VECTOR FACTS & FORMULAS

MAURICE J. DUPRÉ

1. VECTOR SPACES

1.1. **VECTOR SPACE or LINEAR SPACE.** A set V of objects which can be added and multiplied by scalars. The objects of V are called **Vectors**. The addition rule is called **Vector Addition** and the rule for multiplication by scalars is called **Scalar Multiplication**. The set of scalars can be taken here to be the set of all real numbers, denoted $\mathbb{R}$, but can also be any number system forming an algebraic system called a field, for instance the complex number system, $\mathbb{C}$, is used as the field of scalars for the vector spaces in quantum physics. Formulas **2-8** below are assumed as axioms.

1.2. **ASSOCIATIVE LAW OF VECTOR ADDITION.**

$$(x + y) + z = x + (y + z)$$

1.3. **COMMUTATIVE LAW OF VECTOR ADDITION.**

$$x + y = y + x$$

1.4. **ZERO VECTOR.**

$$0_V + x = x \text{ and if } x + x = x, \text{ then } x = 0_V$$

1.5. **ASSOCIATIVE LAW OF SCALAR MULTIPLICATION.**

$$r(sx) = (rs)x, \text{ for any scalars } r, s \text{ and any vector } x$$

1.6. **RIGHT DISTRIBUTIVE LAW OF SCALAR MULTIPLICATION.**

$$(r + s)x = rx + sx, \text{ for any scalars } r, s \text{ and any vector } x$$

1.7. **LEFT DISTRIBUTIVE LAW OF SCALAR MULTIPLICATION.**

$$r(x + y) = rx + ry \text{ for any scalar } r \text{ and any vectors } x, y$$

1.8. **UNIT IDENTITY.**

$$1x = x$$

Date: February 17, 2013.

2. FUNCTIONS & MAPPINGS

2.1. **DEFINITION OF FUNCTION OR MAPPING.** A **Function** or **Mapping** or **Transformation** is a rule which assigns a member of a set to each member of a set. More specifically, we say f is a function or mapping of set A to set B, provided that f is a rule which assigns a member $f(x)$ in B to each member x in A. Note that the notation $f(x)$ is NOT multiplication in general, even though in some situations it might be. We use the notation

$$f : A \longrightarrow B$$

to mean f is a function or mapping from set A to set B. If $f : A \longrightarrow B$ is a function we call A the **Domain** of f and we call B the **Codomain** of f. In order for two functions f and g to be equal it is necessary and sufficient that they have the same domain, the same codomain, and the same rule, that is, $f(x) = g(x)$ for each x in their domain. Think of a function f as an input-output device with domain as the set of allowable inputs and codomain as a set containing the outputs.

2.2. **COMPOSITION OF FUNCTIONS.** If $f : A \longrightarrow B$ and $g : B \longrightarrow C$ are both functions, we can form their **Composition** by using the outputs of f as the inputs for g, resulting in the new function

$$g \circ f : A \longrightarrow C$$

with rule

$$[g \circ f](x) = g(f(x)), \text{ for each } x \text{ in } A.$$

2.3. **ASSOCIATIVE LAW OF COMPOSITION OF FUNCTIONS.** If U, V, W, X are all sets and if $R : U \longrightarrow V,\ S : V \longrightarrow W,\ T : W \longrightarrow X$ are all functions, then

$$(T \circ S) \circ R = T \circ (S \circ R)$$

2.4. **IDENTITY FUNCTION.** If S is any set, then the **Identity Function** on S, denoted $Id_S : S \longrightarrow S$ is defined by the rule

$$Id_S(x) = x, \text{ for any } x \text{ in } S.$$

Note that if S and T are any sets and if $f : S \longrightarrow T$ is any function, then

$$Id_T \circ f = f = f \circ Id_S.$$

2.5. **INVERSE FUNCTION.** If S and T are sets, we say that $f : S \longrightarrow T$ and $g : T \longrightarrow S$ are **Mutually Inverse Functions** provided that

$$g \circ f = Id_S \text{ and } f \circ g = Id_T.$$

We note that this is equivalent to saying that $g(f(x)) = x$ for any x in S and $f(g(y)) = y$, for any y in T. Thus, each undoes what the other does. We say that g is the **Inverse** of f and that f is the inverse of g. Moreover, in this situation, each function completely determines the other, and we write

$$g = f^{-1} \text{ and } f = g^{-1}.$$

2.6. **SETS OF FUNCTIONS.** If R and S are sets, then the set of all functions with domain R and codomain S is denoted by S^R.

$$S^R = \{f \mid f : R \longrightarrow S \text{ is a function }\}.$$

2.7. **IMAGE OF A SUBSET UNDER A FUNCTION.** If $f : B \longrightarrow D$ and if A is a subset of B, in symbols, $A \subset B$, then we define the **Image** of A under f, denoted by $f(A)$, as the subset of D given by

$$f(A) = \{f(x) \mid x \text{ in } A\} \subset D.$$

2.8. **RANGE OR IMAGE OF A FUNCTION.** If $f : B \longrightarrow C$, then the **Range** of f, denoted $\text{Im}(f)$ is the image of the whole domain of f, so

$$\text{Range of } f = \text{Im}(f) = f(B) \subset C.$$

2.9. **INVERSE IMAGE OF A SUBSET UNDER A FUNCTION.** If $f : B \longrightarrow D$ and if C is a subset of D, in symbols, $C \subset D$, then we define the **Inverse Image** of C under f, denoted $f^{-1}(C)$, as the subset of B given by

$$f^{-1}(C) = \{x \text{ in } B \mid f(x) \text{ in } C\}.$$

2.10. **CARTESIAN PRODUCT OF SETS.** If A and B are sets, then their **Cartesian Product**, denoted $A \times B$, is the set of all ordered pairs (x, y) with x in A and y in B.

