omega_{124}\cdot dx^2 + D_3\,\omega_{124}\cdot dx^3 + D_4\,\omega_{124}\cdot dx^4)\wedge dx^1\wedge dx^2\wedge dx^4 +$$
$$(D_1\,\omega_{134}\cdot dx^1 + D_2\,\omega_{134}\cdot dx^2 + D_3\,\omega_{134}\cdot dx^3 + D_4\,\omega_{134}\cdot dx^4)\wedge dx^1\wedge dx^3\wedge dx^4 +$$
$$(D_1\,\omega_{234}\cdot dx^1 + D_2\,\omega_{234}\cdot dx^2 + D_3\,\omega_{234}\cdot dx^3 + D_4\,\omega_{234}\cdot dx^4)\wedge dx^2\wedge dx^3\wedge dx^4$$

We can use Theorem 3.8 Part 1 to distribute the wedge on the right and then Theorem 3.8 Part 4 to make zero any term with a repeated wedge product, for example

$$dx^1\wedge dx^1\wedge dx^2\wedge dx^3 = 0.$$

$$\begin{aligned} d\omega = {} & D_4\,\omega_{123}\cdot dx^4\wedge dx^1\wedge dx^2\wedge dx^3 + \\ & D_3\,\omega_{124}\cdot dx^3\wedge dx^1\wedge dx^2\wedge dx^4 + \\ & D_2\,\omega_{134}\cdot dx^2\wedge dx^1\wedge dx^3\wedge dx^4 + \\ & D_1\,\omega_{234}\cdot dx^1\wedge dx^2\wedge dx^3\wedge dx^4 \end{aligned}$$

We finally rearrange each wedge product using Theorem 3.8 Part 4 and collect like terms; in this case there is only one distinct basis element remaining.

$$d\omega = (-D_4\,\omega_{123} + D_3\,\omega_{124} - D_2\,\omega_{134} + D_1\,\omega_{234})\cdot dx^1\wedge dx^2\wedge dx^3\wedge dx^4$$

The example shows that the result of applying d to a 3-form was a 4-form. That is the idea! The differential operator makes a k-form into a $(k+1)$-form.

Theorem 3.18 If each of ω and ϑ is a k-form on $\mathbb{R}^n$, then $d(\omega+\vartheta) = d\omega + d\vartheta$.

Proof. Let each of ω and ϑ be a k-form on $\mathbb{R}^n$. By Definition 3.19, $d(\omega+\vartheta)$ gives us a sum of terms of the form $d((\omega+\vartheta)_{\iota_1\,\iota_2\,\ldots\,\iota_k})\wedge dx^{\iota_1}\wedge dx^{\iota_2}\wedge\cdots\wedge dx^{\iota_k}$. Now $d((\omega+\vartheta)_{\iota_1\,\iota_2\,\ldots\,\iota_k}) = d(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k} + \vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k})$ since $(\omega+\vartheta)_{\iota_1\,\iota_2\,\ldots\,\iota_k} = \omega_{\iota_1\,\iota_2\,\ldots\,\iota_k} + \vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k}$ for certain functions $\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k}$ and $\vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k}$ that exists since each of ω and ϑ is a k-form on $\mathbb{R}^n$. By Theorem 3.14, $d(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k} + \vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k}) = \sum_{\alpha=1}^{n} D_\alpha(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k} + \vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k})\cdot dx^\alpha$ and by Theorem 2.7, $D_\alpha(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k} + \vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k}) = D_\alpha(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k}) + D_\alpha(\vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k})$ for each α; thus, $d(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k} + \vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k}) = \sum_{\alpha=1}^{n}[D_\alpha(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k}) + D_\alpha(\vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k})]\cdot dx^\alpha$. By repeated use of Theorem 3.8 Part 1, $\left(\sum_{\alpha=1}^{n}[D_\alpha(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k}) + D_\alpha(\vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k})]\cdot dx^\alpha\right)\wedge dx^{\iota_1}\wedge dx^{\iota_2}\wedge\cdots\wedge dx^{\iota_k} =$ $\sum_{\alpha=1}^{n}[D_\alpha(\omega_{\iota_1\,\iota_2\,\ldots\,\iota_k}) + D_\alpha(\vartheta_{\iota_1\,\iota_2\,\ldots\,\iota_k})]\cdot dx^\alpha\wedge dx^{\iota_1}\wedge dx^{\iota_2}\wedge\cdots\wedge dx^{\iota_k}$, and by the distributive property of real numbers,

$\sum_{\alpha=1}^{n} [D_\alpha(\omega_{\imath_1 \imath_2 \ \ \imath_k}) + D_\alpha(\vartheta_{\imath_1 \imath_2 \ \ \imath_k})] \cdot dx^\alpha \wedge dx^{\imath_1} \wedge dx^{\imath_2} \wedge \cdots \wedge dx^{\imath_k} =$

