Infinite bases, and unique expressions for vectors

1. Background. When a vector space V has a finite basis $\beta = \{v_1, \ldots, v_n\}$, we know that each element of V has a unique expression $a_1 v_1 + \ldots + a_n v_n$ $(a_i \in F)$. Though some elements of V can be written as linear combinations of proper subsets of our basis, we can always rewrite such expressions in the above form by throwing in the absent elements with zero coefficients attached. If we included expressions written with different subsets of our basis, then our expressions for elements would be nonunique, but by insisting that all elements appear, we get uniqueness.

An alternative way to get uniqueness is to allow representations in terms of subsets of the basis, but require all coefficients in these representations to be nonzero. This, however, has the disadvantage that the description of the expression for a sum $x+y$ requires subdivision into many cases: for each basis element, one must ask whether it appears in neither summand, in one summand but not in the other, in both, but with coefficients that cancel in the sum, or in both, with coefficients that do not cancel, before one can say whether it appears, and if so, with what coefficient, in the sum.

When a vector space V is infinite-dimensional, we do not have the option of writing representations involving all the elements of a basis, so it appears that to get uniqueness, we must insist on expressions with all coefficients nonzero, and put up with the disadvantages of that approach. This is done by the authors of our text: in §1.7, Exercise 5 they ask us to prove that elements of V have unique representations of that form. The awkwardness of arguments involving linear combinations of such expressions is probably the reason why they have not proved (or asked us to prove) the infinite case of the most important property of a basis of a vector space: Theorem 2.6, saying that a linear transformation $V \to W$ can be determined uniquely by specifying its action on any basis of V.

On the other hand, §2.4, Exercise 25, asks us to prove that if S is a basis for V, and $\mathscr{C}(S,F)$ is the vector space of F-valued functions f on S such that $f(s) = 0$ for all but finitely many $s \in S$, then a map they write $\Psi\colon \mathscr{C}(S,F) \to V$ is an isomorphism. They describe Ψ as taking each nonzero f to the linear combination of elements of S in which each s appears with coefficient $f(s)$ if $f(s) \neq 0$, and does not appear if $f(s) = 0$. Given the unique representations for members of V proved in §1.7, Exercise 5, it is easy to see that Ψ is bijective; but the proof that Ψ is linear again requires messy case-by-case arguments; which is why I deleted Exercise 25 from the homework.

(Another way one might like to do that exercise would be to note that the elements of $\mathscr{C}(S,F)$ with value 1 at one point, and 0 at all other points, form a basis, and then to map $\mathscr{C}(S,F)$ to V by the unique linear map sending each such basis element to the corresponding basis element of V, and show that the resulting map equals the bijection described above, hence is an isomorphism. But that assumes the infinite-dimensional case of Theorem 2.6, which, as noted, the authors do not prove.)

Many mathematicians would ignore these problems, saying, for instance, regarding Theorem 2.6 for arbitrary vector spaces, that the map sending each linear combination of elements of a basis of V to the corresponding linear combination of any specified family of elements of W is ''clearly'' well-defined.

Below, I will give an approach which gets the same results by a more explicit argument, with what I hope is a minimal amount of messiness, by setting up a definition that captures the way one intuitively looks at linear combinations of elements taken from an infinite family.

2. Expressions and coefficients. It will be most convenient to start with a general family S of elements of a vector space V, and only later specialize to the case where S is a basis. So consider a family

$$S = \{s_i : i \in \Lambda\},$$

of elements of V, where Λ is any index-set. By a *linear expression* in elements of S, we shall mean an expression $a_1 s_{i_1} + \ldots + a_n s_{i_n}$, where $i_1, \ldots, i_n$ are distinct elements of Λ, and each a_m is an element of F (possibly 0). Note that different *expressions* can determine the same *value* in V, either because of linear relations satisfied by the elements of S, or because the expressions in question differ by the presence or absence of elements with zero coefficients.