2.11. **GRAPH OF A FUNCTION.** If V and W are any sets and $T : V \longrightarrow W$ is a function, then the **Graph** of T, denoted $\text{Graph}(T)$ is the subset of $V \times W$ consiting of all pairs (v, w) for which $w = T(v)$.

$$(v, w) \text{ belongs to Graph}(T) \text{ if and only if } w = T(v).$$

3. FUNCTIONS AND TRANSFORMATIONS OF VECTORS

3.1. **PARAMETRIZED LINE.** For any vectors u and v the set L consisting of vectors $x(t)$ for t in $\mathbb{R}$ is the **Line** through u with direction (velocity) v provided

$$x(t) = u + tv, \text{ for all scalars } t \text{ in } \mathbb{R}$$

or, more precisely, $x : \mathbb{R} \longrightarrow V$ is the function with rule $x(t) = u + tv$, for t in $\mathbb{R}$, and $L = x(\mathbb{R})$ is the image of $\mathbb{R}$ under the function x.

3.2. **LINEAR TRANSFORMATIONS.** If V and W are vector spaces, then any function $T : V \longrightarrow W$ is called a **Linear Transformation** provided that T preserves all lines:

$$T(x + ry) = T(x) + rT(y) \text{ for any scalar } r \text{ and any vectors } x, y.$$

Equivalently, if $x : \mathbb{R} \longrightarrow V$ is any line through u with velocity v, then $T \circ x : \mathbb{R} \longrightarrow W$ is the line through $T(u)$ with velocity $T(v)$.

NOTE: T is linear if and only if T is both

$$\textbf{Additive: } T(x + y) = T(x) + T(y), \text{ for any vectors } x, y,$$

and

$$\textbf{Homogeneous: } T(rx) = rT(x), \text{ for any scalar } r \text{ and any vector } x.$$

3.3. **THE COMPOSITION OF LINEAR TRANSFORMATIONS IS LINEAR.** If $S : U \longrightarrow V$ and $T : V \longrightarrow W$ are both linear transformations of vector spaces, then so is $T \circ S$. Moreover, the function Id_V is linear if V is any vector space. If $U = W$ and if S and T are mutually inverse, then if either is linear then both are linear and we say each is a **Linear Isomorphism**, and we say that U and V are **Isomorphic**. In this case, U and V are essentially the same vector space with vectors simply given different names via the isomorphism.

3.4. **NOTATION.** If $T : V \longrightarrow W$ is a linear transformation of vector spaces, then the usual function notation is often abreviated to

$$T(v) = Tv, \text{ for any vector } v \text{ in } V$$

and if $R : U \longrightarrow V$ and $T : V \longrightarrow W$ are both linear transformations, the composition notation is often abbreviated to

$$T \circ R = TR,$$

and we think of the composition of linear transformations as a way of multiplying linear transformations, that is, as a generalization of ordinary multiplication but which is not commutative, since the order of composition of functions matters.

3.5. **VECTOR VALUED FUNCTIONS ON A SET FORM A VECTOR SPACE.** If S is any set and V is any vector space, and if f, g both belong to V^S, then $f : S \longrightarrow V$ and $g : S \longrightarrow V$ and we define $f + g$ by

$$[f+g](x) = f(x) + g(x), \text{ for any } x \text{ belonging to } S$$

and we define scalar multiplication rf by

$$[rf](x) = r[f(x)], \text{ for any } x \text{ belonging to } S;$$

it is easy to see the axioms are all true here. Therefore, if S is any set and V is any vector space, then V^S is a vector space. As $\mathbb{R}$ is a vector space, it follows that $[\mathbb{R}]^S$ is a vector space for any set S.

We define V^n to be the vector space $\mathbb{R}^S$ where S is the set of positive integers k with $1 \leq k \leq n$. If x belongs to V^n, then x is completely determined by its list of values and we write these values as x_k instead of using the usual function notation $x(k)$, so we can also write

$$x = (x(1), x(2), x(3), ..., x(n)) = (x_1, x_2, x_3, ..., x_n), \text{ with each } x_k \text{ in } V, \ 1 \leq k \leq n.$$

Such a list enclosed by parenthesis is called an $n-$tuple, where n is of course the length of the list. In particular, as $\mathbb{R}$ is a vector space, it follows that $\mathbb{R}^n$ is a vector space, consisting of all possible lists of n numbers, that is, all possible $n-$tuples of real numbers.

3.6. **THE VECTOR SPACE OF LINEAR TRANSFORMATIONS.** If U and V are both vector spaces, then V^U is the set of all functions from U to V and therefore all linear transformations from U to V also belong to V^U. The subset of V^U consisting of linear transformations is denoted $L(U;V)$.

$$L(U;V) = \{T \text{ in } V^U \mid T \text{ is linear } \}.$$

It follows easily that $L(U;V)$ is a vector subspace of V^U. In particular it is customary to denote

$$V^\dagger = L(V;\mathbb{R}),$$

and $V^\dagger$ is called the **Dual space** to V.

3.7. **DISTRIBUTIVE LAWS.** If U, V, W are all vector spaces, if S, S_1, S_2 are in $L(U;V)$ and T, T_1, T_2 are in $L(V;W)$, then

$$S[T_1 + T_2] = ST_1 + ST_2 \text{ and } [S_1 + S_2]T = S_1T + S_2T.$$

Because of these distributive laws, composition of linear transformations behaves like an ordinary multiplication except that the order matters, so it is not commutative.

