LECTURE 1

Introduction

The rudiments of linear algebra are familar to every scientist who knows what a vector is and every software engineer who knows what an array is. In Math 3013 (Linear Algebra) these rudimentary ideas are abstracted and generalized so that the notions of vector spaces and vector space operations can be seem as uniformly applicable to a wide variety of situations. This course, Math 4063/5023 builds on the abstract point of view developed in Math 3013, extending the concepts developed there to an even broader range of applicability.[1]

So perhaps the most instructive important thing to do at the beginning of this course is to provide an overview of the developments that took place in Math 3013.

The first half of Math 3013 deals with the vector space $\mathbb{R}^n$.

DEFINITION 1.1. *The vector space* $\mathbb{R}^n$ *is the set of ordered lists of* n *real numbers*[2]*. For any real number* λ *and any vector* $\mathbf{v} = [v_1, v_2, \ldots, v_n]$ *in* $\mathbb{R}^n$ *the scalar multiple of* $\mathbf{v}$ *by* λ *is the vector* $\lambda\mathbf{v} := [\lambda v_1, \lambda v_2, \ldots, \lambda v_n]$*. For any pair of vectors* $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ *and* $\mathbf{v} = [v_1, v_2, \ldots, v_n]$*, the vector sum of* $\mathbf{u}$ *and* $\mathbf{v}$ *is defined by* $\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n]$.

While $\mathbb{R}^n$ is only one example of a vector space, it is fundamental in the sense that

- Calculations in $\mathbb{R}^n$ are easy if not rout.
- Because $\mathbb{R}^n$ provides an easy, concrete, computational setting, abstract ideas like
  - linear independence
  - bases
  - dimension
  - linear transformations

  can be given a firm footing by direct computations.

I should also stress at this point that $\mathbb{R}^n$ in not just the set of n-tuples of real numbers. Just as important to the data used to specify elements of $\mathbb{R}^n$ are the operations of scalar multiplication and vector addition.

Also, the specialization to $\mathbb{R}^n$ in the beginning of Math 3013 is merely heuristic and utilitarian. About midway through Math 3013, Definition 1.1 is abandoned for a more general, axiomatic approach to vectors. The first step in this axiomatization process is to display some fundamental properties of $\mathbb{R}^n$. Given Definition 1.1, and the familiar arithmetic properties of the real numbers (commutativity, associativity, distributive law, etc) one has

THEOREM 1.2. *Let* $\mathbb{R}^n$ *be the set of ordered lists of* n *real numbers endowed with the operations of scalar multiplication and vector addition as per Defintion 1.1.*

(1) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ *for all vectors* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ *(commutativity of vector addition);*
(2) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *for all vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ *(associativity of vector addition);*

[1] Another major difference between this course and Math 3013 will be the strong emphasis on proofs.

[2] By an ordered list, we simply mean a list of objects where the order in which the objects appear makes a difference. Thus, $[1, 3]$ does not equal $[3, 1]$ as an ordered list. In the text (and elsewhere) ordered lists of n objects are often referred to as **n-tuples**

(3) *There exists a vector* $\mathbf{0}$ *such that* $\mathbf{v}+\mathbf{0}=\mathbf{v}$ *for all* $\mathbf{v}\in\mathbb{R}^n$*;*
(4) *For each vector* $\mathbf{v}$*, there exists a vector* $-\mathbf{v}$ *with the property that* $\mathbf{v}+(-\mathbf{v})=\mathbf{0}$*;*
(5) $\lambda(\nu\mathbf{v})=(\lambda\nu)\cdot\mathbf{v}$ *for all* $\lambda,\nu\in\mathbb{R}$ *and all* $\mathbf{v}\in\mathbb{R}^n$ *(associativity of scalar multiplication)*
(6) $(\lambda+\nu)\mathbf{v}=(\lambda\mathbf{v})+(\nu\mathbf{v})$ *for all* $\lambda,\nu\in\mathbb{R}$ *and all* $\mathbf{v}\in\mathbb{R}^n$ *(distributivity of scalar addition w.r.t. scalar multiplication)*
(7) $\lambda(\mathbf{u}+\mathbf{v})=(\lambda\mathbf{u})+(\lambda\mathbf{v})$ *for all* $\lambda\in\mathbb{R}$ *and all* $\mathbf{u},\mathbf{v}\in\mathbb{R}^n$ *(distributivity of vector addition w.r.t. scalar multiplications);*
(8) $1\cdot\mathbf{v}=\mathbf{v}$ *for all vectors* $\mathbf{v}$ *(scalar mulitplication by* 1 *is trivial).*

Each of this properties is easy to prove for the vector space $\mathbb{R}^n$ (you just calculate and compare both sides the stated identities using the definitions of scalar multiplication and vector addition appearing in Definition 1.1.) However, the main point is that there exists a grand variety of sets with other notions of scalar multiplications and vector addition and which behave in essentially the the same way.