$\sum_{\alpha=1}^{n} D_\alpha(\omega_{\imath_1 \imath_2 \ldots \imath_k}) \cdot dx^\alpha \wedge dx^{\imath_1} \wedge dx^{\imath_2} \wedge \cdots \wedge dx^{\imath_k} +$

$\sum_{\alpha=1}^{n} D_\alpha(\vartheta_{\imath_1 \imath_2 \ \ \imath_k}) \cdot dx^\alpha \wedge dx^{\imath_1} \wedge dx^{\imath_2} \wedge \cdots \wedge dx^{\imath_k}$. Replacing each term in the original sum by the right-hand-side of the last equation, then arranging the sum and forming two groups—terms involving $\omega_{\imath_1 \imath_2 \ . \ \imath_k}$ and terms involving $\vartheta_{\imath_1 \imath_2 \ \ \imath_k}$—we have our result. ∎

Theorem 3.19 If ω is a k-form on $\mathbb{R}^n$ and η is an l-form on $\mathbb{R}^n$, then $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \, \omega \wedge d\eta$.

Proof. This result follows from the rule for the derivative of a product (Theorem 2.8), the asymmetric commutativity of wedge products (Theorem 3.8), and the previous theorem (Theorem 3.18). ∎

An explicit proof occupies a great deal of space, and it is an exercise in following definitions and manipulating symbols; thus, in this rare case we refrained from giving such a proof.

Theorem 3.20 If ω is a differentiable k-form on $\mathbb{R}^n$, then $d(d(\omega)) = 0$.

Proof. Let ω be a differentiable k-form on $\mathbb{R}^n$. We take each term $d(\omega_{\imath_1 \imath_2 \ .\imath_k}) \wedge dx^{\imath_1} \wedge dx^{\imath_2} \wedge \cdots \wedge dx^{\imath_k}$ from Definition 3.19 , and by Theorem 3.14 $d(\omega_{\imath_1 \imath_2 \ldots \imath_k}) = \left[\sum_{\alpha=1}^{n} D_\alpha(\omega_{\imath_1 \imath_2 \ . \ \imath_k}) \cdot dx^\alpha\right]$. We apply the definition and theorem again so $d(d(\omega_{\imath_1 \imath_2 \ \ \imath_k})) = \left[\sum_{\beta=1}^{n} \sum_{\alpha=1}^{n} D_\beta(D_\alpha(\omega_{\imath_1 \imath_2 \ \ \imath_k})) \cdot dx^\beta \wedge dx^\alpha\right]$. Since the limits of this double sum are the same, then for each α and β, there is a term $D_\beta(D_\alpha(\omega_{\imath_1 \imath_2 \ \ \imath_k})) \cdot dx^\beta \wedge dx^\alpha$ and a term $D_\alpha(D_\beta(\omega_{\imath_1 \imath_2 \ldots \imath_k})) \cdot dx^\alpha \wedge dx^\beta$. Now since $\omega_{\imath_1 \imath_2 \ \ \imath_k}$ is differentiable (C^∞), then by Theorem 2.2, for each α and β, each of $D_\beta(D_\alpha(\omega_{\imath_1 \imath_2 \ \ \imath_k}))$ and $\ _\alpha(D_\beta(\omega_{\imath_1 \imath_2 \ \ \imath_k}))$ is continuous; thus, by Theorem 2.10 $D_\beta(D_\alpha(\omega_{\imath_1 \imath_2 \ \ \imath_k})) = D_\alpha(D_\beta(\omega_{\imath_1 \imath_2 \ \ \imath_k}))$. Since for each α and β $dx^\beta \wedge dx^\alpha = -dx^\alpha \wedge dx^\beta$ by Theorem 3.8 Part 4, then $D_\alpha(D_\beta(\omega_{\imath_1 \imath_2 \ \ \imath_k})) \cdot dx^\alpha \wedge dx^\beta = -D_\beta(D_\alpha(\omega_{\imath_1 \imath_2 \ . \ \imath_k})) \cdot dx^\beta \wedge dx^\alpha$. Finally then in the expansion of $d(d(\omega_{\imath_1 \imath_2 \ \ \imath_k}))$, for each term $\ _\beta(D_\alpha(\omega_{\imath_1 \imath_2 \ . \ \imath_k})) \cdot dx^\beta \wedge dx^\alpha$ there is, by the last conclusion, its additive inverse, the term $D_\alpha(D_\beta(\omega_{\imath_1 \imath_2 \ldots \imath_k})) \cdot dx^\alpha \wedge dx^\beta$; thus,

$d(d(\omega_{i_1 i_2 \ldots i_k})) = 0$. Therefore, for each term in the expansion of $d(d(\omega))$, we have $d(d(\omega_{i_1 i_2 \ldots i_k})) = 0$, making the whole term zero, and we conclude $d(d(\omega)) = 0$. ∎

Theorem 3.21 If ω is a differentiable k-form on $\mathbb{R}^n$ and $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable, then $f^*(d\omega) = d(f^*\omega)$.

