

## Bases and Coordinates

**Definition 1.** Let  $V$  be a vector space over the field  $F$ . Let  $S \subseteq V$  be a subset. The set  $S$  is said to be *linearly dependent over  $F$*  if there exist distinct  $v_1, \dots, v_m \in S$  and  $a_1, \dots, a_m \in F$ , with not all  $a_i = 0$ ,  $m \geq 1$  so that

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0. \quad (1)$$

This equation is called a *dependence relation*. If no such dependence relation exists, the set  $S$  is called *linearly independent over  $F$* .

If equation (1) holds for distinct elements  $v_i$  of a linearly independent set, then all  $a_i$  must be 0.

The following properties are clear:

1. If  $S \subseteq T$ , then  $S$  linearly dependent implies that  $T$  is linearly dependent. Similarly, if  $T$  is linearly independent, then  $S$  is linearly independent.
2. The set  $\{0\}$  is linearly dependent:  $1 \cdot 0 = 0$  being the required dependence relation. Thus any set  $S$  containing 0 is linearly dependent.
3. The empty set  $\phi$  is linearly independent.
4. A set  $S$  is linearly independent if and only if every finite subset of  $S$  is linearly independent.

**Definition 2.** A subset  $\mathcal{B} \subseteq V$  is a *basis* for  $V$  over  $F$  if the following hold:

1.  $\mathcal{B}$  is linearly independent over  $F$ ,
2.  $\text{Span}_F(\mathcal{B}) = V$ .

The vector space  $V$  is called *finite dimensional* if it contains a finite subset  $S$  with  $\text{Span}_F(S) = V$ .

We now introduce notation that will simplify computations, particularly for theoretical arguments. By doing this we manage to avoid awkward arguments where we must renumber lists of vectors which change during the discussion when new ones appear, or might appear. Let  $\alpha_b \in F$  for  $b \in \mathcal{B}$  be a collection of elements of  $F$  with only finitely many being non-zero. We then obtain a vector in  $V$  by taking the finite sum of the  $\alpha_b b$  for those  $\alpha_b$  which are non-zero. We abbreviate this as

$$\sum_{b \in \mathcal{B}} \alpha_b b. \quad (2)$$

That is, although this may appear to be an infinite sum (if  $\mathcal{B}$  is infinite) it is actually a finite sum since only finitely many terms matter.

Our first lemma will give the primary tool for using bases.

**Lemma 3.** *Let  $\mathcal{B}$  be a basis for the vector space  $V$  over the field  $F$ . Then every vector in  $V$  can be written uniquely as a finite linear combination of elements of  $\mathcal{B}$ .*

*Proof.* We must do two things: first show that for any vector  $v \in V$  such an expression is possible, and second, show that there is only one such expression.

Since  $\text{Span}_F(\mathcal{B}) = V$ , given  $v \in V$  there exist  $\alpha_b \in F$ , only finitely many non-zero, so that

$$v = \sum_{b \in \mathcal{B}} \alpha_b b .$$

In order to prove the uniqueness of the expression, we assume that we are given a second expression for  $v$ :

$$v = \sum_{b \in \mathcal{B}} \beta_b b .$$

Subtraction of the second expression from the first yields

$$0 = \sum_{b \in \mathcal{B}} (\alpha_b - \beta_b) b .$$

However,  $\mathcal{B}$  is linearly independent and hence all coefficients of the last equation are 0, that is  $\alpha_b = \beta_b$  for all  $b \in \mathcal{B}$ .  $\square$

We now give a few examples of bases (see the section on “Examples of Vector Spaces”).

1. For  $1 \leq i \leq n$  let  $e_i = (0, \dots, 1, \dots, 0)$  be the vector in  $F^n$  with 1 in the  $i$ -th position and 0 elsewhere and let  $\mathcal{B} = \{e_1, \dots, e_n\}$ . Then  $\mathcal{B}$  is a basis for  $F^n$  called the *standard basis*. Consider the equation

$$(a_1, \dots, a_n) = a_1 e_1 + \dots + a_n e_n .$$

Reading from left to right gives the formula to show that  $\mathcal{B}$  spans; reading from right to left shows that if we assume the linear combination is the 0 vector of  $F^n$ , then all of the  $a_i$  must be 0, giving linear independence.

2. For  $1 \leq i \leq m$  and  $1 \leq j \leq n$  let  $e_{i,j}$  be the matrix in  $F^{m \times n}$  with 1 in position  $(i, j)$  and 0 elsewhere. If  $A \in F^{m \times n}$  denotes the matrix with entries  $a_{i,j} \in F$  in positions  $(i, j)$ , then we have

$$A = \sum_{i,j} a_{i,j} e_{i,j} .$$

As in the previous example, we conclude that  $\mathcal{B} = \{e_{i,j} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$  is a basis, also called the *standard basis*.

3. Let  $S$  be a finite, non-empty set and let  $F^S$  be the vector space of all functions from  $S$  to  $F$ . For  $s \in S$ , define the function  $\delta_s : S \rightarrow F$  by

$$\delta_s(t) = \begin{cases} 0 & t \neq s \\ 1 & t = s \end{cases}$$

We then have for  $f \in F^S$  the formula

$$f = \sum_{s \in S} f(s)\delta_s .$$

This is easy to verify: given  $t \in S$  the left-hand side is just  $f(t)$  while the right-hand side has 0 for all terms where  $t \neq s$ ; for the term where  $t = s$ , it has  $f(t) \cdot 1$ . And as in the previous examples, this one equation implies that  $\mathcal{B} = \{\delta_s \mid s \in S\}$  is a basis for  $F^S$ . WARNING: If  $S$  is infinite, the  $\delta_s$  do not form a basis of  $F^S$ . They in fact span the subspace  $F^{(S)}$  and form a basis for it. Further, the dimension of  $F^S$  over  $F$  is strictly larger than  $|S|$ , but that's a bit harder to prove.

4. For our last example, consider  $F[x]$ , the formal polynomials with coefficients in  $F$ , and let  $\mathcal{B} = \{1, x, x^2, x^3, \dots, x^i, \dots\}$ . This is a basis:  $\mathcal{B}$  spans as every polynomial is a finite linear combination of the monomials  $x^i$ , by definition;  $\mathcal{B}$  is linearly independent since the representation of a (formal) polynomial is unique, by definition.

**Remark 4.** The first three examples are really all the “same”, or at least the same idea is involved. For  $F^n$  one can think of a vector  $(a_1, \dots, a_n)$  as a function  $a : \{1, \dots, n\} \rightarrow F$ , with  $a(i) = a_i$ . Similarly, for  $A$  with entries  $a_{i,j}$ , one can think of  $a$  as a function on the set of pairs  $\{(i, j) \mid 1 \leq i \leq m, 1 \leq j \leq n\}$  with values in  $F$  given by  $a(i, j) = a_{i,j}$ . And of course the third example is such a set of functions. It should be clear how to give isomorphisms between such examples. In any case, we'll do this explicitly following our discussion of the Universal Mapping Property for Bases in what follows.

Two sets  $X$  and  $Y$  are said to be the same “size”, or more precisely to have the same *cardinality* if there exists a one-to-one, onto function  $f : X \rightarrow Y$ . It is easy to check that cardinality is reflexive (the identity function), symmetric (use the inverse), and transitive (use composition). One obtains thereby the analogue of an equivalence relation on the class of all sets. We write  $|X|$  to denote the cardinality of the set  $X$ ; so we write  $|X| = |Y|$  to mean that they have the same cardinality. There is an “arithmetic” for cardinal numbers which we will not develop here. However, we will informally use this notion in our discussions and proofs – we will be careful to state all results on dimension in a manner so that they either remain valid for vector spaces of infinite dimension, or we will add the assumption “finite dimension” to the statements.

In addition we will say that the cardinality of the set  $X$  is less than or equal to the cardinality of the set  $Y$  if there exists a one-to-one function  $f : X \rightarrow Y$ , and we

denote this by  $|X| \leq |Y|$ . An equivalent condition is that there is an onto function  $g : Y \rightarrow X$ . The Schroeder-Bernstein Theorem (sometimes the name Cantor is also added) asserts that  $|X| \leq |Y|$  and  $|Y| \leq |X|$  imply that  $|X| = |Y|$  (that is, the existence of two one-to-one functions going in opposite directions implies the existence of a one-to-one, onto function). The proof turns out to be surprisingly easy. See the section on the “Axiom of Choice and Zorn’s Lemma” for further information.

The next result is the main theorem about bases we use. A proof for the finite-dimensional case uses the theory behind the solutions of homogeneous systems of equations and is given below. The proof in the general case requires the use of the Axiom of Choice, usually in the form of Zorn’s Lemma. One may think of this as a way (analogous to mathematical induction) to handle very large sets. See the section on the “Axiom of Choice and Zorn’s Lemma” for further explanation and proof of the following theorem.

