Math 121A — Linear Algebra

Neil Donaldson

Winter 2022

1 Vector Spaces

1.1 Introduction: What is Linear Algebra and why should we care?

Linear algebra is the study of *vector spaces* and *linear maps* between them. We'll formally define these concepts later, though they should be familiar from a previous class.
A function, or map, $\mathrm{T}: V \to W$ between vector spaces is *linear* if for all vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ and all scalars λ, we have the properties:

(a) $\mathrm{T}(\mathbf{v}_1 + \mathbf{v}_2) = \mathrm{T}(\mathbf{v}_1) + \mathrm{T}(\mathbf{v}_2)$

(b) $\mathrm{T}(\lambda \mathbf{v}_1) = \lambda \mathrm{T}(\mathbf{v}_1)$

Examples 1.1. You have seen many examples of these in your mathematical career.

1. $\mathrm{T}(x) = 3x$ defines a linear map $\mathrm{T}: \mathbb{R} \to \mathbb{R}$.

   More generally, $V = \mathbb{R}^m$, $W = \mathbb{R}^n$, with T being multiplication by a real $n \times m$ matrix.

2. Differentiation: Let $\mathrm{T} = \frac{\mathrm{d}}{\mathrm{d}x}$ be the usual differential operator and V the vector space of differentiable functions $f: \mathbb{R} \to \mathbb{R}$. More generally, T could be a *linear differential operator* such as $\mathrm{T} = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + 2x\frac{\mathrm{d}}{\mathrm{d}x} + x^2 + 1$ whence

   $$\mathrm{T}(y) = y'' + 2xy' + (x^2 + 1)y$$

   The standard methods for solving *linear differential equations* seen in a lower-division class are based on linear algebra.

3. Integration: let V be a vector space of integrable functions then $\mathrm{T}(f) = \int_a^x f(t)\,\mathrm{d}t$ defines a linear map to a vector space of continuous functions.

The ubiquity of linear structures is one reason to study linear algebra. Another is that *linear problems* often admit systematic techniques that give us at least a fighting chance of finding a solution. By contrast, *non-linear* problems are typically much more difficult: if such can be solved, it is often due to some trickery or luck.

What makes linear problems easy? The core idea is to use simple solutions as building blocks to construct more complex solutions. Here is a, hopefully familiar, example.

Example 1.2. The power law tells us how to integrate monomials. The linearity of integration allows us to combine these building blocks to compute the integral of any polynomial:

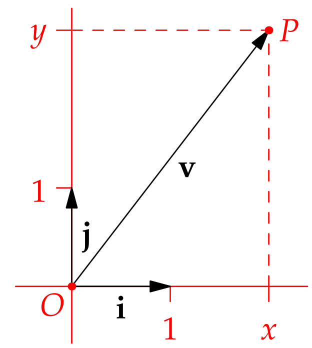
$$\begin{aligned}\int x^2 + 5x^3 \,\mathrm{d}x &= \int x^2 \,\mathrm{d}x + 5\int x^3 \,\mathrm{d}x && \text{(linearity)}\\ &= \frac{1}{3}x^3 + \frac{5}{4}x^4 + c && \text{(power law)}\end{aligned}$$

By contrast, the integration of *products* is a non-linear problem. The fact that

$$\int e^x \sin x \,\mathrm{d}x \neq \left[\int e^x \,\mathrm{d}x\right]\left[\int \sin x \,\mathrm{d}x\right]$$

and the resulting need for integration-by-parts is a major source of difficulty in freshman calculus.

A Brief Review of $\mathbb{R}^2$ and $\mathbb{R}^3$ In these standard spaces, we often visualize vectors as arrows. In the picture, the *vector* $\mathbf{v}$ points from the *origin* O with *co-ordinates* $(0,0)$ to the point $P = (x,y)$. Writing $\mathbf{i}$, $\mathbf{j}$ for the standard basis vectors, there are several common notations for $\mathbf{v}$:

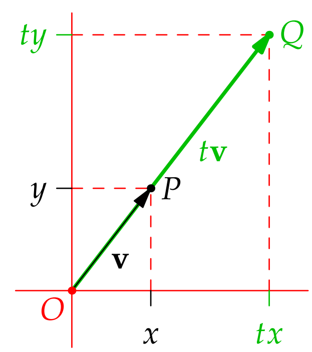


$$\mathbf{v} = \overrightarrow{OP} = \begin{pmatrix} x \\ y \end{pmatrix} = x\mathbf{i} + y\mathbf{j} = x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Column vector notation helps distinguish a *vector* from a *point* (x,y): we call x and y the *components* of the column vector $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$.

The *vector space* $\mathbb{R}^2$ is simply the set of all such vectors. There is no need for a vector to have its tail at the origin, only direction and magnitude matter. In $\mathbb{R}^3$ things are similar, a point has three co-ordinates and we need the three standard basis vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$.

Scalar multiplication involves lengthening or contracting a vector by a real multiple: the vector $t\mathbf{v}$ has components tx and ty and we write



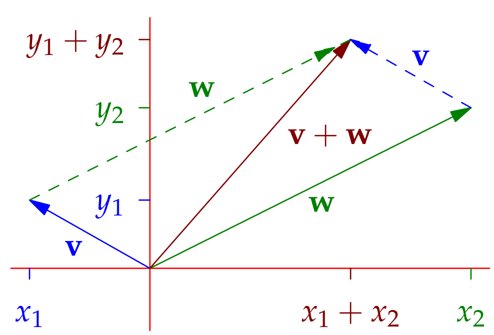
$$t\mathbf{v} = \overrightarrow{OQ} = \begin{pmatrix} tx \\ ty \end{pmatrix} = tx\mathbf{i} + ty\mathbf{j}$$

Note that if $t < 0$, then $t\mathbf{v}$ points in the opposite direction to $\mathbf{v}$.

Vector addition is defined by the *parallelogram law.* Simply add components: if $\mathbf{v}_1 = \left(\begin{smallmatrix} x_1 \\ y_1 \end{smallmatrix}\right)$ and $\mathbf{v}_2 = \left(\begin{smallmatrix} x_2 \\ y_2 \end{smallmatrix}\right)$, then

$$\mathbf{v}_1 + \mathbf{v}_2 := \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} = (x_1 + x_2)\mathbf{i} + (y_1 + y_2)\mathbf{j}$$

The intuitive *nose-to tail* interpretation of vector addition is immediate.

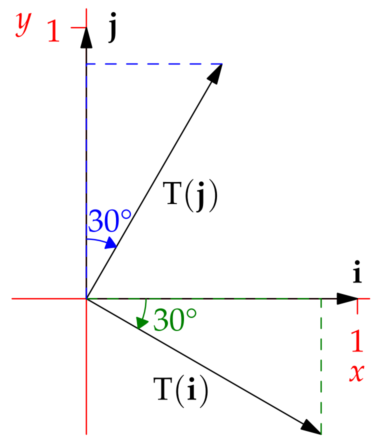


Example 1.3 (Rotation and the standard basis). We finish by reviewing an approach you should have seen in a previous course. By considering how to transform a basis, we obtain a complete formula for a linear map.

Consider the function $T:\mathbb{R}^2\to\mathbb{R}^2$ which rotates a point $30°$ clockwise around the origin. You should believe, though a proof is tricky at the present, that T is indeed linear. To discover a formula for T it is enough to consider what it does to the standard basis $\{\mathbf{i},\mathbf{j}\}$ of $\mathbb{R}^2$.

This is because if $\mathbf{v}=\binom{x}{y}$ is any vector, then, by linearity

$$\mathbf{v}=x\mathbf{i}+y\mathbf{j} \implies T(\mathbf{v})=xT(\mathbf{i})+yT(\mathbf{j})$$



Using the picture and a little trigonometry, you should be convinced that that

$$T(\mathbf{i})=\begin{pmatrix}\cos 30°\\ -\sin 30°\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{3}}{2}\\ -\frac{1}{2}\end{pmatrix}\qquad T(\mathbf{j})=\begin{pmatrix}\sin 30°\\ \cos 30°\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\\ \frac{\sqrt{3}}{2}\end{pmatrix}$$

from which we obtain

$$T\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{3}}{2}x+\frac{1}{2}y\\ -\frac{1}{2}x+\frac{\sqrt{3}}{2}y\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{3}}{2} & \frac{1}{2}\\ -\frac{1}{2} & \frac{\sqrt{3}}{2}\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

where we've written the final expression as a matrix multiplication.[a]

[a] This is one of the advantages of column vector notation. Indeed one of the major goals of the course is to see that every linear map between finite-dimensional vector spaces can be represented in such a fashion.

Because of linearity, we were able to completely determine T merely by understanding how it acted on the basis vectors $\mathbf{i}$ and $\mathbf{j}$. This property is *not* shared by non-linear functions. For instance, if $|\mathbf{v}|$ returns the length of a vector $\mathbf{v}\in\mathbb{R}^2$, then the function

$$f:\mathbb{R}^2\to\mathbb{R}^2:\mathbf{v}\mapsto\left(|\mathbf{v}|^2+1\right)\mathbf{v} \tag{$*$}$$

is non-linear. Simply knowing that $f(\mathbf{i})=2\mathbf{i}$ and $f(\mathbf{j})=2\mathbf{j}$ is insufficient to completely understand the function.

Exercises 1.1

1. Using the same approach as in Example 1.3, explicitly find a formula for the linear map $T:\mathbb{R}^2\to\mathbb{R}^2$ which *reflects* in the line $y=x$.
2. A linear map is not the same thing as a straight line! Explain why the function $f:x\mapsto 3x+2$ is non-linear.
3. Give a reason why the function $f:\mathbb{R}^2\to\mathbb{R}^2$ defined in $(*)$ above is non-linear.
4. (a) Give an algebraic proof (use components!) of the distributive law $\lambda(\mathbf{v}+\mathbf{w})=\lambda\mathbf{v}+\lambda\mathbf{w}$ in the vector space $\mathbb{R}^2$.

   (b) Give a *pictorial* argument for the distributive law?

   (*Hint: Consider similar triangles or parallelograms and channel your inner Euclid...*)

1.2 Vector Spaces: Basic Results, Examples and Subspaces

Vector spaces generalize the intuitive structure of $\mathbb{R}^2$, where identities such as commutativity

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$$

are geometrically obvious. The *axioms* of a vector space merely assert that such identities hold generally.

Definition 1.4. Let V be a non-empty set (elements *vectors*) and $\mathbb{F}$ a *field* (elements *scalars*), and suppose we have two operations:

Vector Addition If $\mathbf{v}$ and $\mathbf{w}$ are vectors, we can form the *sum* $\mathbf{v} + \mathbf{w}$.

Scalar Multiplication If $\mathbf{v}$ is a vector and λ a scalar, we can form the *product* $\lambda\mathbf{v}$.

We say that V is a *vector space* over $\mathbb{F}$ if the following axioms are satisfied:[a]

G1: Closure under addition	$\forall \mathbf{v}, \mathbf{w} \in V, \quad \mathbf{v} + \mathbf{w} \in V$
G2: Associativity of addition	$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \quad (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
G3: Identity for addition	$\exists \mathbf{0} \in V, \quad \text{such that} \quad \forall \mathbf{v} \in V, \quad \mathbf{v} + \mathbf{0} = \mathbf{v}$
G4: Inverse for addition	$\forall \mathbf{v} \in V, \exists -\mathbf{v} \in V, \quad \text{such that} \quad \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$
G5: Commutativity of addition	$\forall \mathbf{v}, \mathbf{w} \in V, \quad \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$
A1: Closure under scalar multiplication	$\forall \mathbf{v} \in V, \lambda \in \mathbb{F}, \quad \lambda\mathbf{v} \in V$
A2: Identity for scalar multiplication	$\forall \mathbf{v} \in V, \quad 1\mathbf{v} = \mathbf{v}$
A3: Action of scalar multiplication	$\forall \lambda, \mu \in \mathbb{F}, \mathbf{v} \in V, \quad \lambda(\mu\mathbf{v}) = (\lambda\mu)\mathbf{v}$
D1: Distributivity I	$\forall \mathbf{v}, \mathbf{w} \in V, \lambda \in \mathbb{F}, \quad \lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$
D2: Distributivity II	$\forall \mathbf{v} \in V, \lambda, \mu \in \mathbb{F}, \quad (\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$

[a] This is easier to remember if you've studied group theory: the 'G' axioms say that $(V, +)$ is an Abelian group, the 'A' axioms say that the field $\mathbf{F}$ has a *left action* on V. The distributivity axioms explain how the two operations interact.