Definition 3.20 A k-form ω is called *closed* if $d\omega = 0$.

Definition 3.21 A differentiable k-form ω is called *exact* if $\omega = d\eta$ for some form η.

Theorem 3.22 If ω is an exact k-form on $\mathbb{R}^n$, then ω is closed.

Proof. Let ω be an exact k-form on $\mathbb{R}^n$. Since ω is exact, then by Definition 3.21, then $\omega = d\eta$, for some form η. Applying d we have $d\omega = d(d\eta)$ and by Theorem 3.20, $d(d\eta) = 0$; thus, $d\omega = 0$, which is to say that ω is closed by Definition 3.20. ∎

CHAPTER IV

INTEGRATION ON CHAINS

In this chapter we press forward most expeditiously to the final result of this thesis, Stokes' Theorem. We make a number of definitions to make precise what is meant by integrating forms over chains.

n-chains

Definition 4.1 A *singular n-cube* is a continuous function $c : [0, 1]^n \longrightarrow A$, for $A \subset \mathbb{R}^m$.

We are familiar with some singular 1-cubes. Examples of these are number functions such as $f : [0, 1] \longrightarrow [0, 2]$ where $f = \{(x, 2x) \mid 0 \leq x \leq 1\}$. The standard n-cube, which will be most important for the purposes of this paper, is the inclusion mapping $I^n : [0, 1]^n \longrightarrow \mathbb{R}^n$, that is $I^n(\boldsymbol{x}) = \boldsymbol{x}$ for each $\in \mathbb{R}^n$. We note that in Definition 4.1 since $[0, 1]^k$ is compact and c is continuous then $c([0, 1]^k)$ is compact, however, n is not necessarily m as is the case for the standard n-cube.

Definition 4.2 For each $i \in \{1, 2, \ldots, n\}$ and $\alpha \in \{1, 2\}$ we define the *(i, α)-face* of the standard n-cube I^n as the $(n - 1)$-cube $I^n_{(i,\alpha)} : [0, 1]^{n-1} \longrightarrow \mathbb{R}^n$ defined for each $\boldsymbol{x} \in [0, 1]^{n-1}$ by $I^n_{(i,\alpha)}(\boldsymbol{x}) = I^n(x^1, \ldots, x^{i-1}, \alpha, x^i, \ldots, x^{n-1})$.

Note since there are two faces for each $i \in \{1, 2, \ldots, n\}$, then for the standard n-cube, there are $2n$ different faces. To understand this important concept we investigate the faces of I^2. The four faces $I^2_{(j,\alpha)} : \mathbb{R} \longrightarrow \mathbb{R}^2$ are as follows:

$$I^2_{(1,0)}(x^1) = (0,\, x^1),$$
$$I^2_{(1,1)}(x^1) = (1,\, x^1),$$
$$I^2_{(2,0)}(x^1) = (x^1,\, 0),$$
$$I^2_{(2,1)}(x^1) = (x^1,\, 1).$$

Each of these has a natural orientation as x^1 increases from 0 to 1, but they are differing, so when we make the definition of the boundary of I^2, for example, we introduce $(-1)^{i+\alpha}$ to give each face the same orientation.

Definition 4.3 For a singular n-cube $c : [0, 1]^n \longrightarrow A$ ($A \subset \mathbb{R}^m$) the *boundary* of c is defined $\partial c = \sum_{i=1}^{n} (-1)^i\, c \circ I^n_{(i,0)} + (-1)^{i+1}\, c \circ I^n_{(i,1)}$.

Definition 4.4 An *n-chain* in $A \subset \mathbb{R}^m$ is a formal linear combination of singular n-cubes in $A \subset \mathbb{R}^m$ with integer coefficients.

Definition 4.5 For a singular n-chain $c = a_1\, c_1 + a_2\, c_2 + \cdots + a_m\, c_m$ the *boundary* of c is defined $\partial c = a_1\, \partial c_1 + a_2\, \partial c_2 + \cdots + a_m\, \partial c_m$.

Definition 4.6 For a k-form ω on $[0, 1]^k$, ω has the form $f \cdot dx^1 \wedge \cdots \wedge dx^k$ and we define $\int_{[0,1]^k} \omega = \int_{[0,1]^k} f$.