**Theorem 5** (Main Theorem for Bases). *Let  $V$  be a vector space over the field  $F$ .*

1. *There exists a basis for  $V$ .*
2. *If the subset  $S \subseteq V$  spans  $V$ , then  $S$  contains a basis for  $V$ .*
3. *If the subset  $S \subseteq V$  is linearly independent, then  $S$  is contained in some basis for  $V$ .*
4. *If  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are two bases for  $V$ , then  $|\mathcal{B}_1| = |\mathcal{B}_2|$ .*

**Corollary 6.** *Let  $V$  be a vector space over the field  $F$ . If  $I \subseteq V$  is a linearly independent subset and  $S \subseteq V$  is a subset which spans  $V$  over  $F$ , then*

$$|I| \leq |S|$$

(that is, there exists a one-to-one function  $h : I \rightarrow S$ ).

*Proof.* By the third part of the Main Theorem, there exists a basis  $\mathcal{B}_1$  of  $V$  with  $I \subseteq \mathcal{B}_1$ . By the second part of the Main Theorem there exists a basis  $\mathcal{B}_2$  of  $V$  with  $\mathcal{B}_2 \subseteq S$ . By the fourth part of the Main Theorem, there exists a one-to-one, onto function  $f : \mathcal{B}_1 \rightarrow \mathcal{B}_2$ . Hence, let  $h : I \rightarrow S$  be the composition

$$I \xrightarrow{\text{inc}} \mathcal{B}_1 \xrightarrow{f} \mathcal{B}_2 \xrightarrow{\text{inc}} S$$

where inc denotes the inclusion function in both cases. That is,  $h = \text{inc} \circ f \circ \text{inc}$  is a one-to-one function since it is the composition of three one-to-one functions. Equivalently,

$$|I| \leq |S| .$$

□

**Definition 7.** Let  $V$  be a vector space over the field  $F$  and let  $W$  be a subspace of  $V$ . A subspace  $W'$  of  $V$  is called a *complement* of  $W$  if  $V$  is the internal direct sum of  $W$  and  $W'$ .

**Corollary 8.** *Let  $V$  be a vector space over the field  $F$  and let  $W$  be a subspace of  $V$ . Then a complement of  $W$  in  $V$  exists.*

*Proof.* Let  $\mathcal{A}$  be a basis for  $W$ . Clearly  $\mathcal{A}$  is a linearly independent subset of  $V$  and hence by the third part of the Main Theorem, there exists a basis  $\mathcal{B}$  for  $V$  containing  $\mathcal{A}$ . Let  $\mathcal{A}'$  be the set  $\mathcal{B}$  with  $\mathcal{A}$  removed. Then  $\mathcal{B}$  is the disjoint union of  $\mathcal{A}$  and  $\mathcal{A}'$ . Let  $W' = \text{Span}_F(\mathcal{A}')$ . Clearly  $W + W' = V$  as every element of a basis of  $V$  is contained in  $W + W'$ . Further  $W \cap W' = 0$  since a vector  $v$  in the intersection can be expressed as a linear combination of elements in  $\mathcal{A}$  and similarly as a linear combination of elements in  $\mathcal{A}'$ . The difference of these two expressions is 0 and as it is a linear combination of elements of the linearly independent set  $\mathcal{B}$ , all coefficients must be 0.  $\square$

**Remark 9.** If  $W \subseteq V$  is a proper subspace (i.e., neither 0 nor  $V$ ), then  $W$  will have more than one complement. For example, in  $V = \mathbb{R}^2$ , if  $W$  is the span of a non-zero vector, then the span of any vector not in  $W$  will be a complement to  $W$ .

**Corollary 10.** *Every short exact sequence of vector spaces splits.*

*Proof.* This is an easy consequence of the preceding corollary, [See the section on “Short Exact Sequences” for the terminology, and for exercises related to these two corollaries.]  $\square$

We now give a proof of the Main Theorem for Bases (Theorem 5) in the case where the vector space  $V$  is finite-dimensional.

*Proof.* The first part of the theorem (that bases exist) is a consequence of the second part, that there is a subset of  $S$  which is a basis. Similarly, the first part is also a consequence of the third part since one can start with the linearly independent set  $\emptyset$ , the empty set.

As this vector space  $V$  is finite dimensional, there is a finite subset  $S \subseteq V$  which spans  $V$  over  $F$ .

We now prove that there is a subset of  $S$  which is a basis for  $V$ . If  $S$  is linearly independent, the  $S$  itself is a basis since we already know it spans  $V$ . If  $S = \{v_1, \dots, v_n\}$  is not linearly independent, then there will exist  $a_i \in F$ , not all 0, which give a dependence relation:

$$a_1 v_1 + \dots + a_n v_n = 0.$$

Since not all of the coefficients are 0, there will be some  $j$  with  $a_j \neq 0$ . Thus

$$v_j = \frac{-a_1}{a_j} v_1 + \dots + \frac{-a_n}{a_j} v_n$$

where the  $j$ -th term has been omitted from the right-hand side. Thus the smaller set  $S_1 = S \setminus \{v_j\}$  will still span  $V$ . If  $S_1$  is linearly independent, then it forms a basis

for  $V$ , and if not, some vector in  $S_1$  will be a linear combination of the remaining ones. Then we can construct a strictly smaller spanning set  $S_2$  by removing that vector as well. Continuing this process we will end up with a basis for  $V$  in at most  $|S|$  steps.

Let  $I = \{u_1, \dots, u_m\}$  be a linearly independent subset of  $V$ . If  $I$  spans  $V$ , then  $I$  itself is a basis and we are done. If it does not span  $V$ , then there will be some vector  $v_j \in S$  which is not contained in  $\text{Span}_F(I)$ . Let  $I_1 = I \cup \{v_j\}$ . Then this set  $I_1$  is strictly larger, and is in fact linearly independent. For if not, there exist some  $b_i, a \in F$ , not all 0, with

$$b_1 u_1 + \dots + b_m u_m + a v_j = 0.$$

Now if  $a \neq 0$ , then solving the equation shows that  $v_j \in \text{Span}_F(I)$ , a contradiction. Thus  $a = 0$ . But then all  $b_k = 0$  since  $I$  is a linearly independent set. This contradiction shows that in fact  $I_1$  must be linearly independent (and hence is strictly larger than  $I$ ). If  $I_1$  spans  $V$ , then it is a basis, and we're done. If not, then there is yet another element of  $S$  which did not lie in  $\text{Span}_F(I)$ . We let  $I_2$  be the set obtained as the union of  $I_1$  and this additional vector of  $S$  which was not in the span. The same argument as before shows that  $I_2$  is linearly independent, and a strictly larger set. If it spans  $V$  we are done. If not, we continue in the same fashion. After at most  $|S|$  steps the process must cease since there will be no more vectors from  $S$  to add, and hence none that are not in the span of the just constructed set.

Finally, we will verify that any two bases for a finite dimensional  $V$  must have the same number of elements. Let  $I = \{u_1, \dots, u_m\}$  be a linearly independent subset of  $V$  and let  $S = \{v_1, \dots, v_n\}$  be a spanning set. We now show that  $m = |I| \leq |S| = n$ , as stated in Corollary 6 for the case of arbitrary  $V$ . Since  $S$  spans, for each  $u_j \in I$  there will exist  $a_{ij} \in F$  such that

$$u_j = a_{1j} v_1 + \dots + a_{nj} v_n.$$

Let  $A = (a_{ij})$  be the  $n \times m$  matrix with entries the  $a_{ij}$ . If  $n < m$ , then the matrix equation  $AX = 0$  would have a non-trivial solution, say with entries  $b_1, \dots, b_m$  (not all 0) so that

$$b_1 u_1 + \dots + b_m u_m = 0$$

which is impossible as the set  $I$  is linearly independent. Thus  $n \geq m$  must hold, that is  $|S| \geq |I|$  as asserted.

Now let  $\mathcal{B}_1$  and  $\mathcal{B}_2$  be two bases for  $V$ . As  $\mathcal{B}_1$  is linearly independent and  $\mathcal{B}_2$  spans, we then have  $|\mathcal{B}_1| \leq |\mathcal{B}_2|$ . Switching the roles of the two bases yields the inequality in the other direction, and thus we must have  $|\mathcal{B}_1| = |\mathcal{B}_2|$ .  $\square$

For a different proof that can be generalized to the infinite dimensional case see Exercise 19.

In our argument we proved the following lemma that is useful in many situations:

**Lemma 11.** *Let  $V$  be a vector space over the field  $F$ . If  $I \subseteq V$  is a linearly independent subset and  $v \in V$  is not in  $\text{Span}_F(I)$ , then the set  $I \cup \{v\}$  is linearly independent.*

## Coordinates

Our first application of Theorem 5 will be via Equation (2) to obtain the *coordinates* of a vector with respect to a basis.

