Definition. A **vector space** is a non-empty set V, whose elements are called vectors, on which there are defined two operations:

1. **addition**, which to any two vectors v, w assigns a vector $v + w$, called the sum of v and w;
2. **scalar multiplication**, which to any number c and any vector v assigns a vector $c \cdot v$ (we will often write just cv);

and these operations satisfy the following properties:

V1. addition is associative, i.e. $(u + v) + w = u + (v + w)$ for any vectors u, v, w.
V2. addition is commutative, i.e. $v + w = w + v$ for any vectors v, w.
V3. there is a vector $0 \in V$ such that $0 + v = v$ for every vector v.
V4. for any vector v there is a vector w such that $v + w = 0$.
V5. $1 \cdot v = v$ for any vector v.
V6. $c \cdot (d \cdot v) = (cd) \cdot v$ for any numbers c, d and any vector v.
V7. $c \cdot (v + w) = c \cdot v + c \cdot w$ for any number c and any vectors v, w.
V8. $(c + d) \cdot v = c \cdot v + d \cdot v$ for any numbers c, d and any vector v.

It is not hard to see that the addition and scalar multiplication in vector spaces have additional familiar properties, for example:

- the vector 0 defined in V3. is unique with the required property. We call it the **zero vector**.
- for a given vector v, the vector w defined in V4. is unique with the required property. We call it the **negative** of v and denote it by $-v$.
- for any vector v we have $0 \cdot v = 0$. Note that the zero on the left is the number 0 and the zero on the right is the vector 0.
- for any vector v we have $-v = (-1) \cdot v$.
- $c \cdot 0 = 0$ for any number c.

We can justify the above properties as follows. Suppose first that 0 and 0_1 both have the property stated in V3. Then $0 + 0_1 = 0_1$ (apply V3 with $v = 0_1$). On the other hand $0 + 0_1 = 0_1 + 0$ by V2 and applying V3 with $v = 0$ yields $0 + 0_1 = 0_1 + 0 = 0$. Thus $0 = 0_1$.

Suppose now that w and w_1 both have the property stated in V4. Thus $0 = v + w$. Adding w_1 to both sides and using associativity we get

$$w_1 = w_1 + 0 = w_1 + (v + w) = (w_1 + v) + w = 0 + w = w.$$

Next note that $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v$. Adding $-0 \cdot v$ to both sides we get

$$0 = 0 \cdot v + (-0 \cdot v) = (0 \cdot v + 0 \cdot v) + (-0 \cdot v) = 0 \cdot v + (0 \cdot v + (-0 \cdot v)) = 0 \cdot v + 0 = 0 \cdot v.$$

Finally, we have $v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 + (-1)) \cdot v = 0 \cdot v = 0$, so $(-1) \cdot v = -v$ by uniqueness of the negative.

We leave the justification of the last property and other similar properties as exercise.

Definition. A subset U of a vector space V is called a **subspace**, if it is non-empty and for any $u, v \in U$ and any number c the vectors $u + v$ and cu are are also in U (i.e. U is closed under addition and scalar multiplication in V).

Clearly a subspace of a vector space is itself a vector space under the addition and scalar multiplication inherited from V.

Vector spaces appear often in mathematics and its applications and linear algebra provides tools to work with vector spaces.

Examples of vector spaces. The sets $\mathbb{R}^n$ with the usual addition and multiplication by scalars are fundamental examples of vector spaces.

The set of all functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space with the usual addition of functions and multiplication by numbers. The subset of all continuous functions is clearly a subspace. A subset of all differentiable functions is a subspace of the space of continuous functions. The subset of all polynomial functions is a subspace of the space of all differentiable functions. The subset of all polynomials of degree at most d (where d is any non-negative integer) is a subspace of the space of all polynomial functions. All these spaces naturally appear in calculus.

The set of all $m \times n$ matrices with the addition and multiplication by numbers we discussed in previous notes is a vector space.

If A is an $m \times n$ matrix then the set of all solutions to the homogeneous system of linear equations $Ax = 0$ is a subspace of $\mathbb{R}^n$.

If $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation then the image of L is a subspace of $\mathbb{R}^m$.

The last two examples are special cases of a more general result. To state it we introduce the following fundamental definition.

Definition. A **linear transformation** between vector spaces V and W is a function $L : V \longrightarrow W$ such that $L(u + v) = L(u) + L(v)$ and $L(cu) = cL(u)$ for any vectors u, v in V and any number c.

Clearly this definition is a straightforward generalization of the concept of linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. As in this special case, the linear transformations have the following properties.