Definition. (a) *If* $a_1 s_{i_1} + \ldots + a_n s_{i_n}$ *is a linear expression in the elements of* S, *and* i *is any element of* Λ, *then the* coefficient *of* s_i *in the above expression will be defined to be* a_i *if* $i \in \{a_1, \ldots, a_n\}$, *and to be* 0 *if* $i \notin \{a_1, \ldots, a_n\}$.

(b) *Two linear expressions in the elements of* S *will be called* trivially equivalent *if for each* $i \in \Lambda$, *the coefficient of* a_i *is the same in both expressions.*

Thus if, following §2.4, Exercise 25, we write $\mathscr{C}(\Lambda, F)$ for the vector space of functions $\Lambda \to F$ taking nonzero values at only finitely many points, then for any linear expression $a_1 s_{i_1} + \ldots + a_n s_{i_n}$ in the elements of S, we get an element of $\mathscr{C}(\Lambda, F)$ taking each $s \in S$ to its coefficient in the given expression. Clearly, two expressions will be trivially equivalent if and only if they yield the same element of $\mathscr{C}(\Lambda, F)$.

Lemma. *Let* S, *as above, be an indexed family of elements of* V. *Then*

(i) *Any two trivially equivalent expressions in the elements of* S *have the same* value *in* V (*i.e., when multiplied out and summed they give the same element of* V).

(ii) *The function* Ψ *associating to every element* $f \in \mathscr{C}(\Lambda, F)$ *the common value of all expressions* *whose family of coefficients is the function* f *is a linear map* $\mathcal{C}(\Lambda, F) \to \mathsf{V}$.

(iii) S *is a basis of* V *if and only if the above map is an isomorphism.*

Proof. (i) If two expressions are trivially equivalent, then they differ only in the presence or absence of certain terms consisting of elements with coefficient 0. These terms do not change the element of V that we get by multiplying out and summing; so trivially equivalent expressions yield the same value in V.

(ii) The one verification that takes some thought is that Ψ respects addition. Suppose $f, g \in \mathcal{C}(\Lambda, F)$. Let $\{i_1, \dots, i_n\}$ be a nonempty set of elements of Λ which includes the union of the finite set of $i \in \Lambda$ such that $f(s_i) \neq 0$ and the finite set of $i \in \Lambda$ such that $g(s_i) \neq 0$. Then we see that

$$\Psi(f) = f(i_1)\, s_{i_1} + \dots + f(i_n)\, s_{i_n}, \qquad \Psi(g) = g(i_1)\, s_{i_1} + \dots + g(i_n)\, s_{i_n},$$

$$\Psi(f+g) = (f+g)(i_1)\, s_{i_1} + \dots + (f+g)(i_n)\, s_{i_n}.$$

The first two expressions sum to the second, so $\Psi(f) + \Psi(g) = \Psi(f+g)$.

The other verification required, that Ψ respects scalar multiplication, is immediate.

(iii) Recall that, as I said in class, the statement that a *family* S indexed by Λ is a basis of V means that it spans V, and is linearly independent as an *indexed family;* where the latter condition means that the only relations $a_1 s_{i_1} + \dots + a_n s_{i_n} = 0$ with distinct elements $i_1, \dots, i_n \in \Lambda$ that hold in V are the trivial relations, where all $a_m = 0$.

Now it is clear that S spans V if and only if Ψ is surjective. It remains to show that S is linearly independent as an indexed family if and only if Ψ is one-to-one, equivalently, that it is linearly dependent if and only if Ψ is not one-to-one. Linear dependence is equivalent to the existence of a nontrivial linear relation $a_1 s_{i_1} + \dots + a_n s_{i_n} = 0$ holding in V; and the coefficients of such a relation correspond to a nonzero element of $\mathcal{C}(\Lambda, F)$ in the null space of Ψ, giving the desired equivalence. □

Statement (iii) of the above lemma gives §2.4, Exercise 25. We can now also get

Corollary. (Theorem 2.6 for arbitrary bases) *Let* V *and* W *be vector spaces over* F, *suppose* $S = \{s_i : i \in \Lambda\}$ *is an indexed basis of* V, *and let* $\{w_i : i \in \Lambda\}$ *be an indexed family of elements of* W. *Then there exists a unique linear transformation* $\mathsf{T}: \mathsf{V} \to \mathsf{W}$ *such that* $\mathsf{T}(s_i) = w_i$ *for all* $i \in \Lambda$.