Let  $\mathcal{B} \subseteq V$  be a basis for the vector space  $V$  over  $F$ . Lemma 3 asserts that for a vector  $v \in V$  the coefficients  $\alpha_b \in F$  in Equation 2 are uniquely determined:

$$v = \sum_{b \in \mathcal{B}} \alpha_b b .$$

Hence we obtain a function from  $\mathcal{B}$  to  $F$ , which we denote by  $[v]_{\mathcal{B}}$ , that is  $[v]_{\mathcal{B}}$  evaluated at  $b \in \mathcal{B}$  is just the  $\alpha_b$  in this equation.

Recall (see the section on “Direct Sums and Products”) that the direct sum of  $\mathcal{B}$  copies of  $F$ , denoted  $\bigoplus_{b \in \mathcal{B}} F$ , is just the set of functions from  $\mathcal{B}$  to  $F$ , which are non-zero for at most finitely many elements of  $\mathcal{B}$ . This is also the same as  $F^{(\mathcal{B})}$ . Hence we obtain a function

$$[\ ]_{\mathcal{B}} : V \longrightarrow \bigoplus_{b \in \mathcal{B}} F = F^{(\mathcal{B})} \quad (3)$$

which when evaluated at  $v \in V$ , gives the function  $[v]_{\mathcal{B}}$ .

To restate explicitly, we have a function whose values are also functions:

$[\ ]_{\mathcal{B}}$  sends  $v$  to the function  $[v]_{\mathcal{B}}$

$[v]_{\mathcal{B}}$  is the function that sends  $b \in \mathcal{B}$  to  $\alpha_b$

– the last statement, using ordinary notation for functions, is just  $[v]_{\mathcal{B}}(b) = \alpha_b$ .

**Theorem 12** (Coordinates with Respect to a Basis). *Let  $V$  be a vector space over the field  $F$  with basis  $\mathcal{B}$ . Taking coordinates with respect to the basis  $\mathcal{B}$  gives an isomorphism of  $V$  with the direct sum of  $\mathcal{B}$  copies of  $F$ ; that is*

$$[\ ]_{\mathcal{B}} : V \longrightarrow \bigoplus_{b \in \mathcal{B}} F = F^{(\mathcal{B})}$$

*is an isomorphism of vector spaces.*

*Proof.* Let  $v$  and  $w$  be two vectors in  $V$  and express them as linear combinations of the basis:

$$\begin{aligned} v &= \sum_{b \in \mathcal{B}} \alpha_b b \\ w &= \sum_{b \in \mathcal{B}} \beta_b b . \end{aligned}$$

Adding the two yields

$$v + w = \sum_{b \in \mathcal{B}} (\alpha_b + \beta_b) b$$

and thus for any  $b \in \mathcal{B}$

$$\begin{aligned} [v + w]_{\mathcal{B}}(b) &= \alpha_b + \beta_b \\ &= [v]_{\mathcal{B}}(b) + [w]_{\mathcal{B}}(b) \\ &= ([v]_{\mathcal{B}} + [w]_{\mathcal{B}})(b) . \end{aligned}$$

Hence  $[v + w]_{\mathcal{B}} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}$ , since two functions are equal precisely when they have the same value for each element of the domain  $\mathcal{B}$ .

Similarly for  $c \in F$ ,

$$cv = \sum_{b \in \mathcal{B}} (c\alpha_b) b$$

and thus

$$\begin{aligned} [cv]_{\mathcal{B}}(b) &= c\alpha_b \\ &= c([v]_{\mathcal{B}}(b)) \\ &= (c[v]_{\mathcal{B}})(b) . \end{aligned}$$

Hence the two functions are equal:  $[cv]_{\mathcal{B}} = c[v]_{\mathcal{B}}$ .

We've now verified that  $[\ ]_{\mathcal{B}}$  is a linear transformation.

The function  $[\ ]_{\mathcal{B}}$  is one-to-one since a vector  $v$  is determined by its coordinates  $\alpha_b$ . Finally,  $[\ ]_{\mathcal{B}}$  is onto for a similar reason; given a function  $f \in \bigoplus_{b \in \mathcal{B}} F$ , the vector  $u \in V$  given by the finite sum

$$u = \sum_{b \in \mathcal{B}} f(b) b$$

will satisfy  $[u]_{\mathcal{B}} = f$  since the two functions do the same thing on elements of  $\mathcal{B}$ .  $\square$

Up to this point coordinates with respect to a basis have been treated abstractly, which is a good way to do it if one is interested in proving things. However, sometimes one wants to make concrete computations, and even write them on the blackboard or a piece of paper. How does one write down a function on a basis? One very inefficient way to do it would be to write down the basis element in one column and next to it the value of the function. As we write on a two-dimensional surface, this means in addition we have to make a choice as to the order we write things down and we also need to list the basis elements. This is usually abbreviated by deciding in advance the order we pick, fixing that order throughout the discussion, and taking advantage of the geometry of a two-dimensional surface to indicate the order (as well as our own internal concept of 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, ...).

We will restrict our attention now to finite dimensional vector spaces. An *ordered basis*  $\mathcal{B}$  for  $V$  is simply a basis where we have chosen in advance an order for the elements:  $\mathcal{B} = \{b_1, \dots, b_n\}$ . We will now use the same notation as used above to

denote the coordinates of a vector with respect to an ordered basis  $\mathcal{B}$ . If  $v \in V$  can be written as  $v = \sum_{i=1}^n \alpha_i b_i$  we will write  $[v]_{\mathcal{B}} \in F^{n \times 1}$  as follows:

$$[v]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

We restate the finite dimensional case of Theorem 12:

**Corollary 13** (Coordinates with Respect to a Basis). *Let  $V$  be a finite dimensional vector space over the field  $F$  with ordered basis  $\mathcal{B}$ . Then taking coordinates with respect to the ordered basis  $\mathcal{B}$  gives an isomorphism of  $V$  with  $F^{n \times 1}$ ; that is*

$$[\ ]_{\mathcal{B}} : V \longrightarrow F^{n \times 1}$$

*is an isomorphism of vector spaces.*

## Universal Mapping Property for Bases

We now give the usual theorem which describes all linear transformations in terms of a given basis for a vector space stated as a *universal mapping property*.

**Theorem 14** (Universal Mapping Property for Bases). *Let  $U$  and  $V$  be vector spaces over  $F$ . Let  $\mathcal{B} \subseteq V$  be a basis for  $V$  and let  $i : \mathcal{B} \rightarrow V$  be the inclusion map. For any function  $t : \mathcal{B} \rightarrow U$ , there exists a unique linear transformation  $T : V \rightarrow U$  such that the following diagram commutes:*

$$\begin{array}{ccc} \mathcal{B} & \xrightarrow{i} & V \\ & \searrow t & \vdots \\ & & U \end{array}$$

*that is,  $T \circ i = t$ .*

*Proof.* In outline, in almost all cases, proofs of universal mapping properties take the following form: first show uniqueness, next use the result of uniqueness (typically a formula) to show existence of the sought-after function, and finally verify that the function just constructed has all of the right properties.

**Uniqueness:** Our assumption means that  $T \circ i(b) = t(b)$  or  $T(b) = t(b)$  for any  $b \in \mathcal{B}$ . Given  $v \in V$  we can write  $v = \sum_{b \in \mathcal{B}} \alpha_b b$  uniquely as noted earlier. Hence if  $T$  exists and is a linear transformation,

$$\begin{aligned} T(v) &= T\left(\sum_{b \in \mathcal{B}} \alpha_b b\right) \\ &= \sum_{b \in \mathcal{B}} \alpha_b T(b) \end{aligned}$$

yielding

$$T(v) = \sum_{b \in \mathcal{B}} \alpha_b t(b). \quad (4)$$

Thus  $T(v)$ , if it exists, is uniquely determined by  $t$  (since the  $\alpha_b$  are uniquely determined by  $\mathcal{B}$  and  $v$ ).

**Existence:** In view of the previous part, we define  $T$  by the formula in Equation (4). In contrast to the situation we found for quotient spaces, there is no ambiguity at all (no choices were made, no representatives were used) in this definition. Hence  $T$  exists as a *function*.

**Properties:** Finally, we must show that  $T$  is the right kind of function, a linear transformation. Suppose we also have a vector  $w \in V$  and we write  $w = \sum_{b \in \mathcal{B}} \beta_b b$  for unique  $\beta_b \in F$ . Then  $v + w = \sum_{b \in \mathcal{B}} (\alpha_b + \beta_b) b$

$$\begin{aligned} T(v + w) &= T\left(\sum_{b \in \mathcal{B}} (\alpha_b + \beta_b) b\right) \\ &= \sum_{b \in \mathcal{B}} (\alpha_b + \beta_b) t(b) \\ &= \sum_{b \in \mathcal{B}} \alpha_b t(b) + \sum_{b \in \mathcal{B}} \beta_b t(b) \\ &= T(v) + T(w) \end{aligned}$$

And finally if  $c \in F$ , then  $cv = \sum_{b \in \mathcal{B}} c\alpha_b b$ . Hence

$$\begin{aligned} T(cv) &= T\left(\sum_{b \in \mathcal{B}} c\alpha_b b\right) \\ &= \sum_{b \in \mathcal{B}} c\alpha_b t(b) \\ &= c\left(\sum_{b \in \mathcal{B}} \alpha_b t(b)\right) \\ &= cT(v) \end{aligned}$$

□

**Remark 15.** 1. The term “commutes” will more generally mean the following: if one has a diagram with a number of objects (e.g., vector spaces, fields, or whatever) with a number of functions (arrows) between some of the objects, we say that the diagram *commutes* if for every pair of objects which can be connected by a path (all arrows pointing the same direction so that composition of the corresponding functions is possible) in more than one way, the compositions of the functions along the various possible paths must always be equal. Note that in our situation above, this didn’t say a lot as there were only two objects that could be connected by two paths this way and hence only one resulting equation.