Definition 4.7 If ω is a k-form on A and c is a singular k-cube in A, then we define $\int_c \omega = \int_{[0,1]^k} c^*\, \omega$.

Definition 4.8 If ω is a k-form on A and c is a singular k-chain in A with $c = \sum_{i=1}^{m} a_i\, c_i$ for $a_i \in \mathbb{Z}$ and singular k-cubes c_i in A, then we define $\int_c \omega = \sum_{i=1}^{m} a_i \int_{c_i} \omega$.

Theorem 4.1 If dx^i is the 1-form on $\mathbb{R}^n$ with $dx^i\,(\boldsymbol{p}) \in \Lambda^1((\mathbb{R}^n)_{\boldsymbol{p}})$ for each $\boldsymbol{p} \in \mathbb{R}^n$ defined by $dx^i\,(\boldsymbol{p})\,(\boldsymbol{v_p}) = v^i$ for $\boldsymbol{v} \in \mathbb{R}^n$, and $k = n$, then $I^k_{(j,\alpha)}\ (dx^i) = 0$ if $i = j$ and ${I^k_{(j,\alpha)}}^*(dx^i) = dx^i$ if $i \neq j$.

Proof. Since $I^k_{(j,\alpha)} : [0, 1]^{k-1} \longrightarrow \mathbb{R}^k$ and dx^i is a 1-form on $\mathbb{R}^k$, then by Definition 3.18, $I^k_{(j,\alpha)}\ : \Lambda^1\left((\mathbb{R}^k)_{I^k_{(j,\alpha)}(\boldsymbol{p})}\right) \longrightarrow \Lambda^1\,((\mathbb{R}^{k-1})_{\boldsymbol{p}})$. Let $\boldsymbol{p} \in \mathbb{R}^{k-1}$ and $\boldsymbol{v_p} \in (\mathbb{R}^{k-1})$. Definition 3.18 defines $({I^k_{(j,\alpha)}}^*\, dx^i)\,(\boldsymbol{p})\,(\boldsymbol{v_p})$ to be $dx^i\,(I^k_{(j,\alpha)}\,(\boldsymbol{p}))\left((I^k_{(j,\alpha)})_*\,(\boldsymbol{v_p})\right)$, and

Definition 3.17 defines $(I^k_{(j,\alpha)})_*(v_p)$ to be $(DI^k_{(j,\alpha)}(p)(v))_{I^k_{(j,\alpha)}(p)}$. We note the component functions of $I^k_{(j,\alpha)}$ are as follows:

$$I^k_{(j,\alpha)}{}^{(1)} : \mathbb{R}^{k-1} \longrightarrow \mathbb{R} \text{ by } I^k_{(j,\alpha)}{}^{(1)}(x) = x^1$$
$$\vdots$$
$$I^k_{(j,\alpha)}{}^{(j-1)} : \mathbb{R}^{k-1} \longrightarrow \mathbb{R} \text{ by } {}^k_{(j,\alpha)}{}^{(j-1)}(x) = x^{j-1}$$
$$I^k_{(j,\alpha)}{}^{(j)} : \mathbb{R}^{k-1} \longrightarrow \mathbb{R} \text{ by } I^k_{(j,\alpha)}{}^{(j)}(x) = \alpha$$
$${}^k_{(j,\alpha)}{}^{(j+1)} : \mathbb{R}^{k-1} \longrightarrow \mathbb{R} \text{ by } I^k_{(j,\alpha)}{}^{(j+1)}(x) = x^j$$
$$\vdots$$
$$I^k_{(j,\alpha)}{}^{(k)} : \mathbb{R}^{k-1} \longrightarrow \mathbb{R} \text{ by } {}^k_{(j,\alpha)}{}^{(k)}(x) = x^{k-1}$$

So $DI^k_{(j,\alpha)}(p)$ has the following $k \times (k-1)$ Jacobian Matrix form.

$$\begin{array}{r l}
 & \\
 & \\
 & \\
 & \\
 & \\
j\text{th row} \rightarrow & \\
 & \\
 & \\
 & \\
 & \\
 &
\end{array}
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1
\end{pmatrix}$$

This is the same as inserting a row of zeros before the jth row of the identity matrix. We dot this matrix with the vector $v \in \mathbb{R}^{k-1}$ to arrive at the vector $(v^1, v^2, \ldots, v^{j-1}, 0, v^j, \ldots, v^{k-1}) \in \mathbb{R}^k$. Thus $dx^i\,(I^k_{(j,\alpha)}(p))\big((I^k_{(j,\alpha)})\ (v_p)\big)$ is the ith position of $(v^1, v^2, \ldots, v^{j-1}, 0, v^j, \ldots, v^{k-1})$. If $i = j$, the ith position is 0 and if $i \neq j$, the ith position is v^i. Since each of   and p is arbitrary then we have established our result. ∎