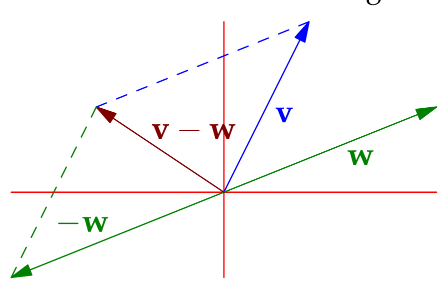
Notation You can use another notation (e.g. $\vec{v}$ or $\underline{v}$) for abstract vectors, but use something: distinguishing vectors and scalars helps avoid common mistakes like dividing by a vector. This notation might not be appropriate in certain examples, (e.g. polynomials, matrices) so take extra care.

Fields A *field* $\mathbb{F}$ is a set which behaves very like the real numbers under addition and multiplication. In almost all examples, $\mathbb{F}$ will be either the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$. The symbols 0 and 1 (e.g. axiom A2) refer to the additive and multiplicative identities in $\mathbb{F}$. Be careful to distinguish the scalar $0 \in \mathbb{F}$ from the *zero vector* $\mathbf{0} \in V$.

Inverses and subtraction Subtraction of vectors is taken to mean *addition of the inverse*, namely

$$\mathbf{v} - \mathbf{w} := \mathbf{v} + (-\mathbf{w})$$

In $\mathbb{R}^2$ this can be viewed pictorially.



Essentially every example we will encounter falls into one of two classes.

Theorem 1.5 (Matrices & Sets of Functions). *Let $\mathbb{F}$ be a field.*

1. *The set $M_{m\times n}(\mathbb{F})$ of $m \times n$ matrices with entries in $\mathbb{F}$*

$$M_{m\times n}(\mathbb{F}) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} : a_{ij} \in \mathbb{F} \right\}$$

forms a vector space over $\mathbb{F}$ under component-wise addition and scalar multiplication: given matrices $A = (a_{ij})$ and $B = (b_{ij})$ and $\lambda \in \mathbb{F}$, the ij^{th} entries of the matrices $A + B$ and λA are

$$(A+B)_{ij} := a_{ij} + b_{ij}, \qquad (\lambda A)_{ij} := \lambda a_{ij} \tag{$*$}$$

2. *Let D be a set and V a vector space over $\mathbb{F}$. The set of functions*

$$\mathcal{F}(D,V) = \{f : D \to V\}$$

forms a vector space over $\mathbb{F}$ with addition and scalar multiplication defined by[a]

$$(f+g)(x) := f(x) + g(x), \qquad (\lambda f)(x) := \lambda(f(x))$$

[a] The *function* $f + g \in \mathcal{F}(D,V)$ is defined by what it does to an element $x \in D$. In particular, $f(x) \in V$ is *not* a function. However, it is acceptable to write 'the function $f(x)$', just make sure you know that this is an abuse of notation.

To prove the theorem, each axiom (G1–5, A1–3, D1,2) should be checked explicitly for each part of the theorem: this is tedious! For instance, axiom D2 may be verified for matrices as follows:

$$\big((\lambda+\mu)A\big)_{ij} = (\lambda+\mu)a_{ij} = \lambda a_{ij} + \mu a_{ij} = (\lambda A)_{ij} + (\mu A)_{ij}$$

Definitions ($*$) provide the red equalities, while the blue is distributivity in the field $\mathbb{F}$.

Examples 1.6. 1. The *column vectors* (*n-tuples*) are a special case: $\mathbb{F}^n := M_{n\times 1}(\mathbb{F})$. E.g., in $\mathbb{R}^3$

$$2\begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} + 7\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -8 \end{pmatrix} + \begin{pmatrix} -7 \\ 14 \\ 7 \end{pmatrix} = \begin{pmatrix} -5 \\ 14 \\ -1 \end{pmatrix}$$

2. In $M_{2\times 3}(\mathbb{C})$, we have

$$\begin{pmatrix} 1 & i & 0 \\ -3 & 1-i & 2+3i \end{pmatrix} + i\begin{pmatrix} 2 & -3 & 1 \\ 3-i & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1+2i & -2i & i \\ -2+3i & 1-i & 2+5i \end{pmatrix}$$

3. A field is a vector space over itself! In particular, $\mathcal{F}(\mathbb{R},\mathbb{R})$ is a vector space. For instance, if f, g are defined by $f(x) = x^2$ and $g(x) = \sin x$, then $4f - 2g$ is the function given by

$$(4f - 2g)(x) = 4x^2 - 2\sin x$$

We'll shortly restrict to certain types of functions (e.g. *continuous functions*, *differentiable functions*, *polynomials*) and see that these also form vector spaces.

Basic Results Here we gather several basic facts about vector spaces that you'll use without thinking. Since these are not axioms, they do require *proof*.

Lemma 1.7.
1. *Cancellation law:* $\mathbf{x}+\mathbf{z}=\mathbf{y}+\mathbf{z} \implies \mathbf{x}=\mathbf{y}$
2. *Uniqueness of Identity: The zero vector* $\mathbf{0}$ *posited in axiom G3 is unique.*
3. *Uniqueness of Inverse: Given* $\mathbf{v}\in V$*, the vector* $-\mathbf{v}$ *posited in axiom G4 is unique.*
4. *Scalar multiplication by zero:* $\forall \mathbf{v}\in V$*, we have* $0\mathbf{v}=\mathbf{0}$*.*
5. *Action of negatives:* $\forall \mathbf{v}\in V,\ \lambda\in\mathbb{F}$*, we have* $(-\lambda)\mathbf{v}=-(\lambda\mathbf{v})$*.*
6. *Action on zero vector:* $\forall \lambda\in\mathbb{F}$*, we have* $\lambda\mathbf{0}=\mathbf{0}$*.*

Proof. We prove number 4, leaving the remainder as exercises: they are easiest if tackled in order! Since $0=0+0$ in $\mathbb{F}$, apply axioms D2, G3, G5 and the cancellation law to see that

$$\begin{aligned} &0\mathbf{v}=(0+0)\mathbf{v}=0\mathbf{v}+0\mathbf{v} && \text{(Distributivity D2)}\\ \implies\ &\mathbf{0}+0\mathbf{v}=0\mathbf{v}+0\mathbf{v} && \text{(Identity G3 and Commutativity G5)}\\ \implies\ &\mathbf{0}=0\mathbf{v} && \text{(Cancellation law)} \end{aligned}$$

■

Subspaces As in other areas of algebra (subgroup, subring, subfield, etc.) the prefix *sub* means that an object is a *subset*, while retaining the algebraic structure of the original set.

Definition 1.8. Let V be a vector space over a field $\mathbb{F}$.

A non-empty subset $W\subseteq V$ is a *subspace* of V (written $W\le V$) if it is a vector space over the *same* field $\mathbb{F}$ with respect to the *same* addition and scalar multiplication operations as V.

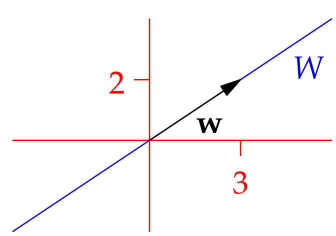
A subspace is *proper* if it is a proper subset (i.e., $W\neq V$).

The *trivial subspace* of V is the point set $\{\mathbf{0}\}$.

The subset approach allows us to quickly construct many more examples.

Example 1.9. Consider the line containing $\mathbf{w}=\binom{3}{2}\in\mathbb{R}^2$:

$$W:=\left\{a\mathbf{w}=\begin{pmatrix}3a\\2a\end{pmatrix}:a\in\mathbb{R}\right\}$$



It is almost trivial, if tedious, to check that W satisfies the axioms and is therefore a subspace of $\mathbb{R}^2$. For instance:

$$\text{G1:}\quad \begin{pmatrix}3a\\2a\end{pmatrix}+\begin{pmatrix}3b\\2b\end{pmatrix}=\begin{pmatrix}3(a+b)\\2(a+b)\end{pmatrix}\in W$$

by appealing to the distributivity laws in $\mathbb{R}$.

Thankfully, as the next theorem shows, there is no need to check all the axioms to determine when we have a subspace.

Theorem 1.10. *Suppose W is a non-empty subset of a vector space V over $\mathbb{F}$. Then W is a subspace of V if and only if it is closed under addition and scalar multiplication:*[a]

G1: $\forall \mathbf{w}_1, \mathbf{w}_2 \in W$, *we have* $\mathbf{w}_1 + \mathbf{w}_2 \in W$.

A1: $\forall \mathbf{w} \in W,\ \lambda \in \mathbb{F}$, *we have* $\lambda\mathbf{w} \in W$.

[a] That $\mathbf{w}_1 + \mathbf{w}_2$ and $\lambda\mathbf{w}$ lie in W and not just V is what makes these genuine conditions.

There are two common variants of this result: feel free to use these in examples if you prefer.

- Explicitly verify axiom G3 ($\mathbf{0}_V \in W$) instead of checking the non-emptiness of W. You still need to check that W is a subset of V!
- Combine the closure axioms into a single statement: $\forall \mathbf{x}, \mathbf{y} \in W,\ \lambda \in \mathbb{F}$, we have $\lambda\mathbf{x} + \mathbf{y} \in W$.

Proof. If W is a subspace, then it is a vector space and so axioms G1 and A1 hold as written.

Conversely, assume that W is a non-empty subset of V satisfying G1 and A1. Reread the axioms (Definition 1.4) and observe that all except perhaps G1, G3, G4 and A1 hold on any subset of V. Under our assumptions therefore, it remains only to verify G3 and G4.

G3: Choose any $\mathbf{w} \in W$. By Lemma 1.7 (part 4) and axiom A1 (for W!), we see that the zero vector of V satisfies

$$\mathbf{0}_V = 0\mathbf{w} \in W$$

Since $\mathbf{0}_V$ satisfies axiom G3 for V, it necessarily does on any subset: we therefore have $\mathbf{0}_W = \mathbf{0}_V$.

G4: Given $\mathbf{w} \in W$, let $-\mathbf{w} \in V$ be its additive inverse in V. Now observe that

$$-\mathbf{w} = (-1)\mathbf{w} \in W$$

by Lemma 1.7 (part 5) and axiom A1. ∎

Examples 1.11. 1. Returning to Example 1.9, recall that we already checked axiom G1. Moreover,

- $\mathbf{0} = 0\mathbf{w} \in W$ so that W is non-empty.
- A1: $\lambda(a\mathbf{w}) = (\lambda a)\mathbf{w} \in W$ by axiom A3.

so that W is a subspace of $\mathbb{R}^2$. Alternatively, if $\lambda \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in W$, then $\exists a, b \in \mathbb{R}$ for which

$$\lambda\mathbf{x} + \mathbf{y} = \lambda a\mathbf{w} + b\mathbf{w} = (\lambda a + b)\mathbf{w} \in W$$

2. For any field $\mathbb{F}$, let $P_n(\mathbb{F}) = \{a_0 + a_1x + \cdots + a_nx^n : a_i \in \mathbb{F}\}$ be the set of polynomials of degree $\leq n$ with coefficients in $\mathbb{F}$. By considering axioms G1 and A1, this is plainly a subspace of the space of functions $\mathcal{F}(\mathbb{F}, \mathbb{F})$:

$$\lambda(a_0 + a_1x + \cdots + a_nx^n) + (b_0 + b_1x + \cdots + b_nx^n) = (a_0 + b_0) + \cdots + (a_n + b_n)x^n$$

Non-emptiness is guaranteed by considering the *zero polynomial* $0(x) = 0 + 0x + \cdots + 0x^n$.