2. A universal mapping property such as the one just described always gives a one-to-one correspondence (bijection) between two collections of functions. In this case the first is just a collection of ordinary functions, while the second is a set of linear transformations.

$$\theta: \{t \mid t: \mathcal{B} \longrightarrow U\} \longrightarrow \text{Hom}_F(V, U)$$

where we write  $\text{Hom}_F(V, U)$  for the set of all linear transformations from  $V$  to  $U$  (this is denoted  $\mathcal{L}(V, U)$  in some textbooks).

Here  $\theta(t) = T$  is given by the theorem. This map is one-to-one as two different functions give two different linear transformations. Further,  $\theta$  is onto, since given any  $T \in \text{Hom}_F(V, U)$  we can define the required  $t$  by  $t = T \circ i$ , i.e., just  $T$  restricted to  $\mathcal{B}$ :  $\theta(T \circ i) = T$ .

## Dimension and Dimension Formulas

Our next application of Theorem 5 will be to define the dimension of a vector space: If  $V$  is a vector space over  $F$ , then the *dimension* of  $V$  is the cardinality of any basis for  $V$ . So if  $\mathcal{B}$  is a basis for  $V$ , we write  $\dim_F V = |\mathcal{B}|$ . So the examples we had earlier yield  $\dim_F F^n = n$ ,  $\dim_F F^{m \times n} = mn$ ,  $\dim_F F^S = |S|$  for a finite non-empty set  $S$ , and  $\dim_F F[x] = |\mathbb{N}|$  where  $\mathbb{N} = \{0, 1, 2, \dots, i, \dots\}$ . The set  $\mathbb{N}$ , or any other set with the same cardinality, is said to be *countable*. Many times this cardinal number is denoted by the Hebrew letter aleph with a subscript zero:  $|\mathbb{N}| = \aleph_0$ . See the section on the “Axiom of Choice and Zorn’s Lemma” for further information.

We begin by listing a few simple facts that are easily derived from our previous discussion. Below  $V$  and  $V'$  will denote vector spaces over the field  $F$ . We write  $\dim V$  for  $\dim_F V$ .

**Fact 1.** Let  $\mathcal{B}$  be a basis for  $V$  and let  $S \subseteq V$ .

- a. If  $S$  is linearly independent, then  $|S| \leq |\mathcal{B}|$ .
- b. If  $S$  spans  $V$ , then  $|S| \geq |\mathcal{B}|$ .
- c. If  $|S| > \dim V$ , then  $S$  is linearly dependent.
- d. If  $|S| < \dim V$ , then  $\text{Span}_F(S) \neq V$ .

*Proof.* By the Main Theorem on Bases (Theorem 5) any linearly independent set can be enlarged to a basis, so part a. is clear as all bases have the same number of elements. Similarly, any spanning set contains a basis, yielding part b.

Part c. follows as  $S$  would contain a basis which is not all of  $S$  and any of the other elements of  $S$  would have to be a linear combination of some of these these (which gives a dependence relation).

If  $\text{Span}_F(S) = V$ , then  $S$  would contain a basis, so  $|S| \geq \dim V$ , contradicting the assumption. □

**Fact 2.** Let  $W \subseteq V$  be a subspace. Assume  $\dim V < \infty$ . Then  $W = V$  if and only if  $\dim W = \dim V$ .

*Proof.* A basis  $\mathcal{A}$  for  $W$  can be enlarged to a basis  $\mathcal{B}$  for  $V$  since  $\mathcal{A}$  is a linearly independent set. If  $\mathcal{A}$  has fewer elements than  $|\mathcal{B}| = \dim V$ , then  $W \neq V$ . If  $\mathcal{A}$  has the same number of elements as  $\mathcal{B}$ , then  $\mathcal{A} = \mathcal{B}$ , so  $W = V$ .  $\square$

**Fact 3.** Let  $W \subseteq V$  be a subspace, with  $\dim W = \dim V$ . Assume  $\dim V$  is not finite. Then  $W$  may or may not be equal to  $V$ .

*Proof.* Consider  $V = F[x]$  with basis  $\mathcal{B} = \{1, x, x^2, x^3, \dots, x^i, \dots\}$  and proper subset  $\mathcal{A} = \{x, x^2, x^3, \dots, x^i, \dots\}$ . Then the function  $f: \mathcal{B} \rightarrow \mathcal{A}$  given by  $f(x^i) = x^{i+1}$  is one-to-one and onto. Hence the proper subspace  $W = \text{Span}_F(\mathcal{A})$  has the same dimension as  $V$ . For an arbitrary infinite dimensional vector space  $V$  with basis  $\mathcal{B}$  one can embed  $\{1, 2, \dots, n, \dots\}$  into  $\mathcal{B}$  and use the same idea to construct a proper subspace with the same dimension as  $V$ .  $\square$

**Fact 4.** Let  $\dim V = n < \infty$  and let  $S \subseteq V$ . Assume  $|S| = n$ . Then the following are equivalent:

- The set  $S$  is linearly independent.
- The set  $S$  spans  $V$ :  $\text{Span}_F(S) = V$ .
- The set  $S$  is a basis for  $V$ .

*Proof.* This is immediate from the first Fact; or follows directly from the Main Theorem on Bases by using the same ideas: enlarge  $S$  to a basis in the first part, find a subset which is a basis in the second part. In either case one ends up with just  $S$  as otherwise we have a basis that is too large, or too small. And of course the last part implies the other two.  $\square$

**Fact 5.** If  $\mathcal{B}$  is a basis for  $V$ , then  $V$  is isomorphic to a direct sum of  $|\mathcal{B}|$  copies of  $F$ .

*Proof.* This is just the content of Theorem 12 giving coordinates with respect to a basis.  $\square$

**Fact 6.** We have  $\dim V = \dim V'$  if and only if  $V \approx V'$ .

*Proof.* If  $\mathcal{B}$  is a basis for  $V$  and  $\mathcal{B}'$  is a basis for  $V'$ , there is a one-to-one onto function  $t: \mathcal{B} \rightarrow \mathcal{B}'$  by the definition of equality of dimension (and the fact that any two bases for a vector space have the same cardinality). The Universal Mapping Property for Bases implies that there is a linear transformation  $T: V \rightarrow V'$  which is  $t$  on the basis  $\mathcal{B}$ . Applying this again for the inverse function  $t^{-1}: \mathcal{B}' \rightarrow \mathcal{B}$  gives a linear transformation  $T^{-1}$  which is the inverse of  $T$ .

Conversely if  $T: V \rightarrow V'$  is an isomorphism, it is easy to check that  $T(\mathcal{B})$  is a basis for  $V'$  (so  $T$  gives the one-to-one, onto function between bases), yielding  $\dim V = \dim V'$ .  $\square$

**Fact 7.** Let  $A \in F^{n \times n}$  for  $n > 0$  an integer. Then the following are equivalent:

- a. The rows of  $A$  are linearly independent.
- b. The columns of  $A$  are linearly independent.
- c. The rows of  $A$  span  $F^{1 \times n}$ .
- d. The columns of  $A$  span  $F^{n \times 1}$ .
- e.  $A$  has a left inverse.
- f.  $A$  has a right inverse.
- g.  $A$  has an inverse.
- h. The row reduced echelon form of  $A$  is  $I$  (the  $n \times n$  identity matrix).
- i. The column reduced echelon form of  $A$  is  $I$  (the  $n \times n$  identity matrix).
- j.  $A$  is a product of elementary matrices.

*Proof.* The column space of  $A$  is contained in  $F^{n \times 1}$  and the row space of  $A$  is contained in  $F^{1 \times n}$ , both of which have dimension  $n$ . In view of the previous facts, as  $A$  has  $n$  rows and  $n$  columns, a. through d. are equivalent (and are also equivalent to the statements that the rows are a basis for the row space and the columns are a basis of the column space).

Now the first chapter of Hoffman and Kunze, Theorem 12 on page 23 shows that g. through j. are equivalent.

Now  $A$  is invertible (part g.) if and only if  $A$  has a left inverse (part e.) and  $A$  has a right inverse (part f.).