Theorem 4.2 If $f : [0,1]^k \longrightarrow \mathbb{R}$ and $\alpha = 0$ or $\alpha = 1$ then for $i, j \in \{1, 2, \ldots, k\}$,

$\int_{[0,1]^{k-1}} I^k_{(j,\alpha)}{}^*(f \cdot dx^1 \wedge \cdots \wedge dx^{i-1} \wedge dx^{i+1} \wedge \cdots \wedge dx^k)$ is 0 if $j \neq i$ and

$\int_{[0,1]^k} f(x^1, \ldots, \alpha, \ldots, x^k)\, dx^1 \cdots dx^k$ if $j = i$.

Proof. Let $f : [0, 1]^k \longrightarrow \mathbb{R}$ and $\alpha = 0$ or $\alpha = 1$. Let $\imath, \jmath \in \{1, 2, \ldots, k\}$. By Theorems 3.17 Part 3 and 4, ${}^{k}_{(\jmath,\alpha)}{}^{*}(f \cdot dx^1 \wedge \cdots \wedge dx^{\imath-1} \wedge dx^{\imath+1} \wedge \cdots \wedge dx^k) =$
$f \circ I^k_{(\jmath,\alpha)} \cdot I^k_{(\jmath,\alpha)}{}^{*}(dx^1) \wedge \cdots \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^{\imath-1}) \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^{\imath+1}) \wedge \cdots \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^k)$.

In one case assume $\imath \neq \jmath$. Thus $\jmath \in \{1, \ldots, i-1, \imath+1, \ldots, k\}$ and by Theorem 4.1 $I^k_{(\jmath,\alpha)}{}^{*}(dx^1) \wedge \cdots \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^{\imath-1}) \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^{\imath+1}) \wedge \cdots \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^k) = 0$, so in particular,

$$\int_{[0,1]^{k-1}} I^k_{(\jmath,\alpha)}{}^{*}(f \cdot dx^1 \wedge \cdots \wedge dx^{\imath-1} \wedge dx^{\imath+1} \wedge \cdots \wedge dx^k) = 0.$$

In the alternative case, assume $i = j$. We have $f \circ I^k_{(\jmath,\alpha)}(x) = f(x^1, \ldots, \alpha, \ldots, x^k)$, with α in the ith position, and by Theorem 4.1,

$$I^k_{(\jmath,\alpha)}{}^{*}(dx^1) \wedge \cdots \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^{\imath-1}) \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^{\imath+1}) \wedge \cdots \wedge I^k_{(\jmath,\alpha)}{}^{*}(dx^k) =$$
$$dx^1 \wedge \cdots \wedge dx^{\imath-1} \wedge dx^{\imath+1} \wedge \cdots \wedge dx^k, \text{ so}$$

$$\int_{[0,1]^{k-1}} I^k_{(\jmath,\alpha)}{}^{*}(f \cdot dx^1 \wedge \cdots \wedge dx^{\imath-1} \wedge dx^{\imath+1} \wedge \cdots \wedge dx^k) =$$
$$\int_{[0,1]^{k-1}} f(x^1, \ldots, \alpha, \ldots, x^k) \cdot dx^1 \wedge \cdots \wedge dx^{\imath-1} \wedge dx^{\imath+1} \wedge \cdots \wedge dx^k,$$

which reduces to the integrated integral
$\int_0^1 \cdots \int_0^1 \int_0^1 \cdots \int_0^1 f(x^1, \ldots, \alpha, \ldots, x^k)\, dx^1 \cdots dx^{\imath-1}\, dx^{\imath+1} \cdots dx^k$ by Definition 4.6 and Theorem 2.33 At the same time $\int_{[0,1]^k} f(x^1, \ldots, \alpha, \ldots, x^k)\, dx^1 \cdots dx^k$ is $\int_0^1 \cdots \int_0^1 \int_0^1 \cdots \int_0^1 \left(\int_0^1 f(x^1, \ldots, \alpha, \ldots, x^k)\, dx^\imath\right) dx^1 \cdots dx^{\imath-1}\, dx^{\imath+1} \cdots dx^k$ where α is in the ith position, and since $\int_0^1 f(x^1, \ldots, \alpha, \ldots, x^k)\, dx^\imath = f(x^1, \ldots, \alpha, \ldots, x^k) \int_0^1 dx^\imath$ by Theorem 2.26 and $\int_0^1 dx^\imath = 1$ by Theorem 2.30, then our iterated integral reduces to $\int_0^1 \cdots \int_0^1 \int_0^1 \cdots \int_0^1 f(x^1, \ldots, \alpha, \ldots, x^k)\, dx^1 \cdots dx^{\imath-1}\, dx^{\imath+1} \cdots dx^k$. We have shown the result in the alternate case, and established the theorem. ∎

Theorem 4.3 If ω is a differentiable $(k-1)$-form on $[0, 1]^k$ and I^k is the standard k-cube in $[0, 1]^k$, then $\int_{I^k} d\omega = \int_{\partial I^k} \omega$.