3.8. **VECTOR SUBSPACE OR LINEAR SUBSPACE.** A subset U of the vector space V is called a **Vector Subspace** or **Linear Subspace** provided that for any vectors u, v which belong to U it is the case that the line through u with velocity v is entirely contained in U. In order that U be a linear subspace, it is necessary and sufficient that the sum of any two vectors belonging to U again belongs to U and any scalar multiple of a vector in U again belongs to U. For example, the set consisting of the zero vector alone is a linear subspace also denoted 0_V and as well, V is a linear subspace of itself. In addition any line through 0_V is a vector subspace of V. If W is also a vector space and if $T : V \longrightarrow W$ is a linear transformation and if U is a vector subspace of V, then

$$TU = \{Tu \mid u \text{ belongs to } U\} \subset W \text{ is a vector subspace of } W,$$

the **Image** of U under T. On the other hand, if X is a vector subspace of W, then

$$T^{-1}X = \{v \text{ in } V \mid Tv \text{ is in } X\} \subset V \text{ is a vector subspace of } V,$$

the **Inverse Image** of X under T. In particular, we define

$$\mathrm{Im}(T) = TV \text{ and } \mathrm{Ker}(T) = T^{-1}0_W.$$

We call $\mathrm{Im}(T)$ the **Image** or **Range** of T and we call $\mathrm{Ker}(T)$ the **kernel** of T. It follows that T is surjective or onto if and only if the image of T is all of W and T is injective or one-to-one if and only if $\mathrm{Ker}(T) = 0_V$. Thus, if $\mathrm{Im}(T) = W$ and $\mathrm{Ker}(T) = 0_V$, then T is a bijection and therefore a vector space isomorphism.

3.9. **CARTESIAN PRODUCT OF VECTOR SPACES.** If V and W are vector spaces, the cartesian product $V \times W$ which consists of all ordered pairs of vectors (v, w) where v belongs to V and w belongs to W is made a vector space where the vector addition rule is:

$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1, w_2),$$

and the scalar multiplication rule is:

$$r(v, w) = (rv, rw).$$

More generally, if $V_1, V_2, V_3, ..., V_n$ are all vector spaces, then we can similarly make the cartesian product $V_1 \times V_2 \times V_3 \times ... \times V_n$ into a vector space. As a set, it has as members all possible $n-$tuples of vectors $(v_1, v_2, v_3, ..., v_n)$ where v_k belongs too V_k for each k with $1 \leq k \leq n$. To add the $n-$tuples v and w, where $v = (v_1, v_2, v_3, ..., v_n)$ and $w = (w_1, w_2, w_3, ..., w_n)$, we use the formula

$$v + w = (v_1 + w_1, v_2 + w_2, v_3 + w_3, ..., v_n + w_n),$$

and to multiply the $n-$tuple v by the scalar r, we use the formula

$$rv = (rv_1, rv_2, rv_3, ..., rv_n).$$

We can say here that the operations are defined "slot wise". Notice that the vectors in the kth slot all come from the same vector space V_k so it makes sense to add them.

3.10. **GRAPH OF A LINEAR TRANSFORMATION.** If U and V are linear spaces and if $T : U \longrightarrow V$ is any function, then T is linear if and only if Graph(T) is a linear subspace of $U \times V$.

3.11. **DUAL OF A LINEAR TRANSFORMATION.** If U and V are linear spaces and $S : U \longrightarrow V$ is a linear transformations, then the **Dual** of S is the linear transformation denoted $S^\dagger : V^\dagger \longrightarrow U^\dagger$ given by the rule

$$S^\dagger\lambda = \lambda S, \text{ for every } \lambda \text{ in } V^\dagger.$$

The linearity of $T^\dagger$ follows from the distributive laws above. Also, from the distributive and associative laws, it follows that the function

$$\mathrm{DAG} : L(U; V) \longrightarrow L(V^\dagger : U^\dagger)$$

with rule

$$\mathrm{DAG}(S) = S^\dagger, \text{ for every } S \text{ in } L(U; V),$$

is linear. Moreover, if T is in $L(V; W)$ then TS is in $L(U; W)$ and

$$(TS)^\dagger = S^\dagger T^\dagger.$$

Also,

$$(Id_V)^\dagger = Id_{V^\dagger},$$

and if T is invertible, then so is $T^\dagger$, and

$$(T^\dagger)^{-1} = (T^{-1})^\dagger.$$

3.12. **THE DOUBLE DUAL.** If V is a vector space, then for v in V we can define a function $e_v : V^\dagger \longrightarrow \mathbb{R}$ by the rule

$$e_v(\lambda) = \lambda v, \text{ for all } \lambda \text{ in } V^\dagger.$$

Then the distributive law guarantees that e_v is linear and therefore belongs to the **Double Dual** of V, namely,

$$V^{\dagger\dagger} = (V^\dagger)^\dagger.$$

Moreover, the distributive laws also guarantee that the mapping

$$E : V \longrightarrow V^{\dagger\dagger}$$

with rule

$$Ev = e_v, \text{ for every } v \text{ in } V,$$

is a linear transformation. If $Ev = 0$, then

$$0 = [Ev]\lambda = e_v\lambda = \lambda v, \text{ for every } \lambda \text{ in } V^\dagger.$$

This in fact guarantees that $v = 0$. Thus, E is injective and defines a vector space isomorphism of V onto a linear subspace of $V^{\dagger\dagger}$. In particular, if V is finite dimensional, then

$$\dim(V) = \dim(V^{\dagger}) = \dim(V^{\dagger\dagger}).$$

Thus, for V a finite dimensional vector space, we can, via the very natural transformation E, regard

$$V = V^{\dagger\dagger}.$$

3.13. **LINEAR COMBINATION AND SPAN.** If V is a vector space and if S is a subset of V we say that a vector v in V is a **Linear Combination** of vectors in S provided that there are vectors $v_1, v_2, v_3, ..., v_n$ belonging to S and scalars $r_1, r_2, r_3, ..., r_n$ such that

$$v = r_1 v_1 + r_2 v_2 + r_3 v_3 + ... + r_n v_n.$$

The numbers $r_1, r_2, r_3, ..., r_n$ are called the **Coefficients** of the linear combination. The set of all vectors in V which are expressible as linear combinations of vectors in S is called the **Linear Span** of the subset S and denoted $\mathrm{Span}(S)$. Obviously the sum and scalar multiple of any two vectors in $\mathrm{Span}(S)$ is again in $\mathrm{Span}(S)$. It follows that $\mathrm{Span}(S)$ is a vector subspace of V. If $\mathrm{Span}(S)$ is all of V, the we say that S **Spans** V.