We now tie these different sets of statements together. If we call the rows of  $A$   $v_1, \dots, v_n$  and assume part c. (that the  $v_i$  span the row space), then there exist  $b_{ij} \in F$  so that

$$\begin{aligned} b_{11}v_1 + \cdots + b_{1n}v_n &= e_1 \\ b_{21}v_1 + \cdots + b_{2n}v_n &= e_2 \\ &\vdots \\ b_{n1}v_1 + \cdots + b_{nn}v_n &= e_n \end{aligned}$$

where  $\{e_1, \dots, e_n\}$  is the standard basis of  $F^{1 \times n}$ . Letting  $B$  be the matrix with entries  $b_{ij}$  the system of equations above is just  $BA = I$ , the  $n \times n$  identity matrix. This is just part e. On the other hand the equation  $BA = I$  gives a system of equations as above that shows the rows of  $A$  span. Similarly part d. (columns span) implies part f, and vice versa.  $\square$

The proofs of the next five theorems use the same ideas. We will thus give the most general (abstract) statement and derive the others as consequences. On the other hand, the outline of the proof in this part easily gives the usual (e.g., in many textbooks) proofs of the other results.

**Theorem 16.** *Let  $W$ ,  $V$ , and  $U$  be vector spaces over the field  $F$  and assume that*

$$0 \longrightarrow W \longrightarrow V \longrightarrow U \longrightarrow 0$$

*is a short exact sequence. Then*

$$\dim V = \dim W + \dim U .$$

*Proof.* In outline we do the following: We take a basis  $\mathcal{A}$  for  $W$ , apply the linear transformation to get a linearly independent set  $\mathcal{B}_1$  of the same cardinality in  $V$ , enlarge to a basis  $\mathcal{B}$  for  $V$ . We write  $\mathcal{B} = \mathcal{B}_1 \dot{\cup} \mathcal{B}_2$  where  $\mathcal{B}_2$  is the set of additional vectors we had to add (disjoint from  $\mathcal{B}_1$ ). We apply the next linear transformation to this basis  $\mathcal{B}$ . The linear transformation sends  $\mathcal{B}_1$  to 0 and we define  $\mathcal{C}$  to be the image of  $\mathcal{B}_2$  in  $U$ . It will have the same cardinality as  $\mathcal{B}_2$ . This will yield the result:

$$\begin{aligned} \dim V &= |\mathcal{B}| \\ &= |\mathcal{B}_1| + |\mathcal{B}_2| \\ &= |\mathcal{A}| + |\mathcal{C}| \\ &= \dim W + \dim U . \end{aligned}$$

Let  $S : W \longrightarrow V$  be the first linear transformation and  $T : V \longrightarrow U$  the second. The sequence being exact means

- $S$  is one-to-one,
- $\text{im } S = \ker T$ ,
- $T$  is onto.

We set  $\mathcal{B}_1 = S(\mathcal{A})$ . Since  $S$  is one-to-one, the two sets have the same cardinality. Further, if  $w = \sum_{a \in \mathcal{A}} \alpha_a a$  is a linear combination of elements of  $\mathcal{A}$ , then  $S(w) = \sum_{a \in \mathcal{A}} \alpha_a S(a)$ , is a linear combination of elements of  $\mathcal{B}_1 = S(\mathcal{A})$ . This linear combination will be 0 in  $V$  if and only if the original linear combination,  $w$ , is 0 in  $W$  since  $S$  is one-to-one. That is,  $\mathcal{B}_1$  is a linearly independent subset of  $V$ . By the Main Theorem on Bases, there exists a basis  $\mathcal{B}$  for  $V$  which contains the linearly independent set  $\mathcal{B}_1$ . We write  $\mathcal{B} = \mathcal{B}_1 \dot{\cup} \mathcal{B}_2$  where  $\mathcal{B}_2$  (disjoint from  $\mathcal{B}_1$ ) consists of the extra vectors that were needed.

Next we let  $\mathcal{C} = T(\mathcal{B}_2)$ . As  $T$  is onto,  $T(\mathcal{B})$  spans  $U$ , but  $T(\mathcal{B}_1) = TS(\mathcal{A}) = 0$ , so  $\mathcal{C} = T(\mathcal{B}_2)$  spans  $U$ . The only thing left to show is that  $\mathcal{C} = T(\mathcal{B}_2)$  is linearly

independent. Assume that there exist  $\beta_{b,2} \in F$  such that

$$\begin{aligned} 0 &= \sum_{b \in \mathcal{B}_2} \beta_{b,2} T(b) \\ &= T\left(\sum_{b \in \mathcal{B}_2} \beta_{b,2} b\right). \end{aligned}$$

Thus  $\sum_{b \in \mathcal{B}_2} \beta_{b,2} b \in \ker T = \text{im } S$ . Hence there exist  $\beta_{b,1} \in F$  so that

$$\sum_{b \in \mathcal{B}_1} \beta_{b,1} b = \sum_{b \in \mathcal{B}_2} \beta_{b,2} b$$

or

$$0 = \sum_{b \in \mathcal{B}_1} \beta_{b,1} b - \sum_{b \in \mathcal{B}_2} \beta_{b,2} b.$$

Now as  $\mathcal{B} = \mathcal{B}_1 \dot{\cup} \mathcal{B}_2$  we have produced a linear combination of the basis elements of  $V$  which is 0. Hence all  $\beta_{b,j} = 0$ , and in particular it follows that  $\mathcal{C} = T(\mathcal{B}_2)$  is linearly independent, completing the proof.  $\square$

**Theorem 17.** *Let  $V_1$  and  $V_2$  be vector spaces over the field  $F$ . Then*

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2.$$

*Proof.* Earlier we showed that the following sequence is exact:

$$0 \longrightarrow V_1 \xrightarrow{\iota} V_1 \oplus V_2 \xrightarrow{\pi} V_2 \longrightarrow 0$$

and hence this result follows from the short exact sequence version, Theorem 16. Here  $\iota$  is inclusion into  $V_1$  and  $\pi$  is projection onto  $V_2$ .

If one wants a specific basis, then given bases  $\mathcal{B}_1$  for  $V_1$  and  $\mathcal{B}_2$  for  $V_2$ , a basis  $\mathcal{B}$  constructed as above is the union of the two collections of the form  $\mathcal{B}'_1 = \{(b_1, 0) \mid b_1 \in \mathcal{B}_1\}$  and  $\mathcal{B}'_2 = \{(0, b_2) \mid b_2 \in \mathcal{B}_2\}$ .  $\square$

**Remark 18.** Subtraction of cardinal numbers is neither well-defined nor useful in most cases (e.g., removing an infinite subset from  $\{1, 2, \dots, n, \dots\}$  can leave a finite or an infinite subset). However, for finite sets we do it all the time which is the content of the last equation in the next two results.

**Theorem 19.** *Let  $V$  be a vector space over the field  $F$  with  $W \subseteq V$  a subspace. Then*

$$\dim V = \dim W + \dim V/W.$$

*If  $V$  has finite dimension, then*

$$\dim V/W = \dim V - \dim W.$$

*Proof.* The construction of quotient spaces gives the exact sequence:

$$0 \longrightarrow W \xrightarrow{i} V \xrightarrow{p} V/W \longrightarrow 0$$

and hence the result follows from the general one.  $\square$

**Theorem 20.** *Let  $V$  be a vector space over  $F$  with two subspaces  $W_1, W_2 \subseteq V$ . Then*

$$\dim(W_1 \cap W_2) + \dim(W_1 + W_2) = \dim W_1 + \dim W_2 .$$

*If  $W_1$  and  $W_2$  have finite dimension, then*

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2) .$$

*Proof.* Consider the linear transformation  $s : W_1 \oplus W_2 \longrightarrow W_1 + W_2$  given by  $s(w_1, w_2) = w_1 + w_2$  (the formal pair  $(w_1, w_2)$  goes to the actual sum  $w_1 + w_2$  in  $V$ ). Thus  $s$  is onto by definition of  $W_1 + W_2$  and the kernel consists of pairs  $(w_1, w_2)$  such that  $w_1 + w_2 = 0$ . That is,  $w_1 = -w_2$ ; call this vector  $w$ . It is clearly in  $W_1 \cap W_2$ . We then have an exact sequence

$$0 \longrightarrow W_1 \cap W_2 \xrightarrow{j} W_1 \oplus W_2 \xrightarrow{s} W_1 + W_2 \longrightarrow 0$$

where  $j(w) = (w, -w)$ , and hence the result follows from the general one.  $\square$

**Theorem 21.** *Let  $U$  and  $V$  be vector spaces over the field  $F$  with  $\dim_F U = \dim_F V$  finite. Let  $T : U \longrightarrow V$  be a linear transformation. Then the following are equivalent:*

- a.  $T$  has an inverse.
- b.  $T$  is one-to-one.
- c.  $T$  is onto.