More generally, if $n \leq m$, then $P_n(\mathbb{F}) \leq P_m(\mathbb{F}) \leq P(\mathbb{F})$, where the last denotes the space of *all* polynomials of any degree.

3. If $U \subseteq \mathbb{R}$ is an interval, then $V = \mathcal{F}(U, \mathbb{R})$ is a vector space over $\mathbb{R}$. The subset $C(U, \mathbb{R})$ of continuous functions is a subspace of V. Indeed, as is verified in any analysis course,

   If $f, g : I \to \mathbb{R}$ are continuous and $\lambda \in \mathbb{R}$, then $\lambda f + g : I \to \mathbb{R}$ is continuous.

   This also extends to sets of differentiable functions, etc.

4. The *trace* $\operatorname{tr} : M_n(\mathbb{F}) \to \mathbb{F}$ of an $n \times n$ matrix is defined by summing the main diagonal:

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

   The subset of *trace-free* matrices is denoted

$$\mathfrak{sl}_n(\mathbb{F}) = \{A \in M_n(\mathbb{F}) : \operatorname{tr} A = 0\}$$

   It is easy to check that $\mathfrak{sl}_n(\mathbb{F}) \leq M_n(\mathbb{F})$:

$$\operatorname{tr}(\lambda A + B) = \sum_{i=1}^{n} \lambda a_{ii} + b_{ii} = \lambda \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii} = \lambda \operatorname{tr} A + \operatorname{tr} B = 0$$

Intersections and Direct Sums Since vector spaces are sets, we may take intersections...

Theorem 1.12. *If V and W are subspaces of some vector space U, then their* intersection *$V \cap W$ is a subspace of both V and W.*

Proof. Since V and W are subspaces of U, they both contain $\mathbf{0}$ and so $V \cap W$ is non-empty.
Now suppose $\mathbf{x}, \mathbf{y} \in V \cap W$ and $\lambda \in \mathbb{F}$. Since V and W are both vector spaces, they are closed under addition and scalar multiplication (in U!): in particular,

$$\mathbf{x} + \mathbf{y} \in V, \qquad \mathbf{x} + \mathbf{y} \in W, \qquad \lambda \mathbf{x} \in V, \qquad \lambda \mathbf{x} \in W$$

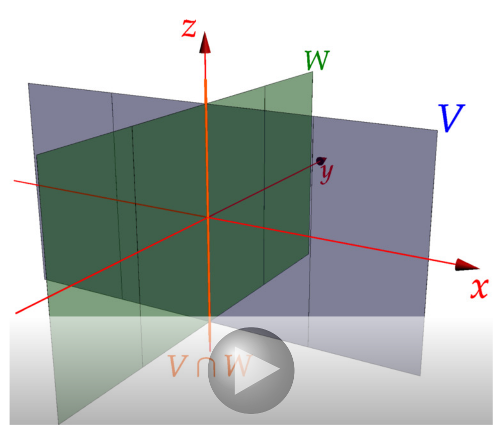
But then $\mathbf{x} + \mathbf{y} \in V \cap W$ and $\lambda \mathbf{x} \in V \cap W$, whence $V \cap W$ is closed and thus a subspace. ■

Example 1.13. Suppose that

$$V = \{x\mathbf{i} + z\mathbf{k} : x, z \in \mathbb{R}\}$$
$$W = \{y\mathbf{j} + z\mathbf{k} : y, z \in \mathbb{R}\}$$

are the xz- and yz-planes respectively. Plainly, V and W are subspaces of $\mathbb{R}^3$ with intersection the z-axis

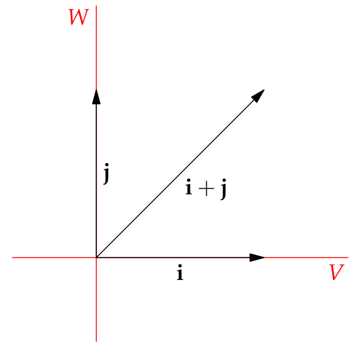
$$V \cap W = \{z\mathbf{k} : z \in \mathbb{R}\}$$



Attempting the same thing for unions results in a problem.
For a simple counterexample, let

$$V = \{x\mathbf{i} : x \in \mathbb{R}\} \qquad W = \{y\mathbf{j} : y \in \mathbb{R}\}$$

be the x- and y-axes in $\mathbb{R}^2$, whose intersection is the trivial subspace $V \cap W = \{\mathbf{0}\}$. Their union



$$V \cup W = \{x\mathbf{i}, y\mathbf{j} : x, y \in \mathbb{R}\}$$

is not a subspace of $\mathbb{R}^2$ since it is not closed under addition:

$$\mathbf{i} \in V \text{ and } \mathbf{j} \in W \text{ but } \mathbf{i}+\mathbf{j} \notin V \cup W$$

Instead we search for the *smallest* vector space containing $V \cup W$.

Definition 1.14. Suppose V and W are subspaces of U. Their *sum* is the set

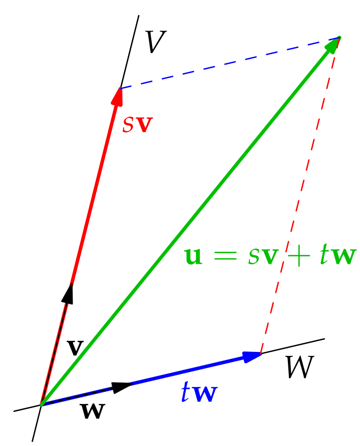
$$V + W := \{\mathbf{v} + \mathbf{w} : \mathbf{v} \in V, \mathbf{w} \in W\}$$

In addition, if $V \cap W = \{\mathbf{0}\}$, we call this the *direct sum* and write $V \oplus W$.

Examples 1.15. 1. The x- and y-axes, $V = \{x\mathbf{i} : x \in \mathbb{R}\}$ and $W = \{y\mathbf{j} : y \in \mathbb{R}\}$ are clearly subspaces of $\mathbb{R}^2$ with trivial intersection. It is immediate that

$$V \oplus W = \{x\mathbf{i} + y\mathbf{j} : x, y \in \mathbb{R}\} = \mathbb{R}^2$$

2. More generally, let $V = \{s\mathbf{v} : s \in \mathbb{R}\}$ and $W = \{t\mathbf{w} : t \in \mathbb{R}\}$ be distinct, non-trivial subspaces of $\mathbb{R}^2$ (i.e. $\mathbf{v}, \mathbf{w}$ are non-parallel). Observe:



- If $V \cap W \neq \{\mathbf{0}\}$, then $\exists s, t \neq 0$ such that $s\mathbf{v} = t\mathbf{w}$, whence $\mathbf{v}$ and $\mathbf{w}$ would be parallel: contradiction.
- Writing $\mathbf{v} = \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right)$ and $\mathbf{w} = \left(\begin{smallmatrix} c \\ d \end{smallmatrix}\right)$, we see that for any given $\mathbf{u} = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2$,

$$\begin{aligned} \mathbf{u} = s\mathbf{v} + t\mathbf{w} &\iff \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix} \\ &\iff \begin{pmatrix} s \\ t \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

Since $\mathbf{v}, \mathbf{w}$ are non-parallel, we are not dividing by zero: every $\mathbf{u}$ can be written in the form $s\mathbf{v} + t\mathbf{w}$ and so $\mathbb{R}^2 = V + W$.

Putting both parts together, we conclude that $\mathbb{R}^2 = V \oplus W$ is a direct sum.

Indeed we see that every $\mathbf{u} \in \mathbb{R}^2$ can be written *uniquely* in terms of the subspaces: as the next result shows, this is a defining property of direct sums.

Theorem 1.16. *Let V, W be subspaces of U with trivial intersection. Then:*

1. *$V \oplus W$ is a subspace of U.*
2. *V and W are subspaces of $V \oplus W$.*
3. *If X is a subspace of U such that both V and W are subspaces of X, then $V \oplus W \leq X$.*
4. *$U = V \oplus W \iff \forall \mathbf{u} \in U, \exists$unique $\mathbf{v} \in V$ and $\mathbf{w} \in W$ such that $\mathbf{u} = \mathbf{v} + \mathbf{w}$.*

The proofs are exercises. Note how the third property says that $V \oplus W$ is the smallest space containing V and W, while the fourth says that direct sums are synonymous with unique decompositions.

Exercises 1.2

1. Let $S = \{0, 1\}$ and $\mathbb{F} = \mathbb{R}$. In the vector space of functions $\mathcal{F}(S, \mathbb{R})$, let

   $$f(t) = 2t + 1, \qquad g(t) = 1 + 4t - 2t^2, \qquad h(t) = 5^t + 1$$

   Show that $f = g$ and $f + g = h$.

2. (a) If $p(x) = 2x + 3x^2$ and $q(x) = 4 - x^2$, compute the polynomial $p(x) + 3q(x)$.
   (b) Explain why the set of degree two polynomials with coefficients in $\mathbb{R}$ is *not* a vector space.
   (c) Prove explicitly that $P_1(\mathbb{R})$ is a subspace of $P_3(\mathbb{R})$.

3. Prove parts 2, 3 and 6 of Lemma 1.7.

4. Consider the vector space $\mathbb{C}^2 = \{\left(\begin{smallmatrix} w \\ z \end{smallmatrix}\right) : w, z \in \mathbb{C}\}$ over the field $\mathbb{C}$ of complex numbers.

   (a) Show that $\mathbf{v} = \left(\begin{smallmatrix} i \\ 2+3i \end{smallmatrix}\right)$ and $\mathbf{w} = \left(\begin{smallmatrix} 1+2i \\ 7+4i \end{smallmatrix}\right)$ are parallel.
   (b) $\mathbb{C}^2$ is automatically a vector space over $\mathbb{C}$. Prove that it is a vector space over $\mathbb{R}$.

5. Is $M_{m \times n}(\mathbb{R})$ a vector space over the rational numbers $\mathbb{Q}$? Explain.

6. Let $V = \{(a_1, a_2) : a_1, a_2 \in \mathbb{R}\}$. For $(a_1, a_2), (b_1, b_2) \in V$ and $\lambda \in \mathbb{R}$, define

   $$(a_1, a_2) + (b_1, b_2) := (a_1 + 2b_1, a_2 + 3b_2) \quad \text{and} \quad \lambda(a_1, a_2) = (\lambda a_1, \lambda a_2)$$

   Is V a vector space over $\mathbb{R}$ with respect to these operations? Explain.

7. Let $V = \mathbb{R}^2$, define vector addition as usual and scalar multiplication (by $\lambda \in \mathbb{R}$) by

   $$\lambda \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} \lambda x \\ \lambda^{-1} y \end{pmatrix} \quad \text{if } \lambda \neq 0 \qquad \text{or} \qquad \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{if } \lambda = 0$$

   Is V a vector space with respect to these operations? Why/why not?

8. Prove or disprove:

   (a) $V := \{\left(\begin{smallmatrix} 4a \\ -a \end{smallmatrix}\right) : a \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$.
   (b) $W := \{\left(\begin{smallmatrix} 4a+1 \\ -a \end{smallmatrix}\right) : a \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^2$.
   (c) $X := \left\{\left(\begin{smallmatrix} 4a+b \\ -a \\ 2a-b \end{smallmatrix}\right) : a, b \in \mathbb{R}\right\}$ is a subspace of $\mathbb{R}^3$.