If $R, T : V \longrightarrow W$ are linear maps and S is a subset of V and if $Rv = Tv$ for every v in S, then in fact $Rv = Tv$ for every vector in $\mathrm{Span}(S)$. Thus if S spans V, then every linear transformation on V is determined on S.

If $v = (v_1, v_2, v_3, ..., v_n)$ is in V^n, then we define $T_v : \mathbb{R}^n \longrightarrow V$ by the rule

$$T_v(r_1, r_2, r_3, ..., r_n) = r_1 v_1 + r_2 v_2 + r_3 v_3 + ... + r_n v_n.$$

Then T_v is linear, and $T_v(\mathbb{R}^n)$ is the linear subspace of V spanned by the vectors in the set $\{v_1, v_2, v_3, ..., v_n\}$.

4. COORDINATE AND MATRIX DESCRIPTIONS OF VECTORS AND LINEAR TRANSFORMATIONS

4.1. **MATRIX AND TRANSPOSE.** If L is any horizontal list, then $L^\dagger$ denotes the same list written vertically and we call $L^\dagger$ the **Transpose** of L. We define $\mathbb{R}_n$ to be the result of transposing all the vectors in $\mathbb{R}^n$, so $\mathbb{R}_n$ is a vector space isomorphic to $\mathbb{R}^n$. If K is any vertical list, then $K^\dagger$ denotes the horizontal list obtained by writing the same list horizontally instead of vertically. Thus $S : \mathbb{R}^n \longrightarrow \mathbb{R}_n$ defined by $Sv = v^\dagger$ and $T : \mathbb{R}_n \longrightarrow \mathbb{R}^n$ defined by $Tv = v^\dagger$ are mutually inverse vector space isomorphisms.

More generally, we define $\mathbb{R}^n_m = [\mathbb{R}_m]^n$, so that $\mathbb{R}^n_m$ consists of horizontal lists of vertical lists forming a rectangular array having m rows and n columns. The members of $\mathbb{R}^n_m$ are called **Matrices**, and more specifically, they are called m by n matrices. Thus the set of all m by n matrices is a vector space. If M is an m by n matrix with columns $(K_1, K_2, K_3, ..., K_n)$, then

$$M = [K_1, K_2, K_3, ..., K_n],$$

and $M^\dagger$ is the n by m matrix having row i given by $K_i^\dagger$. Thus all the columns of M become the rows of $M^\dagger$ and the rows of M become the columns of $M^\dagger$. Notice that

$$[M^\dagger]^\dagger = M.$$

If $T^n_m : \mathbb{R}^n_m \longrightarrow \mathbb{R}^m_n$ is defined by $T(M) = M^\dagger$, for M in $\mathbb{R}^n_m$, then T^n_m and $T^m_n : \mathbb{R}^m_n \longrightarrow \mathbb{R}^n_m$ are mutually inverse isomorphisms of vector spaces. A useful notation is to let M_k be the kth column of M and let M^k be the kth row of M. Thus,

$$[M^\dagger]_k = [M^k]^\dagger \text{ and } [M^\dagger]^k = [M_k]^\dagger.$$

4.2. **FRAME IN A VECTOR SPACE.** If V is a vector space, a **Frame** is an ordered list of n vectors $F = (b_1, b_2, b_3, ..., b_n)$ in V^n, also called an $n-$frame, provided that $T_F : \mathbb{R}^n \longrightarrow V$ is injective or one-to-one. Thus F is a frame if and only if $\text{Ker}(T_F) = 0$. Then the only way to form 0_V as a linear combination of frame vectors is to have all coefficients equal to zero. We say that a frame in V is a **Frame For** V provided that the frame vectors also span V. Thus, F is a frame for V if and only if T_F is a vector space isomorphism of $\mathbb{R}^n$ onto V. If $F^\dagger = (b^1, b^2, b^3, ..., b^n)$ is a frame for $V^\dagger$, so that $b^i b_j$ is one or zero according to whether i and j are the same or not, then $F^\dagger$ is said to be the **Dual Frame** to F and is in fact uniquely determined by F. Keep in mind that the superscipts in the preceding notations are not to be exponents but merely tags. If F is a frame for V, then a unique dual frame $F^\dagger$ for $V^\dagger$ always exists.

More generally, if A is any subset of V, we say that A is **Linearly Independent** if any $n-$tuple of distinct vectors in S, no matter how big n, is a frame in V. If B is a linearly independent subset of V which also spans V, then B is called a **Basis** for V.

If A as a subset of C and C is a subset of V which spans V and if A is linearly independent, then there is a subset B of C which contains A as a subset and which forms a basis for V. In particular, if $v \neq 0_V$, then $\{v\}$ is a linearly independent set, and is a subset of V, and V itself certainly spans V, so therefore every vector space has a basis. Thus, given any nonzero vector, we can find a basis for V containing that nonzero vector as one of the basis vectors.

4.3. **DIMENSION OF A VECTOR SPACE.** If B_1 and B_2 are subsets of V and if each is a basis for V, then there is a bijective map of B_1 onto B_2, so they have the same **Cardinality** which we call the **Dimension** of V.

If V is a vector space and $F = (b_1, b_2, b_3, ..., b_n)$ is a frame for V then we say that n is the **Dimension** of V and we write

$$n = \dim(V).$$

We say that V is **Finite Dimensional** if V has a finite basis. The set of vectors in an ordered list which form a frame for V are thus also a basis for V which is finite. Any two bases for V must have the same cardinality. In particular, if F is an $m-$frame for V and if G is an $n-$frame for V, then $m = n$. Thus, if V is finite dimensional, then there is an $n-$frame for V, where $n = \dim(V)$.