*Proof.* In general we have  $\dim_F \ker T = 0$  if and only if  $\ker T = 0$  if and only if  $T$  is one-to-one. And for finite dimensional  $V$  we have in general  $\dim_F \operatorname{im} T = \dim_F V$  if and only if  $\operatorname{im} T = V$  if and only if  $T$  is onto. Since  $\dim_F U = \dim_F V = \dim_F \ker T + \dim_F \operatorname{im} T$  the two conditions are the same. Note that the second equality comes from the isomorphism  $V/\ker T \approx \operatorname{im} T$ . That is, parts b. and c. are equivalent. But part a. is equivalent to both b. and c. holding. This completes the proof.  $\square$

**Remark 22.** This result is false if the two vector spaces have equal but infinite dimension. For example, take  $F[x]$  and define the linear transformation  $\mathbf{up} : F[x] \longrightarrow F[x]$  on the standard basis by  $\mathbf{up}(x^i) = x^{i+1}$ . Similarly define the linear transformation  $\mathbf{down} : F[x] \longrightarrow F[x]$  by  $\mathbf{down}(1) = 0$  and  $\mathbf{down}(x^i) = x^{i-1}$  for  $i > 0$ . It is clear that  $\mathbf{down} \circ \mathbf{up} = 1_{F[x]}$  as it has this property on the basis; so  $\mathbf{down}$  is a left inverse to  $\mathbf{up}$ . Thus  $\mathbf{up}$  is one-to-one and  $\mathbf{down}$  is onto. But  $\mathbf{up}$  is not onto as  $1 \notin \operatorname{im} \mathbf{up}$  and  $\mathbf{down}$  is not one-to-one as  $1 \in \ker \mathbf{down}$ . One can give similar examples for any infinite basis  $\mathcal{B}$  by embedding  $\{1, 2, \dots, n, \dots\}$  into  $\mathcal{B}$  and using exactly the same idea.

## Vector Spaces of Linear Transformations: $\text{Hom}_F(\mathbf{U}, \mathbf{V})$

Let  $U$  and  $V$  be vector spaces over the field  $F$ . We denote by  $\text{Hom}_F(U, V)$  the set of all linear transformations from  $U$  to  $V$  (denoted  $\mathcal{L}(U, V)$  in some texts).

Addition and scalar multiplication are defined on  $\text{Hom}_F(U, V)$  as follows: For  $S, T \in \text{Hom}_F(U, V)$ ,  $u \in U$  and  $c \in F$  by

$$(S + T)(u) = S(u) + T(u)$$

and

$$(cT)(u) = cT(u).$$

Note that this is a generalization of the idea that was used to define a vector space structure on  $F^S$  for a non-empty set  $S$ . It is easy to check (and will be left to the reader) that the resulting functions,  $S + T$  and  $cT$  are indeed linear transformations from  $U$  to  $V$ , and that the axioms for a vector space are satisfied. For example, one easily checks that the  $0$  function is the zero of  $\text{Hom}_F(U, V)$  and  $-T$  is the additive inverse of  $T$ .

**Theorem 23.** *Let  $U$  and  $V$  be finite dimensional vector spaces over the field  $F$ . Then*

$$\dim_F \text{Hom}_F(U, V) = \dim_F U \cdot \dim_F V.$$

*Proof.* Let  $\mathcal{A}$  be a basis for  $U$  and  $\mathcal{B}$  a basis for  $V$ . For each  $a \in \mathcal{A}$  and  $b \in \mathcal{B}$  we define a linear transformation  $T_{a,b} : U \rightarrow V$  via the Universal Mapping Property by specifying it on a basis  $\mathcal{A}$  for  $U$ :

$$T_{a,b}(a') = \begin{cases} 0 & a' \neq a \\ b & a' = a \end{cases}$$

for  $a' \in \mathcal{A}$ .

We will show that  $\mathcal{C} = \{ T_{a,b} \mid a \in \mathcal{A}, b \in \mathcal{B} \}$  is a basis for  $\text{Hom}_F(U, V)$ . We then have

$$\begin{aligned} \dim U \cdot \dim V &= |\mathcal{A}| \cdot |\mathcal{B}| \\ &= |\mathcal{A} \times \mathcal{B}| \\ &= \left| \left\{ T_{a,b} \mid a \in \mathcal{A}, b \in \mathcal{B} \right\} \right| \end{aligned}$$

We first show that the set  $\mathcal{C}$  is linearly independent. Suppose that there exist  $\gamma_{a,b} \in F$  such that

$$0 = \sum_{a \in \mathcal{A}, b \in \mathcal{B}} \gamma_{a,b} T_{a,b}.$$

Note that the 0 on the left is the zero linear transformation from  $U$  to  $V$  (whereas it is the 0 of  $F$  in the next equation below). We apply this linear transformation to some  $a' \in \mathcal{A}$  and obtain

$$\begin{aligned} 0 &= \sum_{a \in \mathcal{A}, b \in \mathcal{B}} \gamma_{a,b} T_{a,b}(a') \\ &= \sum_{b \in \mathcal{B}} \gamma_{a',b} b \end{aligned}$$

and as  $\mathcal{B}$  is a basis yields  $\gamma_{a',b} = 0$ , for any  $a' \in \mathcal{A}$  and any  $b \in \mathcal{B}$ .

We finally show that  $\mathcal{C}$  spans  $\text{Hom}_F(U, V)$  which completes the proof. Given  $S \in \text{Hom}_F(U, V)$  since  $\mathcal{B}$  is a basis for  $V$ , there exist  $\beta_{a,b} \in F$  so that

$$S(a) = \sum_{b \in \mathcal{B}} \beta_{a,b} b$$

and by our definition of  $T_{a,b}$  we thus have

$$S(a) = \sum_{b \in \mathcal{B}} \beta_{a,b} T_{a,b}(a).$$

Thus  $S$  and  $\sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \beta_{a,b} T_{a,b}$  agree on every basis element  $a' \in \mathcal{A}$  and hence are the same linear transformation, that is

$$S = \sum_{a \in \mathcal{A}, b \in \mathcal{B}} \beta_{a,b} T_{a,b}.$$

□

**Remark 24.** This result is false for infinite dimensional vector spaces. See the exercises for a more precise description of what happens to  $\text{Hom}_F(U, V)$  when the entries are replaced by direct sums or products. For example, even  $\text{Hom}_F(F[x], F)$  has strictly greater dimension than that of  $F[x]$ . This is relevant to a topic we'll consider soon: dual spaces.

## Exercises

The phrase “natural isomorphism” means that no arbitrary choices should be made (e.g., by choosing bases); one should give formulas using only the definitions or whatever other information is given. Do not assume that the vector spaces have finite dimension unless that is stated specifically.

**BaseCoord 1.** Let  $V$  be the real vector space spanned by the rows of the matrix

$$A = \begin{bmatrix} 3 & 21 & 0 & 9 & 0 \\ 1 & 7 & -1 & -2 & -1 \\ 2 & 14 & 0 & 6 & 1 \\ 6 & 42 & -1 & 13 & 0 \end{bmatrix}$$

- a. Find a basis for  $V$ .
- b. Tell which vectors  $(x_1, x_2, x_3, x_4, x_5)$  are elements of  $V$ .
- c. If  $(x_1, x_2, x_3, x_4, x_5)$  is in  $V$ , what are its coordinates with respect to the basis chosen in part a?

**BaseCoord 2.** a. Let  $V$  be a vector space over the field  $\mathbb{R}$  of real numbers. Let  $u, v, w$  be linearly independent vectors in  $V$ . Prove that  $u + v$ ,  $v + w$ , and  $u + w$  are linearly independent as well.

- b. Does the same statement hold when  $F$  is replaced by an arbitrary field? Determine precisely what is true.

**BaseCoord 3.** Suppose  $F$  is a finite field of characteristic  $p$ . Prove that the number of elements in  $F$  is equal to  $p^n$  for some integer  $n \geq 1$ . [What is the prime subfield of  $F$ ?]

**BaseCoord 4.** Let  $\mathbb{R}$  be the real numbers. Regard  $\mathbb{R}$  as a vector space over the field  $\mathbb{Q}$  of rational numbers, with the usual operations. Prove that this vector space is not finite-dimensional.

Note that the solution of this problem requires knowing mathematics not developed in this course. Give as many really different proofs as you can find! There at least four which are really different and use ideas from (1) number theory, (2) algebra, (3) analysis, (4) logic. You do not have to prove whatever theorem you use, but when necessary, give the step required to show  $\mathbb{R}$  is infinite dimensional over  $\mathbb{Q}$  using the theorem(s) you choose.

**BaseCoord 5.** Let  $u = (x_1, x_2)$  and  $v = (y_1, y_2)$  be vectors in  $\mathbb{R}^2$  such that

$$\begin{aligned} x_1 y_1 + x_2 y_2 &= 0 \\ x_1^2 + x_2^2 = y_1^2 + y_2^2 &= 1 \end{aligned}$$

Prove that  $\mathcal{B} = \{u, v\}$  is a basis for  $\mathbb{R}^2$ . Find the coordinates of the vector  $(a, b)$  with respect to the ordered basis  $\mathcal{B} = \{u, v\}$ . (The conditions on  $u$  and  $v$  say, geometrically, that  $u$  and  $v$  are perpendicular and that each has length 1.)