9. Let V be the set of differentiable real-valued functions with domain $\mathbb{R}$. Prove that V is a subspace of the set of functions $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
   (*You may quote anything you like from elementary calculus without proof*)

10. With reference to Theorem 1.10, prove that properties G1 & A1 are equivalent to the combined closure property: $\forall \mathbf{x}, \mathbf{y} \in W, \lambda \in \mathbb{F}$, we have $\lambda \mathbf{x} + \mathbf{y} \in W$.

11. (a) Let $V = \{x\mathbf{i} : x \in \mathbb{R}\}$ and $W = \{y\mathbf{j} + z\mathbf{k} : y, z \in \mathbb{R}\}$ be subspaces of $\mathbb{R}^3$. Prove that $V \oplus W = \mathbb{R}^3$.
    (b) Repeat part (a) but this time with $V = \{x(\mathbf{i} + \mathbf{j}) : x \in \mathbb{R}\}$.

12. (a) Prove all four parts of Theorem 1.16.
    (b) If we drop the assumption that V and W have trivial intersection, which parts of Theorem 1.16 must be true for the *sum* $V + W$.

13. A matrix A is *symmetric* if it equals its transpose: $A^T = A$. It is *skew-symmetric* if $A^T = -A$. Let S be the set of symmetric matrices and K the set of skew-symmetric matrices in $M_2(\mathbb{R})$.
    (a) Show that S and K are subspaces of $M_2(\mathbb{R})$.
    (b) Prove or disprove: $M_2(\mathbb{R}) = S \oplus K$.
    (c) Does your argument extends to $M_n(\mathbb{R})$ and, if you've studied fields, to $M_n(\mathbb{F})$?

14. Let $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$ together with addition and multiplication modulo 5.
    (a) Prove that every non-zero element of $\mathbb{Z}_5$ has a multiplicative inverse ($\mathbb{Z}_5$ is a *field*): for all $x \in \mathbb{Z}_5 \setminus \{0\}$, there exists $y \in \mathbb{Z}_5$ such that $xy = 1$.
    (b) By part (a), $\mathbb{Z}_5^n$ is a vector space. Evaluate the expression $4\left(\begin{smallmatrix}3\\2\end{smallmatrix}\right) + 2\left(\begin{smallmatrix}1\\4\end{smallmatrix}\right) \in \mathbb{Z}_5^2$. For any $n \in \mathbb{N}$, how many vectors are there in $\mathbb{Z}_5^n$? (What is the cardinality of $\mathbb{Z}_5^n$?)

15. Let $V \times W = \{(\mathbf{v}, \mathbf{w}) : \mathbf{v} \in V, \mathbf{w} \in W\}$ be the Cartesian product of spaces V, W over $\mathbb{F}$.
    (a) (Briefly!) Argue that $V \times W$ is a vector space over $\mathbb{F}$ with respect to the operations
    $$(\mathbf{v}_1, \mathbf{w}_1) + (\mathbf{v}_2, \mathbf{w}_2) := (\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}_1 + \mathbf{w}_2) \qquad \lambda(\mathbf{v}, \mathbf{w}) := (\lambda\mathbf{v}, \lambda\mathbf{w})$$
    (b) Verify that $\hat{V} := \{(\mathbf{v}, \mathbf{0}_W) : \mathbf{v} \in V\}$ and $\hat{W} := \{(\mathbf{0}_V, \mathbf{w}) : \mathbf{w} \in W\}$ are subspaces of $V \times W$ and that $\hat{V} \oplus \hat{W} = V \times W$.
    ($V \times W$ *is an alternative definition of the direct sum* $V \oplus W$*, which should be familiar if you've seen* direct products *in group theory.*)

16. (Optional: should be familiar if you've studied group theory) Let W be a subspace of V over $\mathbb{F}$. For any $\mathbf{v} \in V$, define the *coset of* W *containing* $\mathbf{v}$ to be the set
    $$\mathbf{v} + W := \{\mathbf{v} + \mathbf{w} : \mathbf{w} \in W\}$$
    (a) If $V = \mathbb{R}^3$ and $W = \operatorname{Span}\{\mathbf{i}, \mathbf{j}\}$, describe the coset $\mathbf{k} + W$ in words.
    (b) Let V be a a general vector space. Prove that $\mathbf{v} + W$ is a subspace of V if and only if $\mathbf{v} \in W$.
    (c) Prove that $\mathbf{v}_1 + W = \mathbf{v}_2 + W \iff \mathbf{v}_1 - \mathbf{v}_2 \in W$.
    (d) Define the *quotient space* $V/W = \{\mathbf{v} + W : \mathbf{v} \in V\}$ to be the set of cosets of W in V together with the operations
    $$(\mathbf{v}_1 + W) + (\mathbf{v}_2 + W) := (\mathbf{v}_1 + \mathbf{v}_2) + W \qquad \lambda(\mathbf{v} + W) := \lambda\mathbf{v} + W$$
    Prove that addition and scalar multiplication are well-defined, and (briefly) convince yourself that V/W is a vector space over $\mathbb{F}$ under these operations.

1.3 Linear Combinations & Linear Independence

In this section we consider what vectors we can generate from a given collection using only the vector space operations af addition and scalar multiplication.

Definition 1.17. Let S be a non-empty subset of a vector space V over a field $\mathbb{F}$. A *linear combination* of vectors in S is any vector of the form

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \tag{$*$}$$

where each $\mathbf{v}_i \in S$ and each $a_i \in \mathbb{F}$. The *span* of S is the set of all linear combinations of vectors in S:

$$\operatorname{Span} S = \{a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n : n \in \mathbb{N}, a_1, \ldots, a_n \in \mathbb{F},\ \mathbf{v}_1, \ldots, \mathbf{v}_n \in S\}$$

By convention, $\operatorname{Span}\emptyset := \{\mathbf{0}\}$ is the trivial subspace.

Important: A linear combination contains *finitely many* terms—no infinite sums!

Our primary goal of this chapter is to identify the *smallest* possible spanning sets for a vector space: such a set will be called a *basis.* The full discussion is difficult and length; for the present, we consider a few simple examples of linear combinations and spanning sets.

Examples 1.18.
1. In $P_2(\mathbb{R})$, the vector $p(x) = 2 - 3x^2$ is a linear combination of the vectors $q(x) = 2x - x^2$ and $r(x) = 1 - x - x^2$, since

$$p = q + 2r$$

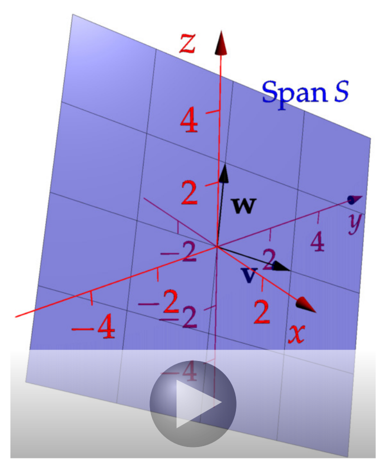
2. In Example 1.9, $W = \operatorname{Span}\{\mathbf{w}\}$. Since this is the span of a single vector, it is common to abuse notation and write $\operatorname{Span}\mathbf{w}$. In this notation, and following Definition 1.14, we see that

$$\operatorname{Span}\{\mathbf{v}, \mathbf{w}\} = \operatorname{Span}\mathbf{v} + \operatorname{Span}\mathbf{w}$$

3. Let $S = \{\mathbf{v}, \mathbf{w}\} \subseteq \mathbb{R}^3$ where $\mathbf{v} = \left(\begin{smallmatrix}1\\2\\-1\end{smallmatrix}\right)$ and $\mathbf{w} = \left(\begin{smallmatrix}-1\\1\\2\end{smallmatrix}\right)$. Then

$$\operatorname{Span} S = \left\{ a\begin{pmatrix}1\\2\\-1\end{pmatrix} + b\begin{pmatrix}-1\\1\\2\end{pmatrix} : a, b \in \mathbb{R} \right\}$$

This is the plane through the origin 'spanned by' $\mathbf{v}$ and $\mathbf{w}$: hence the use of the word *span*!



4. (a) Let $S = \{\mathbf{i}, \mathbf{k}\} \subseteq \mathbb{R}^3$. The span of S is the xz-plane

$$\operatorname{Span} S = \{a\mathbf{i} + b\mathbf{k} : a, b \in \mathbb{R}\}$$

(b) If $T = S \cup \{3\mathbf{i} - 2\mathbf{k}\} = \{\mathbf{i}, \mathbf{k}, 3\mathbf{i} - 2\mathbf{k}\} \subseteq \mathbb{R}^3$, then $\operatorname{Span} T$ remains the xz-plane. The third vector $3\mathbf{i} - 2\mathbf{k}$ is redundant since it is a linear combination of the first two. Indeed

$$a\mathbf{i} + b\mathbf{k} + c(3\mathbf{i} - 2\mathbf{k}) = (a + 3c)\mathbf{i} + (b - 2c)\mathbf{k} \in \operatorname{Span}\{\mathbf{i}, \mathbf{k}\}$$

Part of our concern in this chapter is to more carefully consider such redundancies.

The examples should suggest the following.

Lemma 1.19. *If S is a subset of a vector space V, then* $\operatorname{Span} S$ *is a subspace of V.*

Proof. This is trivial if $S = \emptyset$. Otherwise, we follow the criteria in Theorem 1.10. Let $\mathbf{x}, \mathbf{y} \in \operatorname{Span} S$ and $\lambda \in \mathbb{F}$. Then $\exists m, n \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_m, \mathbf{y}_1, \ldots, \mathbf{y}_n \in S$, and $a_1, \ldots, a_m, b_1, \ldots, b_n \in \mathbb{F}$ such that

$$\mathbf{x} = a_1\mathbf{x}_1 + \cdots + a_m\mathbf{x}_m, \qquad \mathbf{y} = b_1\mathbf{y}_1 + \cdots + b_n\mathbf{y}_n$$

But then

$$\lambda\mathbf{x} + \mathbf{y} = \lambda a_1\mathbf{x}_1 + \cdots + \lambda a_m\mathbf{x}_m + b_1\mathbf{y}_1 + \cdots + b_n\mathbf{y}_n \in \operatorname{Span} S$$

■

Generating or Spanning Sets Part of our goal is to identify subsets S, particularly *small* subsets, of a vector space such that $\operatorname{Span} S$ is the entire space.

Definition 1.20. Let S be a subset of a vector space V. If $\operatorname{Span} S = V$, we say that S is a *spanning* or *generating set* for V. Alternatively, we say that S *spans* V or S *generates* V.

Examples 1.21.

1. $S = \{\mathbf{i}, \mathbf{j}\}$ generates $\mathbb{R}^2$, since every vector $\mathbf{v} \in \mathbb{R}$ is a linear combination $\mathbf{v} = x\mathbf{i} + y\mathbf{j}$ of vectors in S. Indeed $\mathbb{R}^2$ is essentially *defined* as $\operatorname{Span} S$!
2. Many vector spaces are defined via a spanning set: e.g. $P_3(\mathbb{R}) := \operatorname{Span}\{1, x, x^2, x^3\}$.
3. Consider $S = \{1 - 2x^2,\ 1 + x - x^2,\ 1 + 2x + x^3\} \subseteq P_3(\mathbb{R})$.
   (a) The polynomial x^3 lies in $\operatorname{Span} S$.
   For this, we need to find coefficients a, b, c such that
   $$a(1 - 2x^2) + b(1 + x - x^2) + c(1 + 2x + x^3) = x^3$$
   By equating the coefficients of $1, x, x^2$ and x^3 is it enough for us to solve a linear system
   $$\begin{cases} a + b + c = 0 \\ b + 2c = 0 \\ -2a - b = 0 \\ c = 1 \end{cases} \iff (a, b, c) = (1, -2, 1)$$
   (b) $1 + 3x + x^3 \notin \operatorname{Span} S$. It follows that S *does not generate* $P_3(\mathbb{R})$.
   This time, we need to show that there are no coefficients a, b, c such that
   $$a(1 - 2x^2) + b(1 + x - x^2) + c(1 + 2x + x^3) = 1 + 3x + x^3 \iff \begin{cases} a + b + c = 1 \\ b + 2c = 3 \\ -2a - b = 0 \\ c = 1 \end{cases}$$
   Substituting $c = 1$ in the second equation yields $b = 1$, however the remaining equations are now $a + 2 = 1$ and $-2a - 1 = 0$ which are inconsistent.