If V and W are finite dimensional, then so are $V \times W$ and $L(V; W)$. In fact,

$$\dim(V \times W) = [\dim(V)] + [\dim(W)],$$

and

$$\dim(L(V; W)) = [\dim(V)][\dim(W)].$$

Since $\dim(\mathbb{R}) = 1$, it follows that if V is finite dimensional, then

$$\dim(V^\dagger) = \dim(V).$$

More generally, if $V_1, V_2, V_3, ..., V_n$ are all finite dimensional vector spaces, then

$$\dim(V_1 \times V_2 \times V_3 \times ... \times V_n) = \dim(V_1) + \dim(V_2) + \dim(V_3) + ... + \dim(V_n).$$

In particular,

$$\dim(\mathbb{R}^n) = n = \dim(\mathbb{R}_n).$$

4.4. **STANDARD FRAMES.** For $V = \mathbb{R}^n$ we have the **Standard Frame** $F = (e_1, e_2, e_3, ..., e_n)$ where

$$e_1 = (1, 0, 0, 0, ..., 0)$$
$$e_2 = (0, 1, 0, 0, ..., 0)$$
$$e_3 = (0, 0, 1, 0, ..., 0)$$

and so on, so e_k has all zero entries except in slot k which contains a one, for each k, with $1 \leq k \leq n$. For this frame F we have $T_F = Id_V$ is just the identity map of $\mathbb{R}^n = V$, and thus $v_F = v^\dagger$ for each vector v in $\mathbb{R}^n$. Here $v^\dagger$ is the **Transpose** of v that is the vector in $\mathbb{R}^n$ expressed as a vertical list instead of a horizontal list or row.

We define the coordinate projections $e^k : \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$e^k(v_1, v_2, v_3, ..., v_n) = v_k.$$

Then e^k is linear so belongs to $L(\mathbb{R}^n; \mathbb{R}) = [\mathbb{R}^n]^\dagger$. Moreover, $(e^1, e^2, e^3, ..., e^n)$ is the dual frame for $[\mathbb{R}^n]^\dagger$ determined by the standard frame for $\mathbb{R}^n$, as obviously $e^i e_j$ is one or zero according to whether i and j are equal or not. The standard frame for $\mathbb{R}_n$ is likewise $(e_1^\dagger, e_2^\dagger, e_3^\dagger, ..., e_n^\dagger)$.

4.5. **VECTOR COMPONENTS RELATIVE TO A FRAME.** If v is a vector in V and if $F^\dagger = (b^1, b^2, v^3, ..., b^n)$ is the dual frame for $V^\dagger$, then the vector $T_F^{-1}v$ in $\mathbb{R}^n$ is the list of **Components** of v with respect to the frame F, and in fact

$$T_F^{-1}(v) = (b^1(v), b^2(v), b^3(v), ..., b^n(v)), \text{ where } F^\dagger = (b^1, b^2, b^3, ..., b^n)$$

It is customary here to set $v^k = b^k v$, so

$$T_F^{-1}(v) = (v^1, v^2, v^3, ..., v^n) \text{ in } \mathbb{R}^n.$$

Let v^F be the vertical list which is the transpose of the list $T_F^{-1}(v)$, for v in V,

$$v^F = [T_F^{-1}v]^\dagger.$$

Each vector v in V then has a unique expression as a linear combination of frame vectors, namely,

$$v = Fv^F = T_F(v^1, v^2, v^3, ..., v^n) = v^1 b_1 + v^2 b_2 + v^3 b_3 + ... + v^n b_n,$$

which can sometimes be very misleading for an unwary reader, as these superscripts are here not denoting exponential powers, rather they are simply tags just as are the subscripts. Notice the frame F is also defining a linear isomorphism $F : \mathbb{R}_n \longrightarrow V$ given by

$$F[w_1, w_2, w_3, ..., w_n]^\dagger = w_1 b_1 + w_2 b_2 + w_3 b_3 + ... + w_n b_n.$$

Notice that the transformation F is the inverse of the transformation which sends a vector v in V to the component list v^F in $\mathbb{R}_n$. Thus we can write

$$v^F = F^{-1}v, \text{ for } v \text{ in } V.$$

4.6. **LINEAR TRANSFORMATION MATRIX.** If V is a vector space with frame F and of W is a vector space with frame G where $\dim(V) = m$ and $\dim(W) = n$, and if $S : V \longrightarrow W$ is any linear transformation, then

$$M = G^{-1}SF : \mathbb{R}_m \longrightarrow \mathbb{R}_n$$

is a linear transformation completely determined by S and which in turn completely determines S. Any linear transformation is completely determined by its values on the frame vectors, thus S is completely determined as soon as we know $(Sv_1, Sv_2, Sv_3, ..., Sv_m)$ where $F = (b_1, b_2, b_3, ..., b_m)$. Likewise, M is completely determined by the list $Me_1^\dagger, Me_2^\dagger, Me_3^\dagger, ..., Me_m^\dagger$, so as $Me_k^\dagger$ is a vector in $\mathbb{R}_n$ for each k, it is customary to write the entries of Me_k in a vertical list so the list $(Me_1^\dagger, Me_2^\dagger, Me_3^\dagger, ..., Me_m^\dagger)$ becomes a horizontal list of vertical lists forming a square array of numbers called the **Matrix of** M, denoted $[M]$ which completely determines M and hence the matrix $[M]$ together with the frames F and G completely determine S, and vice-versa. Indeed,

$$S = GMF^{-1}.$$

We can denote this relationship by writing

$$[M] = [S_F^G], \text{ so } S_F^G = G^{-1}SF.$$

If $T : W \longrightarrow X$ is also a linear transformation of vector spaces, and if H is a frame for X, then matrix multiplication is defined by the rule

$$[(TS)_F^H] = [T_G^H][S_F^G].$$

Moreover, for v in V, we then have

$$[Sv]^G = [S_F^G][v^F]$$

so the matrix multiplication defined when applied to lists of components of vectors gives the computation of the linear transformation. Thus all computations in linear algebra are turned into matrix computations once frames are chosen. In addition, in terms of vector components, all the operations on vectors are simply the operations in $\mathbb{R}_n$ where n is the dimension of the vector space.

As to the matrix of the dual of a linear transformation, we have the very convenient fact that

$$[(T^\dagger)_{G^\dagger}^{F^\dagger}] = [T_F^G]^\dagger,$$

which is to say, the matrix of the dual transformation with respect to the dual frames is simply the transpose of the transformations's matrix with respect to the given frames.