Proof. Existence: Let $\Psi_{\mathsf{V}}: \mathcal{C}(\Lambda, F) \to \mathsf{V}$ and $\Psi_{\mathsf{W}}: \mathcal{C}(\Lambda, F) \to \mathsf{W}$ be defined as in statement (ii) of the lemma, the first using the family $\{s_i : i \in \Lambda\}$ and the second the family $\{w_i : i \in \Lambda\}$. Since S is assumed an indexed basis, statement (iii) of the lemma shows that Ψ_{V} is invertible. Let $\mathsf{T} = \Psi_{\mathsf{W}}\, \Psi_{\mathsf{V}}^{-1}: \mathsf{V} \to \mathsf{W}$. We see that Ψ_{V}^{-1} carries each s_i to the function with value 1 at i and 0 elsewhere, and that Ψ_{W} carries this to w_i, giving the desired behavior.

Uniqueness: It suffices to show that any two linear transformations $\mathsf{T}, \mathsf{U}: \mathsf{V} \to \mathsf{W}$ which agree on S are equal; I claim that this follows from the fact that S spans V. Indeed if T and U agree on S, this says that S lies in the null space of $\mathsf{T} - \mathsf{U}$. Hence that null space includes the span of S (the least subspace of V containing S), hence it is all of V, so T and U agree everywhere. □

3. Remarks. (i) A variant of the above approach is the following. Given any index set Λ and vector space V, we could look at expressions $v_{i_1} + \dots + v_{i_n}$ $(v \in \mathsf{V})$ indexed by finite subsets $\{i_1, \dots, i_n\}$ of Λ, and call two such expressions ‘‘trivially equivalent’’ if they differ only by the presence or absence of elements of Λ indexing terms v_i equal to zero. Applying to such expressions the same approach used above, we get a map sending every Λ-tuple $(v_i)_{i \in \Lambda}$ of elements of V with all but finitely many components 0 to an element of V, namely the common sum of all finite subfamilies which include all the nonzero terms of the given tuple. It is natural to call this element the sum $\Sigma_{i \in \Lambda}\, v_i$. This gives us a way of defining ‘‘sums’’ of infinite families, *if* all but finitely many of the summands are 0.

With this definition, we can prove that $S = \{s_i : i \in \Lambda\}$ is a basis of V if and only if every element of V has a unique expression as a sum $\Sigma_{i \in \Lambda}\, a_i s_i$, with all but finitely many $a_i = 0$. This is the approach I like best; but at the undergraduate level there is the danger of students’ confusing these ‘‘infinite sums’’ with the convergent series of analysis, and forgetting that the infinite sums we are talking about here are an ad hoc definition, not a given, and are only defined when all but finitely many of the summands are 0.

(ii) I have followed our text in assuming that a linear expression $a_1 s_{i_1} + \dots + a_n s_{i_n}$ must have $n \geq 1$. Recall that a sum of more than 2 terms is, by convention, defined in terms of an iterated application of the binary operation of addition. I like to add to this the convention that the sum of zero elements of a vector space is the element 0. When we make that convention, various statements, such as §1.7, Exercise 5, become slightly nicer, in that one does not have to restrict the statement to nonzero vectors.

(iii) The development given above hangs on an intuitive understanding of the concept of an ‘‘expression’’, as distinct from the element it represents. (However, after the development is complete, that concept of an expression may be replaced by the precise concept of an element of $\mathcal{C}(\Lambda, R)$.)