4. $S = \{1+x^2,\ 2-x^2,\ x,\ 1+4x\}$ generates the vector space $P_2(\mathbb{R})$.

   Given any $a+bx+cx^2 \in P_2(\mathbb{R})$ we need to see that there exist $g,h,j,k \in \mathbb{R}$ such that

$$\begin{aligned} a+bx+cx^2 &= g(1+x^2)+h(2-x^2)+jx+k(1+4x) \\ &= g+2h+k+(j+4k)x+(g-h)x^2 \end{aligned}$$

   By equating coefficients, this amounts to finding a solution (g,h,j,k) (as functions of a,b,c) to the underdetermined linear system

$$\begin{cases} g+2h+k=a \\ j+4k=b \\ g-h=c \end{cases}$$

   Only one solution is required, and $k=0$, $j=b$, $g=\frac{1}{3}(a+2c)$, $h=\frac{1}{3}(a-c)$ does the trick.

   Alternatively, you could try to explicitly construct the elements $1,x,x^2$ from those of S: in this situation it is fairly easy to do by inspection, e.g.,

$$1=(1+4x)-4x, \qquad x=x, \qquad x^2=\frac{2}{3}(1+x^2)-\frac{1}{3}(2-x^2)$$

   It follows that $\{1,x,x^2\} \subseteq \operatorname{Span} S$ and so $P_2(\mathbb{R}) = \operatorname{Span}\{1,x,x^2\} \subseteq \operatorname{Span} S$. Since, plainly, $\operatorname{Span} S \subseteq P_2(\mathbb{R})$ we have equality: $\operatorname{Span} S = P_2(\mathbb{R})$.

Aside: row operations review It should be revision, but the solution to the above linear system would likely have been found very slowly in a previous class. Here are some of the details. The required system can be put in augmented matrix form:

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 4 \\ 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} g \\ h \\ j \\ k \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \leftrightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 0 & 1 & a \\ 0 & 0 & 1 & 4 & b \\ 1 & -1 & 0 & 0 & c \end{array}\right)$$

Applying row operations, we can put this in (reduced) row echelon form:

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{3} & \frac{1}{3}(a+2c) \\ 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3}(a-c) \\ 0 & 0 & 1 & 4 & b \end{array}\right)$$

There is a free variable (k) here, but all solutions can easily be read off:

$$g=\frac{1}{3}(a+2c)-\frac{1}{3}k, \qquad h=\frac{1}{3}(a-c)-\frac{1}{3}k, \qquad j=b-4k$$

Choosing $k=0$ gives the solution referenced above.

Linear systems can always be tackled using augmented matrices, but it is encouraged to avoid them if you can: see, e.g., the alternative method for the last example.

Linear Dependence & Independence: when is a spanning set larger than necessary?

If $\mathbf{w} = 2\mathbf{v}$, then $\operatorname{Span}\{\mathbf{v}, \mathbf{w}\} = \operatorname{Span} \mathbf{v}$; for the purpose of spanning a subspace, the vector $\mathbf{w}$ is therefore redundant. To generalize this idea, we essentially have to extend the notion of *parallel.*

Definition 1.22. A finite non-empty subset $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of a vector space is *linearly dependent* if

$$\exists a_i \in \mathbb{F} \;\; \textit{not all zero}, \text{ for which } \; a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0}$$

Such an equation is a *linear dependence.*[a]

An infinite set is linearly dependent if it has as least one non-empty linearly dependent subset.

[a] The *not all zero* condition is crucial! You can always write $\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n$ (a *trivial representation* of $\mathbf{0}$), but this tells you nothing about the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. A linear dependence is a *non-trivial* representation of the zero vector!

Examples 1.23. 1. Vectors $\mathbf{v}, \mathbf{w}$ are linearly dependent (i.e. $\{\mathbf{v}, \mathbf{w}\}$ *is* linearly dependent) if and only if they are *parallel.*

2. $\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$ and $\mathbf{v}_3 = \begin{pmatrix} 7 \\ 5 \\ 6 \end{pmatrix}$ are linearly dependent since

$$2\mathbf{v}_1 + 3\mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$$

3. The infinite set $S = \{\begin{pmatrix} x \\ y \end{pmatrix} : y > 1\}$ is linearly dependent in $\mathbb{R}^2$. For instance $\{\begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix}\}$ is a finite linearly dependent subset of S, since $3\begin{pmatrix} 0 \\ 2 \end{pmatrix} - 2\begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

We now state the negation of the definition.

Definition (1.22 cont.). A finite subset $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of a vector space is *linearly independent* if

$$\forall a_i \in \mathbb{F}, \quad a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0} \implies a_1 = \cdots = a_n = 0$$

An infinite set is linearly independent if *all* of its finite subsets are linearly independent.

Examples 1.24. 1. The set $S = \{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -5 \end{pmatrix}\}$ is linearly independent in $\mathbb{R}^2$ since

$$a\begin{pmatrix} 2 \\ 1 \end{pmatrix} + b\begin{pmatrix} 3 \\ -5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{cases} 2a + 3b = 0 \\ a - 5b = 0 \end{cases} \implies a = b = 0$$

2. The empty set $\emptyset$ is trivially linearly independent since there is no condition to check.

3. Consider the set $S = \{1 - x^2, -x + 2x^2, 1 + 2x - x^2\}$ in $P_2(\mathbb{R})$. Attempting to find a linear dependence is equivalent to finding a non-trivial solution (a, b, c) to a system of linear equations

$$a(1 - x^2) + b(-x + 2x^2) + c(1 + 2x - x^2) = 0 \iff \begin{cases} a + c = 0 \\ -b + 2c = 0 \\ -a + 2b - c = 0 \end{cases}$$

Since the only solution is trivial $(a, b, c) = (0, 0, 0)$, the set S is linearly independent.

4. $S = \{1, x, x^2, x^3, \ldots\}$ is a linearly independent subset of $P(\mathbb{R})$: we leave this as an exercise.

We consider shrinking or enlarging certain sets of vectors. Prove the next result yourself.

Lemma 1.25. *Suppose $S_1 \subseteq S_2$ are subsets of a vector space. If S_1 is linearly dependent,[a] so is S_2.*

[a] Equivalently (the contrapositive): if S_2 is linearly independent, so is S_1.

Now turn the lemma on its head: if S_2 is linearly independent, when it is possible to find a *larger* linearly independent set $S_2 \supseteq S_1$? What follows is one of the most important results in the course.

Theorem 1.26. *Suppose that S is a linearly independent subset of V and that $\mathbf{v} \notin S$ is given. Then*

$$S \cup \{\mathbf{v}\} \text{ is linearly independent} \iff \mathbf{v} \notin \operatorname{Span} S$$

Be careful reading the proof: we use the contrapositive and prove both directions simultaneously!

Proof. By definition, $S \cup \{\mathbf{v}\}$ is linearly *dependent* if and only if there exist *finitely many* vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in S$ and scalars $a, a_1, \ldots, a_n$ (not all zero), such that

$$a\mathbf{v} + a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{0} \tag{$*$}$$

Plainly $a \neq 0$, for otherwise S would be linearly dependent. By dividing through all coefficients by $-a$ we therefore see that $(*)$ is equivalent to

$$\mathbf{v} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n \iff \mathbf{v} \in \operatorname{Span} S$$

■

Examples 1.27. 1. Let $S = \{\mathbf{i}\} = \{\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\}$ and $\mathbf{v} = \left(\begin{smallmatrix}a\\b\end{smallmatrix}\right)$. Then

$$\{\mathbf{i}, \mathbf{v}\} \text{ linearly independent} \iff \mathbf{v} \notin \operatorname{Span}\{\mathbf{i}\} \iff \mathbf{v} \text{ not parallel to } \mathbf{i} \iff b \neq 0$$

2. Plainly $S = \{\mathbf{v}, \mathbf{w}\} = \left\{\left(\begin{smallmatrix}1\\2\\-1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1\\1\\2\end{smallmatrix}\right)\right\}$ is linearly independent (recall Example 1.18.3).

   (a) Let $\mathbf{u} = \left(\begin{smallmatrix}1\\-3\\0\end{smallmatrix}\right)$: we check that $\mathbf{u} \notin \operatorname{Span} S$. If it were, there would exist $a, b \in \mathbb{R}$ such that

   $$\mathbf{u} = a\begin{pmatrix}1\\2\\-1\end{pmatrix} + b\begin{pmatrix}-1\\1\\2\end{pmatrix} \implies \begin{pmatrix}1\\-3\\0\end{pmatrix} = \begin{pmatrix}1 & -1\\2 & 1\\-1 & 2\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}$$

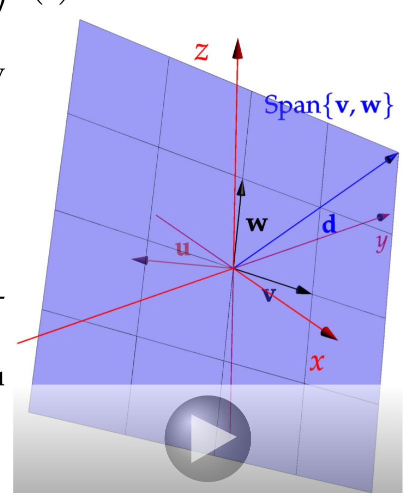
   which has no solutions. It follows that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent. Indeed $\operatorname{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \mathbb{R}^3$.

   (b) If we let $\mathbf{d} = \left(\begin{smallmatrix}0\\6\\2\end{smallmatrix}\right)$, then

   $$\mathbf{d} = 2\mathbf{v} + 2\mathbf{w}$$

   whence $\mathbf{d} \in \operatorname{Span}\{\mathbf{v}, \mathbf{w}\}$ and so $\{\mathbf{d}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

   In the picture, $\mathbf{d}$ lies in the plane spanned by $\mathbf{v}, \mathbf{w}$ while $\mathbf{u}$ does not.



The Theorem should be intuitive in $\mathbb{R}^2$ and $\mathbb{R}^3$ where planes and lines are easy to visualize. Analogues abound elsewhere: indeed the following are reasonable statements according to the RGB (additive) theory of colors,

$$\text{purple} \in \operatorname{Span}\{\text{red}, \text{blue}\}, \qquad \text{brown} \notin \operatorname{Span}\{\text{red}, \text{blue}\}$$

For instance purple, red and blue are not independent colors.

Exercises 1.3

1. In $\mathbb{R}^3$, let $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} 1 \\ 7 \\ 5 \end{pmatrix}$. Show that $\mathbf{x} \in \operatorname{Span}\{\mathbf{v}, \mathbf{w}\}$.