To see in more detail, keep in mind that $[M]$ and M determine each other. Then,

$$[Sv]^G = G^{-1}Sv = G^{-1}SFv^F = S_F^G v^F = [S_F^G]v^F,$$

and

$$[(TS)_F^H]v^F = H^{-1}TSFvF = H^{-1}TGG^{-1}SFv^F = [T_G^H][S_F^G]v^F.$$

If we consider the relationship between $\mathbb{R}^n$ and $\mathbb{R}_n$, we can define the action of a row list on a column list via the rule

$$(r_1, r_2, r_3, ..., r_n)(s_1, s_2, s_3, ..., s_n)^\dagger = r_1 s_1 + r_2 s_2 + r_3 s_3 + ... + r_n s_n.$$

Notice that in this way, we can regard members of $\mathbb{R}_n$ as actually members of $\mathbb{R}^\dagger$, and in fact in this way we regard

$$\mathbb{R}_n = (\mathbb{R}^n)^\dagger \text{ and therefore } (bR_n)^\dagger = (\mathbb{R}^n)^{\dagger\dagger} = \mathbb{R}^n.$$

More generally, when we multiply matrices, in the product AB, the entry in row i and column j of the product, that is, $(AB)^i_j$, is simply $A^i B_j$, where we again recall the notation that A^i is row i of A and B_j is column j of B.

In terms of the frames using the notation

$$T_F^{-1} v = (v^1, v^2, v^3, ..., v^m)$$

we have

$$v = v^1 b_1 + v^2 b_2 + v^3 b_3 + ... + v^m b_m,\ v \text{ in } V$$

and if $G = (c_1, c_2, c_3, ..., c_n)$, in W^n, then

$$T_G^{-1} w = (w^1, w^2, w^3, ..., w^n),\ w \text{ in } W$$

so

$$w = w^1 c_1 + w^2 c_2 + w^3 c_3 + ... + w^n c_n,\ w \text{ in } W$$

and consequently

$$Sv = [Sv]^1 c_1 + [Sv]^2 c_2 + [Sv]^3 c_3 + ... + [Sv]^n c_n.$$

Then for $1 \le k \le n$ we have

$$[Sv]^k = [S(v^1 b_1 + v^2 b_2 + v^3 b_3 + ... + v^m b_m)]^k$$

$$= v^1 [Sb_1]^k + v^2 [Sb_2]^k + v^3 [Sb_3]^k + ... + v^m [Sb_m]^k.$$

Putting

$$S^k_i = [Sb_i]^k$$

we have

$$S^k_i = c^k S b_i,\ \text{where } G^* = (c^1, c^2, c^3, ..., c^n)$$

is the dual frame to G, and

$$[Sv]^k = S^k_1 v^1 + S^k_2 v^2 + S^k_3 v^3 + ... + S^k_m v^m.$$

This means that $[S^G_F]$ is the matrix array which has S^k_i in row k and column i, and also means that the multiplication of matrices is given by rows of the first factor multiplied by columns of the second factor.

5. BILINEAR MAPPINGS AND INNER PRODUCTS

5.1. **BILINEAR MAPPINGS.** If U, V, W are all vector spaces and $B : U \times V \longrightarrow W$, then holding u in U fixed we can define a function

$$Bu : V \longrightarrow W \text{ by the rule } Bu(v) = B(u, v).$$

On the other hand, if instead we hold v in V fixed, we can define a function

$$vB : U \longrightarrow W \text{ by the rule } vB(u) = B(u, v).$$

We say that B is **Bilinear** if Bu is linear for each u in U and vB is linear for each v in V.

Notation: if B is bilinear, we often abbreviate the function notation and write

$$Buv = B(u, v), \text{ for any vectors } u \text{ in } U \text{ and } v \text{ in } V.$$

Notice that the linearity in each slot is making Buv work like a multiplication of vectors as it is then both left and right distributive, so the B is like a coefficient for the multiplication. As there are many linear transformations, there are even more bilinear maps and consequently many ways to multiply vectors, where in this instance we are multiplying a vector in U by a vector in V to get a vector in W. In case $U = V$, then it makes sense to ask if $B(u, v) = B(v, u)$, and if this is always the case, then B is called **Symmetric**. Notice, then the multiplication Buv is commutative if B is symmetric. In this case we have

$$B[u \pm v][u \pm v] = Buu + Bvv \pm 2Buv$$

which gives the

GENERAL POLARIZATION IDENTITY:

$$Buv = \frac{1}{4}[B[u + v][u + v] - B[u - v][u - v]]$$

showing the any commutative multiplication is deterimined by squaring-if you know how to square anything then you can multiply any two things.

5.2. **SEMINORM AND VECTOR LENGTH.** If V is a vector space, we say that $n : V \longrightarrow \mathbb{R}$ is a **Seminorm** provided that

LENGTH SCALING RULE

$$n(rv) = |r|n(v), \text{ for every scalar } r \text{ and every vector } v,$$

and

TRIANGLE INEQUALITY

$$n(v + w) \leq n(v) + n(w), \text{ for all vectors } v, w$$

both hold.

Notation: we write

$$\|v\|_n = n(v), \text{ for every vector } v.$$

When there is no confusion as to n, the subscript is dropped here. These two properties are the minimum that should be expected for properties of length of a vector. The first property in particular implies that the zero vector has zero length.

5.3. **NORM AND LENGTH.** Notice that with a seminorm it is possible for a non-zero vector to have zero length. The seminorm is called a **Norm** if the only vector with zero length is the zero vector.