Proposition. Let $L : V \longrightarrow W$ and $T : W \longrightarrow U$ be linear transformations.

1. $L(0) = 0$
2. $L(a_1v_1 + a_2v_2 + \ldots + a_sv_s) = a_1L(v_1) + a_2L(v_2) + \ldots + a_sL(v_s)$ for any vectors $v_1, v_2, \ldots, v_s \in V$ and any numbers $a_1, \ldots, a_s$.
3. the composition $T \circ L : V \longrightarrow U$ is a linear transformation.
4. if L is a bijection then the inverse function $L^{-1} : W \longrightarrow V$ is a linear transformation.

All these properties are justified in the same way as we did before for linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. For example,

$$(T \circ L)(u + v) = T(L(u + v)) = T(L(u) + L(v)) = T(L(u)) + T(L(v)) = (T \circ L)(u) + (T \circ L)(v)$$

for any u, v in V. This verifies one of the properties needed for $T \circ L$ to be a linear transformation. The other property is verified in a similar way.

To show that $L^{-1}(u + w) = L^{-1}(u) + L^{-1}(w)$ it suffices to show that when we apply L to each side of the last equation, we get equal results (since L is one-to-one):

$$L(L^{-1}(u) + L^{-1}(w)) = L(L^{-1}(u)) + L(L^{-1}(w)) = u + w = L(L^{-1}(u + w))$$

(we use the fact that $L(L^{-1}(x)) = x$ for any x).

The verification of the other properties is left as an exercise.

Definition. The kernel $\ker(L)$ of a linear transformation $L : V \longrightarrow W$ is the set of all vectors v in V such that $L(v) = 0$:

$$\ker(L) = \{v \in V : L(v) = 0\}.$$

Remark. The name **nullity** of L is sometimes used instead of the word "kernel".

The following proposition extends the observation we made for linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ made in our list of examples of vector spaces.

Proposition. Let $L : V \longrightarrow W$ be a linear transformation. Then

1. $\ker(L)$ is a subspace of V.
2. the image $L(V)$ of L is a subspace of W.

Indeed, let u, v be in $\ker(L)$ and let c be a number. Then

$$L(u+v) = L(u) + L(v) = 0 + 0 = 0, \text{ and } L(cv) = cL(v) = c \cdot 0 = 0$$

so $u + v \in \ker(L)$ and $cv \in \ker(L)$. This shows that $\ker(L)$ is a subspace of V (note that $0 \in \ker(T)$, so the kernel is non-empty).

Suppose now that w_1 and w_2 are in the image of L. This means that $w_1 = L(v_1)$ and $w_2 = L(v_2)$ for some $v_1, v_2 \in V$. Thus

$$w_1 + w_2 = L(v_1) + L(v_2) = L(v_1 + v_2), \text{ and } cw_1 = cL(v_1) = L(cv_1)$$

for any number c. This shows that both $w_1 + w_2$ and cw_1 belong to the image of L, i.e. the image is indeed a subspace of W (note that $0 = L(0)$ is in the image, so the image is non-empty).

The following observation is often useful.

Proposition. Let $L : V \longrightarrow W$ be a linear transformation. Then $L(u) = L(v)$ if and only if $u - v$ belongs to $\ker(L)$. In particular, L is one-to one if and only if $\ker(L) = \{0\}$, i.e. 0 is the only vector v in V such that $L(v) = 0$ (we say that the kernel of L is trivial in this case).

Indeed, $L(u) = L(v)$ if and only if $L(u - v) = 0$.

If L is one-to-one then there is at most one v such that $L(v) = 0$, so $\ker(L) = \{0\}$. Conversely, suppose that $\ker(L) = \{0\}$ and that $L(v) = L(u)$. Since $u - v$ is in the kernel of L, we have $u - v = 0$, i.e. $u = v$. This shows that L is one-to-one.

We end with the following example.

Example Let V be the space of all twice differentiable functions from $\mathbb{R}$ to $\mathbb{R}$ and let W be the space of all functions from $\mathbb{R}$ to $\mathbb{R}$. Define $L : V \longrightarrow W$ by $L(f) = f^{''} + f$. It is a simple calculus exercise that L is a linear transformation. Clearly $L(\sin x) = 0 = L(\cos x)$, so $\sin x$ and $\cos x$ are both in the kernel of L. It is not obvious, but can be proved, that

$$\ker(L) = \{a \sin x + b \cos x : a, b \in \mathbb{R}\}.$$

Describing the image of L is a challenging calculus problem.