2. For the following lists of polynomials in $P_3(\mathbb{R})$, determine whether f can be expressed as a linear combination of g and h.
   (a) $f = 4x^3 + 2x^2 - 6, \quad g = x^3 - 2x^2 + 4x + 1, \quad h = 3x^3 - 6x^2 + x + 4$
   (b) $f = x^3 - 8x^2 + 4x, \quad g = x^3 - 2x^2 + 3x - 1, \quad h = x^3 - 2x + 3$

3. Determine whether the vectors $A, B \in M_2(\mathbb{R})$ lie in the span of S:
$$A = \begin{pmatrix} 1 & 2 \\ -3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad S = \left\{ \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \right\}$$

4. Determine whether the set $S = \{1 + x + x^2,\ x - x^2,\ 2 + 3x^2\}$ generates $P_2(\mathbb{R})$.

5. Which of the following sets are linearly independent? Prove your assertions.
   (a) $\{2, 3 - x, 1 - 2x^2\}$ in $P_2(\mathbb{R})$
   (b) $\{1, x^2 - x, x^2 + x, x^2\}$ in $P_2(\mathbb{R})$
   (c) $\{\sin(x), \cos(x), \tan(x)\}$ in $C\big((-\frac{\pi}{2}, \frac{\pi}{2}), \mathbb{R}\big)$ (recall Example 1.11.3 for the notation)
   (d) $\{\cos^2(x), \sin^2(x), \cos(2x)\}$ in $C(\mathbb{R}, \mathbb{R})$

6. Suppose that $S = \{\mathbf{v}\}$ is a linearly *dependent* set. What is $\mathbf{v}$?

7. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be linearly independent vectors in V. Prove that $\operatorname{Span}\{\mathbf{v}_2, \ldots, \mathbf{v}_n\} \neq V$.

8. Explicitly verify the claim on page 16 that $\operatorname{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} = \mathbb{R}^3$.

9. Show that the functions f, g defined by $f(x) = 2x$ and $g(x) = |x|$ are linearly independent in the vector space $C([-1, 1], \mathbb{R})$, but linearly dependent in $C([0, 1], \mathbb{R})$.

10. Suppose that c is a constant, and consider the continuous functions $f, g \in C(\mathbb{R}, \mathbb{R})$ defined by
$$f(x) = \cos(x + c), \qquad g(x) = 2\sin x$$
For what values of c are the functions linearly independent? Draw a picture of what happens.

11. Prove that $\operatorname{Span} S$ is the intersection of all subspaces of V which contain S.

12. Justify Example 1.24.4: the infinite set $\{1, x, x^2, \ldots\}$ is linearly independent.

13. Let X, Y, Z be subsets of a vector space V. Prove that:
    (a) $\operatorname{Span}(X \cup Y) = \operatorname{Span} X + \operatorname{Span} Y$;
    (b) $\operatorname{Span}(X \cup Y) = \operatorname{Span} X \iff Y \subseteq \operatorname{Span} X$
    (c) If $\mathbf{y} \in Z$, then $\operatorname{Span}(Z \setminus \{\mathbf{y}\}) = \operatorname{Span} Z \iff \mathbf{y} \in \operatorname{Span}(Z \setminus \{\mathbf{y}\})$

1.4 Bases and Dimension

We now come, arguably, to the most important definition of the course.

Definition 1.28. A *basis* of a vector space is a linearly independent spanning set.

Our main goals are to see that every vector space has a basis and that all bases of the same space have the same number of elements, what we'll call the dimension.

Standard Bases Many vector spaces have commonly used *standard* bases.

Vector Space V	Standard Basis β
$\mathbb{R}^2$	$\{\mathbf{i},\mathbf{j}\}$
$\mathbb{R}^3$	$\{\mathbf{i},\mathbf{j},\mathbf{k}\}$
$\mathbb{F}^n$	$\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ where $\mathbf{e}_i$ has i^{th} entry 1 and the rest 0
$M_{m\times n}(\mathbb{F})$	$\{E^{ij} : 1\le i\le m, 1\le j\le n\}$ where E^{ij} has ij^{th} entry 1 and remaining entries 0
$P_n(\mathbb{F})$	$\{1,x,x^2,\ldots,x^n\}$
$P(\mathbb{F})$	$\{1,x,x^2,x^3,\ldots\}$

Examples 1.29. 1. The standard bases of $P_3(\mathbb{R})$ and $M_2(\mathbb{R})$ are, respectively,

$$\{1,x,x^2,x^3\} \qquad \text{and} \qquad \{E^{11},E^{12},E^{21},E^{22}\}=\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$$

2. The above spaces have other bases! For example $\beta=\{x,x-1,1+x^2\}$ is a basis of $P_2(\mathbb{R})$: verification of the following should be straightforward:

   - β is linearly independent: $ax+b(x-1)+c(1+x^2)=0 \implies a=b=c=0$
   - β spans $P_2(\mathbb{R})$: $\forall s,t,u,\ \exists a,b,c$ such that $s+tx+ux^2=ax+b(x-1)+c(1+x^2)$

   What matters is that you are comfortable transforming the definitions into algebra!

3. Following the convention that $\operatorname{Span}\emptyset=\{\mathbf{0}\}$, the empty set is a basis of the trivial space $\{\mathbf{0}\}$.

The Unique Co-ordinate Representation We first discuss one of the primary uses of a basis: the representation of vectors in terms of *co-ordinates*.

Definition 1.30. Let $\beta=\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ be a basis of V over $\mathbb{F}$ and suppose $\mathbf{v}\in V$ is given. The *co-ordinate representation of* $\mathbf{v}$ *with respect to* β is the column vector

$$[\mathbf{v}]_\beta := \begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}\in\mathbb{F}^n \quad\text{where}\quad \mathbf{v}=a_1\mathbf{v}_1+\cdots+a_n\mathbf{v}_n$$

To check that this is well-defined, we need to make sure that each vector $\mathbf{v}$ has exactly one co-ordinate representation. We'll deal with this in Theorem 1.32, after seeing an example.

Example 1.31. In $P_2(\mathbb{R})$, consider the bases $\alpha = \{1, x, x^2\}$ and $\beta = \{x, x-1, 1+x^2\}$, and the vector

$$p(x) = 3x - 2(x-1) + 5(1+x^2) = 7 + x + 5x^2$$

The co-ordinate representations with respect to the two bases are then:

$$[p]_\alpha = \begin{pmatrix} 7 \\ 1 \\ 5 \end{pmatrix}, \qquad [p]_\beta = \begin{pmatrix} 3 \\ -2 \\ 5 \end{pmatrix}$$

The advantage of co-ordinates is that we can easily invoke matrix methods. The challenge is to keep in mind the *basis* used in the conversion, so we can properly convert back once we're done!

We now verify the uniqueness of co-ordinate representations. Amazingly, this property essentially characterises the concept of a basis.

Theorem 1.32. *Let $\beta = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a non-empty finite subset of a vector space V. Then β is a basis if and only if each $\mathbf{v} \in V$ can be written as a unique linear combination*

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \tag{$*$}$$

Compare this to Theorem 1.16: we are really saying that $V = \operatorname{Span}\mathbf{v}_1 \oplus \cdots \oplus \operatorname{Span}\mathbf{v}_n$.

Proof. ($\Rightarrow$) If β is a basis, then $V = \operatorname{Span}\beta$ and so every vector can be expressed in the form $(*)$. Now suppose $\exists \mathbf{v} \in V$ with at least two *distinct* representations:

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$$

It follows that

$$(a_1 - b_1)\mathbf{v}_1 + \cdots + (a_n - b_n)\mathbf{v}_n = \mathbf{0}$$

is a linear dependence on β. Contradiction.
($\Leftarrow$) Conversely, suppose β is not a basis. There are two possibilities:

(a) β does not generate V. In this case, $\exists \mathbf{v} \notin \operatorname{Span}\beta$ with no representation.

(b) β generates V but is linearly dependent. In this case there exists a linear dependence

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

from which the zero vector[a] has at least two representations!

Either way, there exists some $\mathbf{v} \in V$ without a unique representation. ∎

[a] $\mathbf{0} = 0\mathbf{v}_1 + \cdots 0\mathbf{v}_n$. Indeed any $\mathbf{v} \in V$ will have multiple representations in this case.

We don't typically refer to *co-ordinates* with respect to infinite bases, but the Theorem can be rephrased so that the uniqueness of representation holds. We will return to co-ordinate representations and their relationship to linear maps and matrix multiplication in the next chapter.

Existence of Finite Bases

If we attempt to enlarge a basis β of V by adding a new vector $\mathbf{v} \notin \beta$, a quick appeal to Theorem 1.26 (let $S = \beta$) shows that

$$\mathbf{v} \in \operatorname{Span}\beta\,(= V) \implies \beta \cup \{\mathbf{v}\} \text{ is linearly dependent}$$

Otherwise said, a basis is a *maximal linearly independent set*. This suggests a simple looped algorithm:

1. Start with a linearly independent set X (even $X = \emptyset$ will do!).
2. Does there exist a vector $\mathbf{s}$ such that $X \cup \{\mathbf{s}\}$ is linearly independent?

   Yes: Repeat step 2 with X replaced with $X \cup \{\mathbf{s}\}$.

   No: Stop. We have a basis.

The algorithm has two problems: how do we find a suitable $\mathbf{s}$, and how do we know that the algorithm will terminate? Both these problems can be addressed by restricting to vector spaces spanned by a *finite* set.

Definition 1.33. A vector space V is *finite-dimensional* if it has a finite spanning set: if there exists a *finite* subset $S \subseteq V$ such that $\operatorname{Span} S = V$.

Theorem 1.34 (Existence of a Basis). *Every finite-dimensional vector space has a basis.*
More specifically, suppose X and S are subsets of V such that:

- *S is a finite spanning set for V;*
- *X is a linearly independent subset of S.*

Then there exists a basis β of V such that $X \subseteq \beta \subseteq S$: in particular β is a finite *set.*

Proof. Suppose V is non-trivial, for otherwise $\emptyset$ is a basis ($X = \emptyset$ and $S = \emptyset$ or $\{\mathbf{0}\}$).
Let $m = |X|$ and $n = |S|$ be the cardinalities so that $m \le n$, and label $X = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$. Then

$$X \subseteq S \implies \operatorname{Span} X \subseteq \operatorname{Span} S = V$$

Loop: If $\operatorname{Span} X = \operatorname{Span} S = V$, we are done: X is a basis.
Otherwise, $\exists \mathbf{s}_{m+1} \in S$ such that $\mathbf{s}_{m+1} \notin \operatorname{Span} X$, whence (Theorem 1.26)

$$X \cup \{\mathbf{s}_{m+1}\} = \{\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{s}_{m+1}\} \text{ is linearly independent.}$$

Now repeat the loop with $X \cup \{\mathbf{s}_{m+1}\}$ in place of X (induction).

The process must terminate with a basis in at most $n - m$ steps since S is a *finite* spanning set.
To establish the primary claim, simply choose any $\mathbf{x} \in S$ and let $X = \{\mathbf{x}\}$. ∎

The finite spanning set S was crucial in resolving the problems with our looped algorithm: it provided a finite list of vectors from which to choose, and guaranteed that only finitely many loops were possible thus forcing the algorithm to terminate.[1] The existence of bases for *infinite-dimensional* spaces (no finite spanning sets) is more technical and will be outlined in the next section.

Example 1.35. We follow the algorithm in $\mathbb{R}^3$, omitting explicit calculations for brevity.

$$S = \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4, \mathbf{s}_5, \mathbf{s}_6\} = \left\{ \begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}2\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\1\end{pmatrix}, \begin{pmatrix}5\\4\\2\end{pmatrix}, \begin{pmatrix}0\\1\\3\end{pmatrix}, \begin{pmatrix}1\\1\\4\end{pmatrix} \right\}$$

1. Let $X = \{\mathbf{s}_1\}$. Since $\mathbf{s}_2 \notin \operatorname{Span} X$ we conclude that $\{\mathbf{s}_1, \mathbf{s}_2\}$ is linearly independent.
2. $\{\mathbf{s}_1, \mathbf{s}_2\}$ does not span $\mathbb{R}^3$ so we need another vector.
   - $\mathbf{s}_3 = \mathbf{s}_2 - \mathbf{s}_1 \in \operatorname{Span}\{\mathbf{s}_1, \mathbf{s}_2\}$ so we reject $\mathbf{s}_3$
   - $\mathbf{s}_4 = \mathbf{s}_1 + 2\mathbf{s}_2 \in \operatorname{Span}\{\mathbf{s}_1, \mathbf{s}_2\}$ so we also reject $\mathbf{s}_4$
   - We accept $\mathbf{s}_5$ since $\mathbf{s}_5 \notin \operatorname{Span}\{\mathbf{s}_1, \mathbf{s}_2\}$.
3. $\beta := \{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_5\}$ is linearly independent and spans $\mathbb{R}^3$: it is a suitable basis.