Proof. Let ω is a differentiable $(k-1)$-form on $[0, 1]^k$ and I^k is the standard k-cube in $[0, 1]^k$. Since ω is a $(k-1)$-form on $[0, 1]^k$, then we can simplify the form of ω in terms of its basis in Definition 3.14 and write

$\omega = \sum_{i=1}^{k} \omega_i \cdot dx^1 \wedge \cdots \wedge dx^{i-1} \wedge dx^{i+1} \wedge \cdots \wedge dx^k$. Starting with the left-hand-side of our result, we have $\int_{I^k} d\omega = \int_{[0,1]^k} I^*(d\omega)$.

First $d\omega = \sum_{i=1}^{k} \sum_{j=1}^{k} D_j\, \omega_i \cdot dx^j \wedge dx^1 \wedge \cdots \wedge dx^{i-1} \wedge dx^{i+1} \wedge \cdots \wedge dx^k$ by Definition 3.19, which simplifies to $\sum_{i=1}^{k} D_i\, \omega_i \cdot dx^i \wedge dx^1 \wedge \cdots \wedge dx^{i-1} \wedge dx^{i+1} \wedge \cdots \wedge dx^k$ by a corollary to Theorem 3.8 Part 4, any wedge product with a repeated basis element is zero, thus the only term of the inner sum to survive the differential operator is the one where $j = i$, since dx^i is omitted in the wedge product. A second application of Theorem 3.8 Part 4 allows us to arrange $dx^i \wedge dx^1 \wedge \cdots \wedge dx^{i-1} \wedge dx^{i+1} \wedge \cdots \wedge dx^k$ in standard form with the coefficient $(-1)^{i-1}$. We have then

$d\omega = \sum_{i=1}^{k} (-1)^{i-1} D_i\, \omega_i \cdot dx^1 \wedge \cdots \wedge dx^k$.

Second $I^*(d\omega) = \sum_{i=1}^{k} ((-1)^{i-1} D_i\, \omega_i \circ I) \cdot I^*\, dx^1 \wedge \cdots \wedge I^*\, dx^k$ by Theorem 3.17 Parts 3 and 4 Part 1 of Theorem 3.17 gives $I^*\, dx^i = \sum_{j=1}^{k} D_j\, I^{(i)} \cdot dx^j$ where $I^{(i)}$ is the ith component function of I, so $I^{(i)}(x) = x^i$. It follows then that $D_j\, I^{(i)} = 0$ if $j \neq i$ and $D_j\, I^{(i)} = 1$ if $j = i$, thus $\sum_{j=1}^{k} D_j\, I^{(i)} \cdot dx^j = dx^i$. Also since $I(x) = x$, then $D_i\, \omega_i \circ I = D_i\, \omega_i$. We see then that $I^*(d\omega) = \sum_{i=1}^{k} (-1)^{i-1} D_i\, \omega_i \cdot dx^1 \wedge \cdots \wedge dx^k$.

Now $\int_{[0,1]^k} I^*(d\omega) = \int_{[0,1]^k} \sum_{i=1}^{k} (-1)^{i-1} D_i\, \omega_i \cdot dx^1 \wedge \cdots \wedge dx^k$, but the right hand side is equivalent to $\sum_{i=1}^{k} (-1)^{i-1} \int_{[0,1]^k} D_i\, \omega_i \cdot dx^1 \wedge \cdots \wedge dx^k$ by Theorems 2.21 and 2.26. We use Definition 4.6 and Theorem 2.33 to write $\int_{[0,1]^k} D_i\, \omega_i \cdot dx^1 \wedge \cdots \wedge dx^k$ as $\int_0^1 \cdots \int_0^1 D_i\, \omega_i\, dx^1 \cdots dx^k$, but the corollary to Theorem 2.33, Theorem 2.34 allows us to equivalently write $\int_0^1 \cdots \int_0^1 \int_0^1 \cdots \int_0^1 \int_0^1 D_i\, \omega_i\, dx^i\, dx^1 \cdots dx^{i-1}\, dx^{i+1} \cdots dx^k$. The Fundamental Theorem of Calculus, Theorem 2.29 gives $\int_0^1 D_i\, \omega_i\, dx^i =$ $\omega_i(x^1, \ldots, 1, \ldots, x^k) - \omega_i(x^1, \ldots, 0, \ldots, x^k)$ with 1 and 0 in the ith positions. Now as was the case in Theorem 4.2, $\int_0^1 \omega_i(x^1, \ldots, \alpha, \ldots, x^k)\, dx^i = \omega_i(x^1, \ldots, \alpha, \ldots, x^k) \int_0^1 dx^i =$