5.4. **INNER PRODUCTS.** If V is a vector space, we say that B is an **Inner Product** for V provided that

$$B : V \times V \longrightarrow \mathbb{R} \text{ is a symmetric bilinear map.}$$

In this case we write

$$\langle u|v\rangle_B = Buv$$

and if B is understood on $V \times V$, we write $Buv = \langle u|v\rangle_V$. If both B and V are understood, we drop the subscript on the angle bracket altogether. We say the inner product is **Positive** if

$$\langle v|v\rangle \geq 0, \text{ for every vector } v.$$

We can notice that using the symmetry and bilinearity that always

$$\langle v \pm w|v \pm w\rangle = \langle v|v\rangle + \langle w|w\rangle \pm 2\langle v|w\rangle$$

and therefore we again have the

POLARIZATION IDENTITY FOR INNER PRODUCTS

$$\langle v|w\rangle = \frac{1}{4}[\langle v + w|v + w\rangle - \langle v - w|v - w\rangle]$$

If V is finite dimensional, then we can find linear subspaces V_+ and V_- such that B is positive on V_+ and $-B$ is positive on V_- with

$$V = V_+ + V_- \text{ and } V_+ \cap V_- = 0_V.$$

5.5. **POSITIVE INNER PRODUCTS AND LENGTH.** If V is vector space with a positive inner product then we can try to define the length of a vector by the rule

$$\|v\| = \sqrt{\langle v|v\rangle}.$$

Since this is equivalent to

$$\|v\|^2 = \langle v|v\rangle$$

we would have easily that the length scaling rule holds but the triangle inequality depends on the

CAUCHY-SCHWARZ INEQUALITY:

$$|\langle v|w\rangle| \leq \|v\|\|w\|, \text{ for all vectors } v, w.$$

To see that the Cauchy-Scwarz Inequality must hold for the case of a positive inner product, we can use the fact that for any real numbers s, t and any vectors v, w, it must be the case that

$$0 \leq \langle sv - tw|sv - tw\rangle = s^2\langle v|v\rangle + t^2\langle w|w\rangle - 2st\langle v|w\rangle$$

and therefore

$$2st\langle v|w\rangle \leq s^2\langle v|v\rangle + t^2\langle w|w\rangle = s^2\|v\|^2 + t^2\|w\|^2.$$

If either $\langle v|v\rangle = \|v\|^2$ or $\langle w|w\rangle = \|w\|^2 = 0$, then the preceding inequality valid for all real numbers s, t will guarantee that $\langle v|w\rangle = 0$, whereas if both are not zero, then choose

$$s = \frac{1}{\|v\|} \text{ and } \frac{1}{\|w\|},$$

and the inequality becomes

$$2\frac{\langle v|w\rangle}{\|v\|\|w\|} = 2st\langle v|w\rangle \leq 2,$$

from which the Cauchy-Schwarz Inequality follows immediately.

5.6. **RIESZ REPRESENTATION.** If v is a vector in the vector space with inner product, V, then we can use it to define a linear map $v^\dagger$ in the dual vector space $V^\dagger = L(V;\mathbb{R})$ by the rule $v^*(w) = \langle v|w\rangle_V$. Since the inner product is bilinear, it follows by definition, that v^* is linear, so in fact, v^* belongs to $V^\dagger$. We can define the linear map $RZ_V : V \longrightarrow V^\dagger$ by the rule $RZ_V v = v^*$. If the inner product is positive definite, and if $v^* = 0$, then $\|v\|^2 = \langle v|v\rangle = v^*v = 0$, and therefore $v = 0$. Thus, in case of a positive definite inner product, $RZ_V : V \longrightarrow V^\dagger$ is injective. More generally, if the inner product is not positive, we say that it is **Non-degenerate** if RZ_V is injective. If λ belongs to $V^\dagger$, we say that λ is **Representable** if λ is in the image of RZ_V, which is to say there is a vector v with $v^* = \lambda$. The Riesz Representation Theorem guarantees that under certain technical conditions, every member of $V^\dagger$ is representable, and therefore RZ_V is a vector space isomorphism of V onto $V^\dagger$. In particular, if V is finite dimensional and the inner product is non-degenerate, then every member or $V^\dagger$ is representable and $RZ_V : V \longrightarrow V^*$ is an isomorphism, since $V^\dagger$ has the same finite dimension as V.

5.7. **CARTESIAN PRODUCT OF INNER PRODUCT VECTOR SPACES.** If $V_1, V_2, V_3, ..., V_m$ are all vector spaces with inner products, V_k having inner product B_k, then we can give the cartesian product

$$V = V_1 \times V_2 \times V_3 \times ... \times V_m$$

the inner product B where

$$\langle (v_1, v_2, v_3, ..., v_m) | (w_1, w_2, w_3, ..., w_m) \rangle_B = \langle v_1 | w_1 \rangle_{B_1} + \langle v_2 | w_2 \rangle_{B_2} + \langle v_3 | w_3 \rangle_{B_3} + ... + \langle v_m | w_m \rangle_{B_m}.$$

The ordinary multiplication of numbers serves as the inner product in $\mathbb{R}$. Thus, $\mathbb{R}^m$ is given the resulting inner product as the cartesian product, where $V_k = \mathbb{R}$ for each $k \leq m$.

5.8. **ORTHONORMAL FRAMES.** A frame $E = (e_1, e_2, e_3, ..., e_m)$ for the vector space V with inner product is said to be **Orthonormal** provided that $|\langle e_i | e_j \rangle|$ is one or zero according to whether $i = j$. If the inner product is non-degenerate and if V is finite dimensional, then there is an orthonormal frame and every vector v in V is easily expressed as a linear combination of these frame vectors using

$$v = \langle v | e_1 \rangle \langle e_1 | e_1 \rangle e_1 + \langle v | e_2 \rangle \langle e_2 | e_2 \rangle e_2 + \langle v | e_3 \rangle \langle e_3 | e_3 \rangle e_3 + ... + \langle v | e_m \rangle \langle e_m | e_m \rangle e_m.$$

Moreover, the dual frame E^* for V^* dual to E is given by

$$E^\dagger = (\langle e_1 | e_1 \rangle e_1^*, \langle e_2 | e_2 \rangle e_2^*, \langle e_3 | e_3 \rangle e_3^*, ..., \langle e_m | e_m \rangle e_m^*).$$

That is to say, $E^\dagger = (e^1, e^2, e^3, ..., e^m)$ where

$$e^k = \langle e_k | e_k \rangle e_k^*.$$

We can easily see that the standard frames for bR^m and $\mathbb{R}_m$ are the duals of each other and are orthonormal.