**BaseCoord 6.** Let  $V$  be an  $n$ -dimensional vector space over a field  $F$  and let  $T$  be a linear transformation from  $V$  to  $V$  such that  $\text{im } T = \ker T$ . Prove that  $n$  is even. Give an example of such a linear transformation  $T$ .

**BaseCoord 7.** Let  $V$  be a finite-dimensional vector space over  $F$  and let  $T : V \rightarrow V$  be a linear transformation. Suppose that  $\text{rank}(T^2) = \text{rank}(T)$ . Prove that the  $\ker T \cap \text{im } T = 0$ .

**BaseCoord 8.** Let  $V$  and  $W$  be vector spaces over a field  $F$ , with  $V$  not trivial. Show that

$$W = \sum \{ \text{im } \alpha \mid \alpha \in \text{Hom}_F(V, W) \} .$$

That is, show that  $W$  is spanned by the collection of subspaces given by the images of *all* linear transformations from  $V$  to  $W$ .

**BaseCoord 9.**  $\boxed{\star}$  You may assume that factorization of integers is unique. Let  $p_1, p_2, \dots, p_k$  be  $k$  distinct primes.

- Prove that the set  $\{ \sqrt{p_i} \mid 1 \leq i \leq k \}$  is linearly independent over  $\mathbb{Q}$ .
- Compute the dimension of  $\mathbb{Q}[\sqrt{p_1}, \dots, \sqrt{p_k}]$  over  $\mathbb{Q}$ .

**BaseCoord 10.** a) Let  $f : V \rightarrow W$  be a linear transformation. Suppose  $W'$  is a finite dimensional subspace of  $W$ , and that  $\ker f$  is finite dimensional. Prove that

$$f^{-1}(W') := \{ v \in V \mid f(v) \in W' \}$$

is a finite dimensional subspace of  $V$  (note that it is always a subspace of  $V$ , regardless of the finiteness assumptions). Show that  $f^{-1}(W')$  will be infinite dimensional if  $\ker f$  is infinite dimensional and might be if  $W'$  is infinite dimensional.

- Let  $f : V \rightarrow W$  and  $g : W \rightarrow Y$  be linear transformations such that  $\ker f$  and  $\ker g$  are finite dimensional. Show that  $\ker gf$  is finite dimensional.

**BaseCoord 11.** Let  $p$ ,  $m$ , and  $n$  be positive integers and  $F$  a field. Let  $V$  be the vector space  $F^{m \times n}$ , of all  $m \times n$  matrices over  $F$  and  $W = F^{p \times n}$ . Let  $A$  be a fixed  $p \times m$  matrix and define  $T(M) = AM$  for  $M \in V$ . Prove that  $T$  has an inverse if and only if  $p = m$  and  $A$  is an invertible  $m \times m$  matrix.

**BaseCoord 12.** Let  $U, V, U_1, U_2, V_1, V_2$  be vector spaces over the same field  $F$ .

- Give a natural isomorphism

$$\text{Hom}_F(U_1 \oplus U_2, V) \longrightarrow \text{Hom}_F(U_1, V) \oplus \text{Hom}_F(U_2, V) .$$

- Give a natural isomorphism

$$\text{Hom}_F(U, V_1 \oplus V_2) \longrightarrow \text{Hom}_F(U, V_1) \oplus \text{Hom}_F(U, V_2) .$$

**BaseCoord 13.** Let  $U, V, U_i, i \in I, V_j, j \in J$  be vector spaces over the same field  $F$ .

a. Give a natural isomorphism

$$\text{Hom}_F\left(\bigoplus_{i \in I} U_i, V\right) \longrightarrow \prod_{i \in I} \text{Hom}_F(U_i, V) .$$

b. Give a natural isomorphism

$$\text{Hom}_F\left(U, \prod_{j \in J} V_j\right) \longrightarrow \prod_{j \in J} \text{Hom}_F(U, V_j) .$$

c. Is there a similar natural description of

$$\text{Hom}_F\left(U, \bigoplus_{j \in J} V_j\right)$$

or of

$$\text{Hom}_F\left(\prod_{i \in I} U_i, V\right) ?$$

d. Can you use these isomorphisms to give the dimension of the “dual space” of  $F[x]$ ? That is, of the vector space

$$F[x]^* = \text{Hom}_F(F[x], F) ?$$

**BaseCoord 14.** If  $F$  is a field with a finite or a countable number of elements and  $V$  is an infinite dimensional vector space over  $F$ , show that  $\dim_F V = |V|$ .

**BaseCoord 15.** Let  $F$  be a field and  $X$  an infinite set. Show that the two vector spaces

$$\bigoplus_{x \in X} F = F^{(X)}$$

and

$$\prod_{x \in X} F = F^X$$

are not isomorphic. Show not only that the natural inclusion is not an isomorphism, but no isomorphism exists.

**BaseCoord 16.** Let  $F$  be any field and let  $F[x]$  be the formal polynomials with coefficients in  $F$ . Let  $\mathcal{A} = \{f_i \mid i \geq 1\}$  be a set of non-zero polynomials.

a. If  $\mathcal{A}$  satisfies  $\deg f_i \neq \deg f_j$  for  $i \neq j$ , then show that  $\mathcal{A}$  is linearly independent.

b. If in addition  $\mathcal{A}$  satisfies  $\{\deg f_i \mid f_i \in \mathcal{A}\}$  is the set of all non-negative integers, then  $\mathcal{A}$  is a basis for  $F[x]$ .

- c. Let  $a \in F$  and let  $\mathcal{B} = \{(x - a)^i \mid i \geq 0\}$ . Prove that  $\mathcal{B}$  is a basis for  $F[x]$ .
- d. Let  $a_j \in F$  for  $j \geq 1$  be a set of distinct elements of  $F$  (so  $F$  must be infinite). Let  $g_j = \prod_{i=1}^{j-1} (x - a_i)$  and let  $g_0 = 1$ . So  $g_1 = x - a_1$ ,  $g_2 = (x - a_1)(x - a_2)$ , etc. Show that  $\{g_j \mid j \geq 0\}$  is a basis for  $F[x]$ .

**BaseCoord 17.** Let  $W$  be the vector space of all continuous real valued functions on  $\mathbb{R}$ .

- a. Let  $\mathcal{E} = \{e^{cx} \mid c \in \mathbb{R}\}$ . Let  $V = \text{Span}_{\mathbb{R}}(\mathcal{E})$ . Show that  $\mathcal{E}$  is a linearly independent subset of  $W$  and hence  $\dim_{\mathbb{R}} V = |\mathbb{R}|$ .
- b. Let  $P = \{r \in \mathbb{R} \mid r > 0\}$ . Give a one-to-one, onto function  $f : \mathbb{R} \rightarrow P$  thus showing that  $|\mathbb{R}| = |P|$ .
- c. Let  $\mathcal{C} = \{\cos(cx) \mid c \in \mathbb{R}, c > 0\}$ . Let  $V_1 = \text{Span}_{\mathbb{R}}(\mathcal{C})$ . Show that  $\mathcal{C}$  is a linearly independent subset of  $W$  and hence  $\dim_{\mathbb{R}} V_1 = |\mathbb{R}|$ .
- d. Let  $\mathcal{S} = \{\sin(cx) \mid c \in \mathbb{R}, c > 0\}$ . Let  $V_2 = \text{Span}_{\mathbb{R}}(\mathcal{S})$ . Show that  $\mathcal{S}$  is a linearly independent subset of  $W$  and hence  $\dim_{\mathbb{R}} V_2 = |\mathbb{R}|$ .
- e. Compute the kernel and cokernel of the linear transformation given by differentiation  $D : V \rightarrow V$ . Let  $V_0 = \text{Span}_{\mathbb{R}}(\mathcal{E} \setminus \{1\})$ . Do the same for  $D : V_0 \rightarrow V_0$ .
- f. Compute the kernel and cokernel of the linear transformation given by differentiation  $D : V_1 \rightarrow V_2$  and  $D : V_2 \rightarrow V_1$ .

**BaseCoord 18.**  $\boxed{\star}$

- a. Let  $F$  be an arbitrary field and let  $F^\infty$  be the vector space of all infinite sequences  $(a_1, a_2, \dots)$  of elements of  $F$ . Addition is coordinatewise and scalar multiplication by  $a \in F$  just multiplies each entry by  $a$ . Define  $\mathbf{L} : F^\infty \rightarrow F^\infty$  by  $\mathbf{L}(a_1, a_2, a_3, \dots) = (a_2, a_3, a_4, \dots)$  ("shift left"). Note that  $\mathbf{L}$  is a linear transformation. It is onto with a kernel of dimension 1. For  $a \in F$  which is non-zero define the vector  $v(a) = (1, a, a^2, a^3, \dots) \in F^\infty$  whose  $i$ -th entry is  $a^{i-1}$ . Show that  $\mathbf{L}(v(a)) = av(a)$ . Prove that  $\dim_F F^\infty \geq |F|$ . (Hint: Show that  $\{v(a) \mid a \in F\}$  is linearly independent.)
- b. Show that  $\dim_F F^\infty$  is uncountable, i.e., bigger than  $\aleph_0 = |\mathbb{Z}| = |\mathbb{Q}|$ . Note that it will suffice to do this in case  $F$  is countable because ...