The Exchange Theorem and its Consequences

Our next goal is *comparison* of the cardinalities of spanning sets and linearly independent set (and thus bases), the key to which is the *Exchange* (or *Replacement*) *Theorem.* Take your time, since this is the trickiest result of the course so far.

Theorem 1.36 (Exchange Theorem). *Let V be a finite-dimensional vector space. If S is a finite spanning set and X is a linearly independent subset of V, then $|X| \le |S|$. More specifically,*

$$\exists T \subseteq S \text{ such that } |T| = |X| \text{ and } \operatorname{Span}(X \cup (S \setminus T)) = V$$

A few observations before we see the proof.

- The hypotheses are the same as for the Existence Theorem (1.34), except that X need not be a subset of S. The result therefore allows us to compare *unrelated* subsets.
- The result shows that *every* linearly independent set is no larger than *every* finite spanning set. In particular, we obtain the important fact that *every linearly independent subset and thus basis is finite*!
- The subset T is sometimes called the *exchange*, since the theorem essentially exchanges T with X while preserving the span.
- Since the proof depends on a tricky induction, it is strongly recommended to work through an example (say Example 1.37) while reading. The exchange is often easy to compute when X and S are small. If you really want to understand the proof, make up more examples! Alternatively, simply skip the proof and come back later; while important, it is technical and hard to use directly in examples.

[1] Imagine applying the algorithm $X = \{1\}$ and $S = \{1, x, x^2, \ldots\}$ in the space of polynomials $P(\mathbb{R})$: what happens?

Proof. Denote $n = |S|$ and $m = \min\{n, |X|\}$: our eventual goal is to see that $|X| = m$, but at present we don't know whether X is finite!

Consider a subset $\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \subseteq X$. We prove the following claim by induction:

$$\forall k \in \{0, 1, \ldots, m\},\ \exists \mathbf{s}_1, \ldots, \mathbf{s}_k \in S \text{ such that } \operatorname{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{s}_{k+1}, \ldots, \mathbf{s}_n\} = V \tag{†}$$

Base case If $k = 0$ then the claim is trivial for S spans V.

Induction step Suppose the claim holds for some $k < m$. Since $\{\mathbf{x}_1, \ldots, \mathbf{x}_k, \mathbf{s}_{k+1}, \ldots, \mathbf{s}_n\}$ spans V, there exist coefficients a_i, b_j such that

$$\mathbf{x}_{k+1} = a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k + b_{k+1}\mathbf{s}_{k+1} + \cdots + b_n\mathbf{s}_n \tag{$*$}$$

Since $(*)$ is a linear dependence, the *independence* of X shows that at least one $b_j \neq 0$: WLOG assume $b_{k+1} \neq 0$. Since

$$\mathbf{s}_{k+1} = -b_{k+1}^{-1}(a_1\mathbf{x}_1 + \cdots + a_k\mathbf{x}_k - \mathbf{x}_{k+1} + b_{k+2}\mathbf{s}_{k+2} + \cdots + b_n\mathbf{s}_n)$$

we may eliminate $\mathbf{s}_{k+1}$ from any linear combination describing an element of V at the cost of including $\mathbf{x}_{k+1}$: we conclude that

$$V = \operatorname{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}, \mathbf{s}_{k+2}, \ldots, \mathbf{s}_n\}$$

By induction, the claim is proved. Taking $k = m$ and setting $T = \{\mathbf{s}_1, \ldots, \mathbf{s}_m\}$ we see that

$$V = \operatorname{Span}\Big(\{\mathbf{x}_1, \ldots, \mathbf{x}_m\} \cup (S \setminus T)\Big)$$

Now suppose, for contradiction, that $|X| > m$. Then $n = m$ and $\exists \mathbf{x}_{m+1} \in X$. Since (†) holds for $k = m = n$, we see that

$$\mathbf{x}_{m+1} \in V = \operatorname{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$$

which contradicts the linear independence of X.

We complete the proof by observing that $|X| = m \leq n = |S|$. ∎

Example 1.37. Let $V = \mathbb{R}^3$, $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$, $X = \left\{\left(\begin{smallmatrix}2\\3\\5\end{smallmatrix}\right), \left(\begin{smallmatrix}6\\9\\12\end{smallmatrix}\right)\right\}$ and apply the induction step twice:

1. Since $\mathbf{x}_1 = 2\mathbf{i} + 3\mathbf{j} + 5\mathbf{k}$ and the coefficient of $\mathbf{i}$ is non-zero, we choose $\mathbf{s}_1 = \mathbf{i}$. Observe that $\operatorname{Span}\{\mathbf{x}_1, \mathbf{j}, \mathbf{k}\} = \operatorname{Span} S = \mathbb{R}^3$.

2. Now find the coefficients of $\mathbf{x}_2$ with respect to $\{\mathbf{x}_1, \mathbf{j}, \mathbf{k}\}$: this might need a little augmented matrix work, but we see that

   $$\mathbf{x}_2 = \begin{pmatrix}6\\9\\12\end{pmatrix} = 3\mathbf{x}_1 + 0\mathbf{j} - 3\mathbf{k}$$

   so we choose $\mathbf{s}_2 = \mathbf{k}$. Again we have $\operatorname{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{j}\} = \mathbb{R}^3$. The exchange is therefore $T = \{\mathbf{i}, \mathbf{k}\}$.

Armed with the Exchange Theorem, the key facts come quickly and easily:

Corollary 1.38. *Given a finite-dimensional vector space:*

1. *(Extension Theorem) Any linearly independent subset may be extended to a basis.*
2. *(Well-definition of Dimension) Any two bases have the same cardinality.*

Proof. If the space is trivial then both statements are immediate. Otherwise:

1. Suppose S is a finite spanning set for V and that X a linearly independent subset. The Exchange Theorem says there exists a finite spanning set $X \cup (S \setminus T)$ containing X. By the Existence theorem, there exists a basis β such that $X \subseteq \beta \subseteq X \cup (S \setminus T)$.
2. If β and γ are bases, take $X = \beta$ and $S = \gamma$ in the Exchange Theorem to see that $|\beta| \leq |\gamma|$. Now repeat the argument with the roles reversed.

■

Definition 1.39. The *dimension* $\dim_{\mathbb{F}} V$ of a finite-dimensional vector space V over a field $\mathbb{F}$ is the cardinality of any basis.[a]

[a] We usually write $\dim V$ if the field is understood, but be careful: see Exercise 1.4.6...

Examples 1.40. 1. The dimension is often part of the name of a vector space or is easily read off:

$$\dim \mathbb{R}^5 = 5, \qquad \dim \mathbb{F}^n = n, \qquad \dim M_{m\times n}(\mathbb{F}) = mn$$

2. Beware of *polynomials*! The standard basis of $P_n(\mathbb{R})$ is $\{1, x, \ldots, x^n\}$, whence $\dim P_n(\mathbb{R}) = n+1$.

Corollary 1.41. *Suppose W is a subspace of a finite-dimensional space V. Then:*

1. $\dim W \leq \dim V$.
2. $\dim W = \dim V \implies W = V$.

Proof. By the Extension Theorem, we may extend any basis α of W to a basis β of V. Since $\alpha \subseteq \beta$, we plainly have $\dim W = |\alpha| \leq |\beta| = \dim V$. Part 2 is an exercise.

■

Summary: bases of finite-dimensional vector spaces We have not quite proved all of the following, but all should now seem at least reasonable.

1. Every such space has a basis and all have the same cardinality (the dimension).
2. We can extend a linearly independent set to a basis: a basis is a *maximal linearly independent set.*
3. Every spanning set contains a basis as a subset: a basis is a *minimal spanning set.*
4. A subset β is a basis of V if it satisfies *any two* of the following (it then satisfies the third):

$$|\beta| = \dim V, \qquad \beta \text{ is linearly independent}, \qquad \operatorname{Span}\beta = V$$

Example 1.42. We verify that $\beta = \left\{ \left(\begin{smallmatrix}1\\3\\4\end{smallmatrix}\right), \left(\begin{smallmatrix}2\\1\\5\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) \right\}$ is a basis of $\mathbb{R}^3$.

Since $|\beta| = 3 = \dim \mathbb{R}^3$, we need only check linear independence:

$$a\begin{pmatrix}1\\3\\4\end{pmatrix} + b\begin{pmatrix}2\\1\\5\end{pmatrix} + c\begin{pmatrix}1\\1\\1\end{pmatrix} = \mathbf{0} \iff \begin{pmatrix}1&2&1\\3&1&1\\4&5&1\end{pmatrix}\begin{pmatrix}a\\b\\c\end{pmatrix} = \mathbf{0} \implies a = b = c = 0$$

since the matrix is invertible.

It is unnecessary, but the invertibility also shows directly that β is a spanning set: given $\mathbf{x} \in \mathbb{R}^3$,

$$\begin{pmatrix}a\\b\\c\end{pmatrix} := \begin{pmatrix}1&2&1\\3&1&1\\4&5&1\end{pmatrix}^{-1}\mathbf{x} \implies \mathbf{x} = a\begin{pmatrix}1\\3\\4\end{pmatrix} + b\begin{pmatrix}2\\1\\5\end{pmatrix} + c\begin{pmatrix}1\\1\\1\end{pmatrix} \in \operatorname{Span}\beta$$

Exercises 1.4 1. Prove carefully that $\beta = \{3\mathbf{i} + 2\mathbf{k}, 2\mathbf{i} + \mathbf{k}, \mathbf{j} + \mathbf{k}\}$ is a basis of $\mathbb{R}^3$.

2. Let $p(x) = 3 - 5x + 7x^2 \in P_2(\mathbb{R})$. With respect to the bases $\beta = \{1 - x, 1 + x^2, x - 2x^2\}$ and $\gamma = \{2 - x, x^2, 1 + x\}$, find the co-ordinate representations $[p]_\beta$ and $[p]_\gamma$.

3. As in Exercise 1.35, find a subset of S which is a basis of the vector space V.

   (a) $V = \mathbb{R}^3,\ S = \left\{ \left(\begin{smallmatrix}1\\1\\2\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\2\\3\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\2\\2\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\0\\1\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\2\\3\end{smallmatrix}\right), \left(\begin{smallmatrix}2\\1\\2\end{smallmatrix}\right) \right\}$

   (b) $V = P_3(\mathbb{R}),\ S = \{1 + 2x,\ 1 + x + x^2,\ 2 + x - x^2,\ 3 + 2x,\ x - 2x^3\}$

4. Find a basis and thus the dimension of the following subspace of $\mathbb{F}^5$:

$$W = \{a_1\mathbf{e}_1 + \cdots + a_5\mathbf{e}_5 \in \mathbb{F}^5 : a_1 - a_3 - a_4 = 0\}$$

5. Let $\mathbf{u}, \mathbf{v}, \mathbf{w}$ be distinct vectors in a vector space V. Prove that if $\beta = \{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a basis of V, then $\gamma := \{\mathbf{u} + \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w}, \mathbf{w}\}$ is also a basis of V.

6. $\mathbb{C}^3$ is a vector space over $\mathbb{C}$ *and* over $\mathbb{R}$. What are the values $\dim_{\mathbb{C}} \mathbb{C}^3$ and $\dim_{\mathbb{R}} \mathbb{C}^3$? State a basis in each case.

7. (a) Define $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. Prove that this is a vector space over $\mathbb{Q}$ and that $\beta = \{1, \sqrt{2}\}$ is a basis.