$\omega_\imath(x^1, \ldots, \alpha, \ldots, x^k)$ for $\alpha = 0, 1$ in the $\imath$th position. Therefore $\int_0^1 D_\imath\, \omega_\imath\, dx^\imath =$ $\int_0^1 \omega_\imath(x^1, \ldots, 1, \ldots, x^k)\, dx^\imath - \int_0^1 \omega_\imath(x^1, \ldots, 0, \ldots, x^k)\, dx^\imath$. We make this substitution back in $\int_0^1 \cdots \int_0^1 \int_0^1 \cdots \int_0^1 \int_0^1 D_\imath\, \omega_\imath\, dx^\imath\, dx^1 \cdots dx^{\imath-1}\, dx^{\imath+1} \cdots dx^k$, move $dx^\imath$ back into position with Fubini's Theorem and have

$$\int_{[0,1]^k} D_\imath\, \omega_\imath \cdot dx^1 \wedge \cdots \wedge dx^k =$$

$$\int_{[0,1]^k} \omega_\imath(x^1, \ldots, 1, \ldots, x^k) - \omega_\imath(x^1, \ldots, 0, \ldots, x^k) \cdot dx^1 \wedge \cdots \wedge dx^k.$$

Using Theorem 2.21 we have

$$\int_{[0,1]^k} \omega_\imath(x^1, \ldots, 1, \ldots, x^k) \cdot dx^1 \wedge \cdots \wedge dx^k -$$

$$\int_{[0,1]^k} \omega_\imath(x^1, \ldots, 0, \ldots, x^k) \cdot dx^1 \wedge \cdots \wedge dx^k,$$

which we put back into our sum $\sum_{\imath=1}^{k} (-1)^{\imath-1} \int_{[0,1]^k} D_\imath\, \omega_\imath \cdot dx^1 \wedge \cdots \wedge dx^k$ to end with

$$\sum_{\imath=1}^{k} (-1)^{\imath-1} \int_{[0,1]^k} \omega_\imath(x^1, \ldots, 1, \ldots, x^k) \cdot dx^1 \wedge \cdots \wedge dx^k +$$

$$(-1)^\imath \int_{[0,1]^k} \omega_\imath(x^1, \ldots, 0, \ldots, x^k) \cdot dx^1 \wedge \cdots \wedge dx^k$$

after distributing $(-1)^{\imath-1}$.

Now we investigate the right-hand-side of our result. First $\int_{\partial I^k} \omega = \int_{[0,1]^{k-1}} (\partial I^k)^* \omega$ by Definition 4.7. Next $(\partial I^k)^* \omega = \sum_{\imath=1}^{k} ((-1)^\imath I^k_{(\imath,0)}{}^* \omega + (-1)^{\imath+1} I^k_{(\imath,1)}{}^* \omega)$ by Definitions 4.3 and 3.12 and $\int_{[0,1]^{k-1}} (\partial I^k)^* \omega = \sum_{\imath=1}^{k} \left((-1)^\imath \int_{[0,1]^{k-1}} I^k_{(\imath,0)}{}^* \omega + (-1)^{\imath+1} \int_{[0,1]^{k-1}} I^k_{(\imath,1)}{}^* \omega \right)$ as the integral of a sum is the sum of integrals by Theorem 2.21. For $\alpha = 0, 1$, $I^k_{(\imath,\alpha)}{}^* \omega = I^k_{(\imath,\alpha)}{}^* \left(\sum_{\jmath=1}^{k} \omega_\jmath \cdot dx^1 \wedge \cdots \wedge dx^{\jmath-1} \wedge dx^{\jmath+1} \wedge \cdots \wedge dx^k \right)$, which is $\sum_{\jmath=1}^{k} I^k_{(\imath,\alpha)}{}^* (\omega_\jmath \cdot dx^1 \wedge \cdots \wedge dx^{\jmath-1} \wedge dx^{\jmath+1} \wedge \cdots \wedge dx^k)$ by Theorem 3.17, and we have two of these sums, one for $\alpha = 0$ and one for $\alpha = 1$. Now

$$\int_{[0,1]^{k-1}} I^k_{(i,\alpha)} = \sum_{j=1}^{k} \int_{[0,1]^{k-1}} I^k_{(i,\alpha)}{}^*(\omega_j \cdot dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^k),$$

but each term in this sum is identically zero except for the ith term which is

$\int_{[0,1]^k} \omega_i(x^1, \ldots, \alpha, \ldots, x^k)\, dx^1 \cdots dx^k$ with α in the ith position by Theorem 4.2. Therefore we have