If $T : V \longrightarrow W$ is a linear transformation of vectors spaces having inner products and if $E = (e_1, e_2, e_3, ..., e_m)$ is an orthonormal frame for V and $G = (g_1, g_2, g_3, ..., g_n)$ is an orthonormal frame for W, then the matrix of T relative to these frames has $\langle g_i | g_i \rangle \langle g_i | Tej \rangle$ in row i and column j.

If f belongs to $[\mathbb{R}^m]^\dagger$, then taking E to be the standard frame for $\mathbb{R}^m$, $(fe_1, fe_2, fe_3, ..., fe_m)$ belongs to $\mathbb{R}^m$, and if $v = (fe_1, fe_2, fe_3, ..., fe_m)$, then it is easy to see that $f = v^*$, giving direct proof of the Riesz Representation Theorem in case of finite dimensions, so

$$[RZ_{\mathbb{R}^m}]^{-1} f = (fe_1, fe_2, fe_3, ..., fe_m).$$

We also have the notations commonly used in calculus here,

$$[RZ_{\mathbb{R}^m}]^{-1} f = \text{ grad } f = \nabla f.$$

5.9. **ADJOINT OF A LINEAR TRANSFORMATION.** If V and W are vector spaces with inner product and if $R : V \longrightarrow W$ and $S : W \longrightarrow V$, then we say that R and S are **Mutually Adjoint** if

$$\langle Rv|w\rangle_W = \langle v|Sw\rangle_V, \text{ for every } v \text{ in } V \text{ and every } w \text{ in } W.$$

Notice this is saying

$$w^* Rv = [Sw]^* v, \text{ for every } v \text{ in } V \text{ and every } w \text{ in } W,$$

or

$$([R^\dagger w^*)v = [Sw]^* v, \text{ for every } v \text{ in } V \text{ and every } w \text{ in } W,$$

and therefore

$$([R^\dagger]w^*) = [Sw]^*, \text{ for every } w \text{ in } W,$$

which is the same as

$$([R^\dagger]RZ_W w) = RZ_V[Sw], \text{ for every } w \text{ in } W,$$

and this in turn is equivalent to

$$[R^\dagger]RZ_W = [RZ_V]S.$$

In particular, any time that the Riesz Representation Theorem holds for V, that is, any time that $RZ_V : V \longrightarrow V^*$ is an isomorphism, it follows that S is determined by R with the formula

$$S = [RZ_V]^{-1}[R^\dagger]RZ_W.$$

For this reason, we define the **Adjoint** of R denoted by R^* by the formula

$$R^* = [RZ_V]^{-1}[R^\dagger]RZ_W,$$

and we say that R has an adjoint.

Thus, whenever $R : V \longrightarrow W$ has an adjoint, which we see must be the case if V is finite dimensional with a non-degenerate inner product, then we have $R^* : W \longrightarrow V$, is linear and

$$\langle Rv|w\rangle_W = \langle v|R^* w\rangle_V, \text{ for every } v \text{ in } V \text{ and every } w \text{ in } W.$$

5.10. **PROPERTIES OF ADJOINTS.** From the equation

$$\langle Rv|w\rangle_W = \langle v|R^*w\rangle_V, \text{ for every } v \text{ in } V \text{ and every } w \text{ in } W,$$

we see clearly that if R has an adjoint, then so does its adjoint R^*, and

$$R^{**} = R,$$

moreover, the same formula also makes it is easy to see that if $R : V \longrightarrow W$ and $Q : U \longrightarrow V$ both have adjoints, then so does RQ and we have

$$[RQ]^* = Q^*R^*.$$

It is also easy to see that if R and S both belong to $L(V;W)$ and have adjoints, then so does $R + tS$ and

$$[R + tS]^* = R^* + tS^*, \text{ for every } t \text{ in } \mathbb{R}.$$

Thus, when all members of $L(U;V)$ have adjoints then the transformation

$$Adj : L(U;V) \longrightarrow L(V;U)$$

given by

$$[Adj]R = R^*, \text{ for every } R \text{ in } L(U;V)$$

is a linear mapping with $(Adj)(Adj) = Id_{L(U;V)}$, so Adj is a vector space isomorphism which, in case $U = V$, is its own inverse.

If $T : V \longrightarrow W$ is a linear transformation and if RZ_V and RZ_W are both isomorphisms, then every linear transformation in $L(V;W)$ has an adjoint and every linear transformation in $L(W;V)$ has an adjoint, and if T is invertible, then so is T^* and

$$[T^{-1}]^* = [T^*]^{-1}.$$

5.11. **THE MATRIX OF THE ADJOINT.** If $R : V \longrightarrow W$ is a linear transformation and $B = (b_1, b_2, b_3, ..., b_m)$ is an orthonormal frame for V and if $C = (c_1, c_2, c_3, ..., c_n)$ is an orthonormal frame for W, then the matrix of R relative to these frames is $[R_B^C]$ having entry $\langle c_i|Rb_j\rangle_W$ in row i and column j. But, from the fact that

$$\langle Rv|w\rangle_W = \langle v|R^*w\rangle_V, \text{ for every } v \text{ in } V \text{ and every } w \text{ in } W,$$

it follows that

$$\langle c_i|Rb_j\rangle_W = \langle b_j|R^*c_i\rangle, \text{ for every } i \leq n \text{ and every } j \leq m.$$

This means that the matrix of R^* is the transpose of the matrix of R, that is

$$[(R^*)_C^B] = [R_B^C]^\dagger.$$

Of course, this means that the adjoint is very easy to compute using matrices.

Department of Mathematics, Tulane University, New Orleans, LA 70118 USA
E-mail address, M. J. Dupré: mdupre@tulane.edu