   (b) More generally, if $d \in \mathbb{Z}$ is not a perfect square, prove that $\dim_{\mathbb{Q}} \mathbb{Q}(\sqrt{d}) = 2$.

8. Explain the observation in the proof of the Existence Theorem 1.34: If $\operatorname{Span} X \neq \operatorname{Span} S$, then $\exists \mathbf{s}_{m+1} \in S$ such that $\mathbf{s}_{m+1} \notin \operatorname{Span} X$.

9. Given subsets X and S of the vector space V, compute the exchange T from the Exchange Theorem by mirroring Example 1.37.

   (a) $X = \left\{ \left(\begin{smallmatrix}1\\1\\2\end{smallmatrix}\right), \left(\begin{smallmatrix}2\\0\\-1\end{smallmatrix}\right) \right\},\ S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\},\ V = \mathbb{R}^3$.

   (b) $X = \{1 - x, 2 + x^2, 1 + x^3\},\ S = \{1, x, x^2, x^3\},\ V = P_3(\mathbb{R})$.

   (c) $X = \left\{ \left(\begin{smallmatrix}0&1\\1&1\end{smallmatrix}\right), \left(\begin{smallmatrix}0&-1\\0&1\end{smallmatrix}\right) \right\},\ S = \{E^{11}, E^{12}, E^{21}, E^{22}\},\ V = M_2(\mathbb{R})$.

10. Let $V = \{(x_n)_{n=1}^{\infty}\}$ be the set of all sequences of real numbers. This is a vector space over $\mathbb{R}$ under elementwise addition and scalar multiplication. For example: if $x_n = \frac{1}{n}$ and $y_n = 1 - \frac{1}{n^2}$, then $(x_n) + (y_n)$ is the sequence (z_n) with nth term

$$z_n = x_n + y_n = \frac{1}{n} + 1 - \frac{1}{n^2}$$

   (a) For each $m \in \mathbb{N}$ define the sequence $E^m = (e_n^m)$ where

$$e_n^m = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases}$$

   Thus $E^1 = (1, 0, 0, 0, \ldots)$ and $E^2 = (0, 1, 0, 0, \ldots)$, etc. Show that the set $X = \{E^m : m \in \mathbb{N}\}$ is a linearly independent subset of V.

   (b) Is X a basis of V? Why/why not?

11. Let V be a vector space with dimension $n \geq 1$, and let S be a generating set.

   (a) Show that S contains a linearly independent subset X.

   (b) If X is a linearly independent subset of S, but X is *not* a basis, prove that $\exists \mathbf{s} \in S$ such that $X \cup \{\mathbf{s}\}$ is linearly independent.

   (c) Prove that there exists a subset of S which is a basis of V.

   (d) Prove that $|S| \geq n$.

   (*This is asking you to modify the proof of the Existence Theorem. Note that you cannot assume that S is a finite set!*)

12. (Optional application) In this question we use linear algebra to find a polynomial of minimal degree through a set of points in the plane. Suppose that a_0, a_1 are distinct real numbers. Define the functions

$$f_0(x) = \frac{x - a_1}{a_0 - a_1}, \qquad f_1(x) = \frac{x - a_0}{a_1 - a_0}$$

   It follows that

$$f_i(a_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{$*$}$$

   (a) Prove that f_0 and f_1 are linearly independent.

   (b) Suppose that b_0, b_1 are real numbers. Show that the straight line passing through the points (a_0, b_0) and (a_1, b_1) lies in $\operatorname{Span}\{f_0, f_1\}$ and that, consequently, $\{f_0, f_1\}$ forms a basis of the vector space of linear polynomials $P_1(\mathbb{R})$.

   (c) Repeat parts (a) and (b) for any set of distinct values $a_0, a_1, \ldots, a_n$ to obtain polynomials $f_0, f_1, \ldots, f_n$ which satisfy $(*)$ and form a basis of $P_n(\mathbb{R})$. Hence or otherwise, prove that there is a unique degree $\leq n + 1$ polynomial passing through any points $(a_0, b_0), \ldots, (a_n, b_n)$ where the a_i are distinct.

   (d) Hence or otherwise, find the unique degree 3 polynomial which passes through the points $(0, 1), (1, 4), (3, -1)$ and $(5, 10)$.

1.5 Maximal linearly independent subsets (non-examinable)

In the previous section, we showed that every finite-dimensional vector space has a basis. What about other vector spaces? Does every vector space have a basis?

Examples 1.43. To see the difficulty, consider two related spaces and the set $\beta = \{1, x, x^2, x^3, \cdots\}$.

1. The space of polynomials $P(\mathbb{R})$ has standard basis β (Exercise 1.3.12), and is therefore an infinite-dimensional space with a *countable* basis: it seems reasonable to write $\dim P(\mathbb{R}) = \aleph_0$.

2. The space V of formal *power series* with coefficients in $\mathbb{R}$ contains the vector

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots$$

an *infinite* combination of the elements of β. Plainly β is not a basis of V. But does V have a basis and, if so, how can we find one?

There are two standard ways to tackle our problem.

1. Broaden the concept of linear combination/span to allow for *infinite sums.*[2] This introduces a new difficulty, *convergence,* which takes us into the realm of analysis and requires further definitions. If you later study Banach and Hilbert spaces, this is the approach you will follow. Indeed, in the context of power series, β is incredibly useful, *even more so than a basis* would be!

2. Appeal to *Zorn's Lemma*, a technical result equivalent to the (somewhat) controversial *axiom of choice.* This is the approach we'll follow for the remainder of the section.

Definition 1.44. Let $\mathcal{F}$ be a set of sets. A subset $\mathcal{C} \subseteq \mathcal{F}$ is a *chain*[a] in $\mathcal{F}$ if

$$\forall A, B \in \mathcal{C} \text{ either } A \subseteq B \text{ or } B \subseteq A$$

A chain $\mathcal{C}$ has an *upper bound* in $\mathcal{F}$ if there is some set $B \in \mathcal{F}$ such that

$$\forall A \in \mathcal{C} \text{ we have } A \subseteq B$$

A set $\beta \in \mathcal{F}$ is *maximal* it is a subset of no member of $\mathcal{F}$ but itself.

[a] Alternatively $\mathcal{C}$ is a *nest*, a *tower,* or is *totally ordered.*

The idea is to let $\mathcal{F}$ to be the set of all linearly independent subsets of a vector space V. Our goal is then to hunt for a maximal member of $\mathcal{F}$, since a basis β is precisely a *maximal linearly independent set* (see Exercise 1.5.1):

1. β is linearly independent.

2. The only linearly independent subset of V containing β is β itself.

[2] Definition 1.17 only allows us to conclude, by induction, that any *finite sum* of vectors $\sum_{i=1}^{n} \mathbf{v}_i$ is well-defined. In the abstract, i.e. without *limits,* an infinite sum $\sum_{n=1}^{\infty} \mathbf{v}_n$ has no meaning.

Examples 1.45. 1. Consider the standard basis $\beta = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ of $\mathbb{R}^3$. Clearly β is an upper bound for the following chain of linearly independent subsets

$$\mathcal{C} = \Big\{\{\mathbf{i}\}, \{\mathbf{i}, \mathbf{j}\}, \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}\Big\}$$

2. The basis $\beta = \{1, x, x^2, \ldots\}$ of $P(\mathbb{R})$ is an upper bound for the chain

$$\mathcal{C} = \Big\{\{1\}, \{1, x\}, \{1, x, x^2\}, \ldots\Big\}$$

Read this example carefully: the ellipsis hides *infinitely many* subsets. In particular the upper bound β does not have to be an element of the chain! It is, however, the *union* $\beta = \bigcup\limits_{U \in \mathcal{C}} U$ of all elements of the chain...

Axiom 1.46 (Zorn's Lemma). Let $\mathcal{F}$ be a non-empty family of sets. If every chain $\mathcal{C} \subseteq \mathcal{F}$ has an upper bound $M_{\mathcal{C}} \in \mathcal{F}$, then $\mathcal{F}$ has a maximal member.

Theorem 1.47. *Every vector space has a basis.*

Proof. If V is non-trivial, let $\mathcal{F} = \{\text{linearly independent subsets of } V\}$. Plainly this is non-empty. Suppose $\mathcal{C} \subseteq \mathcal{F}$ is a chain and define

$$M_{\mathcal{C}} := \bigcup_{U \in \mathcal{C}} U$$

We claim that $M_{\mathcal{C}}$ is an upper bound for $\mathcal{C}$ in $\mathcal{F}$. For this, we need to show two things:

1. $M_{\mathcal{C}} \in \mathcal{F}$: that is, $M_{\mathcal{C}}$ is a linearly independent set.
2. $\forall A \in \mathcal{C}$, we have $A \subseteq M_{\mathcal{C}}$.

The latter is obvious from the definition of union! For the former, suppose that $\mathbf{u}_1, \ldots, \mathbf{u}_n \in M_{\mathcal{C}}$ are distinct vectors such that

$$a_1\mathbf{u}_1 + \cdots + a_n\mathbf{u}_n = \mathbf{0}$$

By the total ordering of $\mathcal{C}$, we see[a] that $\exists U \in \mathcal{C}$ such that $\mathbf{u}_1, \ldots, \mathbf{u}_n \in U$. But each U is linearly independent, whence $a_1 = \cdots = a_n = 0$. It follows that $M_{\mathcal{C}} \in \mathcal{F}$.

Applying Zorn's lemma, we see that $\mathcal{F}$ has a maximal element β, which is necessarily a basis of V. ∎

[a] Since $\mathbf{u}_i \in M_{\mathcal{C}}$, $\exists U_i \in \mathcal{C}$ such that $\mathbf{u}_i \in U_i$. Now let $U = U_1 \cup \cdots \cup U_n$. By total ordering, one of these U_i contains all the others: this is U. Note that this only works because the subscript n is *finite*!

This argument (create an upper bound by taking the union over a chain before invoking Zorn's Lemma) is replicated in other areas of mathematics.[3] The results of the previous section may be generalized to cover infinite-dimensional vector spaces. A couple are outlined in the exercises.

[3] For instance, in abstract algebra to prove the existence of a maximal ideal in a ring.

Exercises 1.5 (Remember these are optional!)

1. As defined above, a set β is a *maximal linearly independent subset* of V if

   - β is linearly independent.
   - The only linearly independent subset of V containing β is β itself.

   The discussion on page 20 shows that every basis is a maximal linearly independent subset. Prove the converse:

   $$\beta \text{ a maximal linearly independent subset} \implies \beta \text{ is a basis}$$

   (*You cannot assume that β is finite: the entire point of this section is that it needn't be!*)

2. Show that categorization 4 on page 23 does not extend to infinite dimensions: specifically, state a linearly independent subset X of a vector space V such that $|X| = \dim V$, but such that X is *not* a basis of V.

3. Prove a more general version of Theorem 1.32: If β is a basis of V, then for all non-zero $\mathbf{v} \in V$ there is a unique *finite* subset $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \beta$ and unique non-zero scalars $a_1, \ldots, a_n$ such that

   $$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

   Our only freedom is in the order of the vectors $\mathbf{v}_i$.

   (*Hint: obtain a contradiction by supposing $\mathbf{v} \in V$ is a non-zero vector which can be written as a linear combination of elements of β in two different ways*)

4. Prove the infinite-dimensional version of the Extension Theorem: if X is a linearly independent subset of a vector space V, then there exists a basis of V which contains X.

   (*Hint: let $\mathcal{F}$ be the set of all linearly independent subsets of V which contain X, and mimic the proof of Theorem 1.47*)

5. Consider the set $X = \{e^{\lambda x} : \lambda \in \mathbb{R}\}$. Investigate the idea that X is a linearly independent set in the vector space of continuous functions on $\mathbb{R}$, and the relationship of this to the *Vandermonde matrix.* It follows that Span X is a subspace of $C(\mathbb{R})$ with *uncountably infinite dimension.*