$$\sum_{i=1}^{k} \left((-1)^i \int_{[0,1]^{k-1}} I^k_{(i,0)}{}^* \omega + (-1)^{i+1} \int_{[0,1]^{k-1}} I^k_{(i,1)}{}^* \omega \right) =$$
$$\sum_{i=1}^{k} (-1)^i \int_{[0,1]^k} \omega_i(x^1, \ldots, 0, \ldots, x^k)\, dx^1 \cdots dx^k +$$
$$(-1)^{i+1} \int_{[0,1]^k} \omega_i(x^1, \ldots, 1, \ldots, x^k)\, dx^1 \cdots dx^k.$$

This is precisely the expression we obtained for the left-hand-side; thus, we have established our result. ∎

Stokes' Theorem

Theorem 4.4 If ω is a differentiable $(k-1)$-form on an open set $A \subseteq \mathbb{R}^n$ and c is a k-chain in A, then $\int_c d\omega = \int_{\partial c} \omega$.

Proof. Let ω be a $(k-1)$-form on an open set $A \subseteq \mathbb{R}^n$ and c be a k-chain in A with $c = \sum_{i=1}^{m} a_i\, c_i$ for $a_i \in \mathbb{Z}$ and singular k-cubes c_i in A. By Definition 3.19, $d\omega$ is a differentiable k-form so by Definition 4.8, $\int_c d\omega = \sum_{i=1}^{m} a_i \int_{c_i} d\omega$. Since $\partial c = \sum_{i=1}^{m} a_i\, \partial c_i$ by Definition 4.5 and ∂c is a $(k-1)$-chain, then by Definition 4.8, $\int_{\partial c} \omega = \sum_{i=1}^{m} a_i \int_{\partial c_i} \omega$. Therefore the theorem is true if $\int_{c_i} d\omega = \int_{\partial c_i} \omega$ for each i. For a particular i, $\int_{c_i} d\omega = \int_{[0,1]^k} c_i{}^*(d\omega)$ by Definition 4.7 and since $c_i{}^*(d\omega) = d(c_i{}^*\, \omega)$ by Theorem 3.21, then $\int_{[0,1]^k} c_i{}^*(d\omega) = \int_{[0,1]^k} d(c_i{}^*\, \omega) = \int_{I^k} d(c_i{}^*\, \omega)$. On the other hand, $\int_{\partial c_i} \omega = \int_{\partial I^k} c_i{}^*\, \omega$. We have shown in Theorem 4.3 that $\int_{k} d(c_i{}^*\, \omega) = \int_{\partial I^k} c_i{}^*\, \omega$; thus, we established the final result of this paper. ∎

REFERENCES

Boyer, Carl B. A History of Mathematics. New York: John Wiley & Sons, Inc., 1991.

O'Conner, J. J. and E. F. Robertson. *George Gabriel Stokes.* Available from http://www-history.mcs.st-andrews.ac.uk/history/index.html.

Royden, H. L. Real Analysis. New York: Macmillian, 1968.

Spivak, Michael. Calculus on Manifolds. New York: Addison-Wesley Publishing Company, 1965.

Spivak, Michael. A Comprehensive Introduction to Differential Geometry. Boston: Publish or Perish, Inc., 1970.

Stoker, J. J. Differential Geometry. New York: Wiley-Interscience, 1969.

Stokes, G. G. Mathematical and Physical Papers. Cambridge: University Press, 1880.

VITA

Christopher Elliot Johnson is the eldest of six children, born 4 March 1977, to Brent Gustave Johnson and Judy Arlene Blossom in Homer, Alaska. He attended high school for a time at Waimea High School on the island of Kauai, Hawaii, but he spent the majority of time at Skyview High School in Soldotna, Alaska, and graduated from there in May 1995. The same year he began study at Sam Houston State University in Huntsville, Texas. In the fall of 1997, he transferred to Texas State University – San Marcos to follow his love, and the following spring, he married Mary Denise Wilson on 14 February 1998.

In the fall of 1999, he began teaching as an undergraduate instructional assistant while continuing his study of mathematics, and the following spring, he earned a Bachelor of Science from Texas State with a major in mathematics and a minor in physics. The very same fall, August 2000, he entered the Graduate College of Texas State with a teaching assistantship.

In the fall of 2001, he took a break from his mathematics study and began a study of the French language and culture at the Université d'Aix-Marseille III in Aix-en-Provence, France. He spent six months abroad then returned to his graduate study and teaching post at Texas State University – San Marcos.

Permanent Address: 20773 Porcupine Lane

Clam Gulch AK 99568

This thesis was typed by Christopher Elliot Johnson.