**BaseCoord 19.** Let  $F$  be a field and  $V$  a vector space over  $F$ . Let  $\mathcal{C} = \{u_1, \dots, u_n\}$  be a subset of  $V$ .

- a. Let  $v \in V$  be a non-zero vector and assume that  $\text{Span}_F(\mathcal{C}) = V$ . Show that there exists an integer  $i$ ,  $1 \leq i \leq n$  so that for

$$\mathcal{C}' = \{v\} \cup \mathcal{C} \setminus \{u_i\}$$

then  $\text{Span}_F(\mathcal{C}') = V$ . That is, for some  $i$ ,  $v$  can *replace*  $u_i$  in the spanning set  $\mathcal{C}$  and the result is also a spanning set.

- b. Let  $\mathcal{A} = \{v_1, \dots, v_m\}$  be a linearly independent set in  $V$ . Assume that  $\text{Span}_F(\mathcal{C}) = V$ . Show that there exists a subset of indices  $I \subseteq \{1, \dots, n\}$  with  $|I| = |\mathcal{A}|$  so that

$$\mathcal{C}' = \mathcal{A} \cup \mathcal{C} \setminus \{u_i \mid i \in I\}$$

spans  $V$ ; that is, the independent set  $\mathcal{A}$  can replace a subset of exactly the same size in the spanning set  $\mathcal{C}$  to yield a new spanning set.

- c. Use this result to prove that any two bases  $\mathcal{B}_1$  and  $\mathcal{B}_2$  have the same number of elements for  $V$  a finite dimensional vector space.

**BaseCoord 20.** a. Let  $V$  be a vector space over the field  $F$ . Let  $f : V \rightarrow F$  be a non-trivial linear transformation. Show that there will exist a vector  $v_0 \in V$  with  $f(v_0) = 1$  and further that  $V = \ker f \oplus Fv_0$ , where  $Fv_0 = \{av_0 \mid a \in F\}$ . Conclude that  $V/\ker f \approx F$  has dimension 1. The subspace  $\ker f$  is called a *hyperplane*, and is said to have codimension 1. In general if  $W \subseteq V$  is a subspace, the number  $\dim V/W$  is called the *codimension* of  $W$  in  $V$ .

- b. Let  $W_1, \dots, W_k \subset V$  be a collection of subspaces. Assume each  $W_i$  has finite codimension in  $V$ . Prove that their intersection has finite codimension in  $V$ . Give an upper bound for that number in terms of the codimensions of the  $W_i$ .
- c. Let  $W_1, \dots, W_k \subset V$  be a collection  $\mathcal{C}$  of hyperplanes in  $V$ . Let  $I, J \subseteq \{1, \dots, k\}$  be subsets. The collection  $\mathcal{C}$  of hyperplanes is said to be in *general position* if  $\bigcap_{i \in I} W_i \supset \bigcap_{j \in J} W_j$  are not equal whenever  $I \subset J$  are not equal. Show that it is always true that  $W = \bigcap_{1 \leq i \leq k} W_i$  has finite codimension in  $V$ . In case the collection  $\mathcal{C}$  is in general position, compute the codimension of  $W$  in  $V$ .

**BaseCoord 21.** The dimension formula

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$$

for finite-dimensional subspaces is analogous to the inclusion-exclusion formula

$$|S_1 \cup S_2| = |S_1| + |S_2| - |S_1 \cap S_2|$$

for sets. For three sets, the inclusion-exclusion formula is:

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + |S_1 \cap S_2 \cap S_3|$$

Is the analogous formula for three finite-dimensional subspaces  $W_1, W_2, W_3$  also true? Prove or provide a counterexample. By analyzing the map

$$W_1 \oplus W_2 \oplus W_3 \longrightarrow W_1 + W_2 + W_3$$

given by

$$(w_1, w_2, w_3) \mapsto w_1 + w_2 + w_3,$$

can you prove that

$$\dim(W_1 + W_2 + W_3) = \dim(W_1) + \dim(W_2) + \dim(W_3) - \dim(W_1 \cap W_2) - \dim((W_1 + W_2) \cap W_3) ?$$

**BaseCoord 22.** Let  $F \subseteq K$  be fields. Let  $V$  be a vector space over  $K$ .

- a. Explain why  $V$  is also a vector space over  $F$ .
- b. If  $\{e_1, \dots, e_n\}$  is a basis for  $K$  over  $F$  and if  $\mathcal{B} = \{v_1, \dots, v_m\}$  is a basis for  $V$  over  $K$ , show that  $\mathcal{A} = \left\{ e_i v_j \mid 1 \leq i \leq n, 1 \leq j \leq m \right\}$  is a basis for  $V$  over  $F$ . This yields the following formula

$$\dim_F V = (\dim_F K) \cdot (\dim_K V)$$

where the subscript on  $\dim$  denotes the field over which the dimension is computed.

- c. For the particular case of the reals contained in the complexes give formulas for the dimensions of the following over both fields:
- i.  $\mathbb{C}^{m \times n}$ ,
  - ii. all polynomials of degree less than  $n$  (include 0) with complex coefficients,
  - iii. all  $n \times n$  symmetric matrices with complex coefficients.
- d. Let  $S : V \rightarrow V$  be a linear transformation on the vector space  $V$  over  $K$ . Explain why  $S$  is also a linear transformation over  $F$ . Assume that the matrix of  $S$  with respect to the basis  $\mathcal{B}$  has entries  $a_{ij}$  for  $1 \leq i, j \leq m$ . Choose an appropriate ordering for the basis  $\mathcal{A}$  and find the matrix of  $S$  considered as a linear transformation over  $F$ . (Note that the matrix may be easier to describe if you choose a nice order for the basis. Hint: Use block matrices!)

**BaseCoord 23.** Let  $\mathbb{F}_q$  be a finite field with  $q$  elements and having characteristic  $p$ . Prove the assertions you make to answer the following questions.

- a. Let  $m, n$  be positive integers. Give a formula for the number of elements in  $\mathbb{F}_q^n$  and  $\mathbb{F}_q^{m \times n}$ .
- b. Give a formula for the number of different ordered bases of  $\mathbb{F}_q^n$ .
- c. Give a formula for the number of invertible matrices in  $\mathbb{F}_q^{n \times n}$ .
- d. Give a formula for the number of ordered, linearly independent sequences of vectors with  $m$  elements in  $\mathbb{F}_q^n$ .
- e. Give a formula for the number of matrices of  $\mathbb{F}_q^{m \times n}$  with rank  $m$ .
- f. For a non-negative integer  $k$  determine the number of subspaces of dimension  $k$  of  $\mathbb{F}_q^n$ .
- g. For a non-negative integer  $k$  determine the number of ordered sequences of  $k$  vectors in  $\mathbb{F}_q^n$  that span.

**BaseCoord 24.** Let  $V$  be a finite dimensional vector space over the field  $F$ . Let  $\mathcal{A}$  and  $\mathcal{B}$  be any bases for  $V$ . Assume we are given a decomposition of the basis  $\mathcal{A} = \mathcal{A}_1 \dot{\cup} \mathcal{A}_2$  of the first basis into two disjoint subsets. Show that there is a way to choose a similar decomposition of the second basis  $\mathcal{B}$  to obtain four bases for  $V$ :

$$\begin{aligned}\mathcal{A} &= \mathcal{A}_1 \dot{\cup} \mathcal{A}_2 \\ \mathcal{B} &= \mathcal{B}_1 \dot{\cup} \mathcal{B}_2 \\ \mathcal{C} &= \mathcal{A}_1 \dot{\cup} \mathcal{B}_2 \\ \mathcal{D} &= \mathcal{B}_1 \dot{\cup} \mathcal{A}_2 .\end{aligned}$$

The Notes for the course *Math 4330, Honors Linear Algebra* at Cornell University have been developed over the last ten years or so mainly by the following (in chronological order):

Gerhard O. Michler

R. Keith Dennis

Martin Kassabov

W. Frank Moore

and

Yuri Berest.

Most sections have been revised so many times the original author may no longer recognize it. The intent is to provide a modern treatment of linear algebra using consistent terminology and notation. Some sections are written simply to provide a central source of information such as those on “Useful Definitions”, “Subobjects”, and “Universal Mapping Properties” rather than as a chapter as one might find in a traditional textbook. Additionally there are sections whose intent is to provide proofs of some results which are not given in the lectures, but rather provide them as part of a more thorough development of a tangential topic (e.g., Zorn’s Lemma to develop cardinal numbers and the existence of bases and dimension in the general case).

A large number of challenging exercises from many different sources have been included. Although most should be readily solvable by students who have mastered the material, a few even more challenging ones still remain.

Much still remains to be done. Corrections and suggestions for additional exercises, topics and supplements are always welcome.