Example 1.3. Let $\mathcal{P}$ be the set of polynomials with real coefficients. If we define "scalar multiplication" of a polynomial $p=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ by real number λ in the natural way

$$(\lambda\cdot p):=(\lambda a_n)x^n+(\lambda a_{n-1})x^{n-1}+\cdots+(\lambda a_1)x+(\lambda a_0)$$

and define "vector addition" of two polynomials $p_1=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ and $p_2=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$

$$(p_1+p_2):=(a_n+b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1}+\cdots+(a_1+b_1)x+(a_0+b_0)$$

The set $\mathcal{P}$ with the operations of scalar multiplication and vector addition enjoys the same properties as the set $\mathbb{R}^n$ of Theorem 1.2 (just replace the vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ appearing the statements of Theorem 1.2 with polynomials p_1,p_2,p_3 to get the analogous statements for $\mathcal{P}$).

Example 1.4. Let $\mathcal{C}(\mathbb{R})$ be the set of functions on the real line. Define scalar multiplication and vector addition for functions in $\mathcal{C}(\mathbb{R})$ by

$$\begin{aligned}\lambda\cdot f\ &:\ =\text{the function whose value at a point } x\in\mathbb{R} \text{ is the number } \lambda\cdot f(x)\\ f+g\ &:\ =\text{the function whose value at a point } x\in\mathbb{R} \text{ is the number } f(x)+g(x)\end{aligned}$$

One then finds, after replacing the vectors $\mathbf{u},\mathbf{v},\mathbf{w}$ in Theorem 1.2 with arbitary functions $f,g,h\in\mathcal{C}(\mathbb{R})$ that each of the 8 properties still hold.

And so sets of polynomials and sets of functions can be made to behave like $\mathbb{R}^n$. However, a much more democratic way of saying this is any set V with notions of scalar multiplication and vector addition that satisfies the 8 properties of Theorem 1.2 is a **vector space over** $\mathbb{R}$.

Indeed,

Definition 1.5. *A* ***real vector space*** *is a set* V *with two operations defined on it,*

$$\begin{aligned}+\ &:\ V\times V\to V\\ \cdot\ &:\ \mathbb{R}\times V\to V\end{aligned}$$

satisfying the following 8 properties:

Theorem 1.6. (1) $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ *for all vectors* $\mathbf{u},\mathbf{v}\in V$ *(commutativity of vector addition);*
(2) $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$ *for all vectors* $\mathbf{u},\mathbf{v},\mathbf{w}\in V$ *(associativity of vector addition);*
(3) *There exists a vector* $\mathbf{0}$ *such that* $\mathbf{v}+\mathbf{0}=\mathbf{v}$ *for all* $\mathbf{v}\in V$*;*
(4) *For each vector* $\mathbf{v}$*, there exists a vector* $-\mathbf{v}$ *with the property that* $\mathbf{v}+(-\mathbf{v})=\mathbf{0}$*;*
(5) $\lambda(\nu\mathbf{v})=(\lambda\nu)\cdot\mathbf{v}$ *for all* $\lambda,\nu\in\mathbb{R}$ *and all* $\mathbf{v}\in V$ *(associativity of scalar multiplication)*
(6) $(\lambda+\nu)\mathbf{v}=(\lambda\mathbf{v})+(\nu\mathbf{v})$ *for all* $\lambda,\nu\in\mathbb{R}$ *and all* $\mathbf{v}\in V$ *(distributivity of scalar addition w.r.t. scalar multiplication)*
(7) $\lambda(\mathbf{u}+\mathbf{v})=(\lambda\mathbf{u})+(\lambda\mathbf{v})$ *for all* $\lambda\in\mathbb{R}$ *and all* $\mathbf{u},\mathbf{v}\in V$ *(distributivity of vector addition w.r.t. scalar multiplications);*
(8) $1\cdot\mathbf{v}=\mathbf{v}$ *for all vectors* $\mathbf{v}\in\mathbf{V}$ *(scalar mulitplication by* 1 *is trivial).*

Now Theorem 1.2 allows us to view $\mathbb{R}^n$ as a particular example of an abstract vector space; and the other possibilities for an abstract vector space arise by looking at other sets with notions of (internal) addition and multiplication by $\mathbb{R}$ and verifying that properties 1 – 8 are satisfied.

QUESTION 1.7. *In Example 1.3 we think of a polynomial as a set of coefficients times various powers of x. Can you think of a way of regarding a polynomial as an ordered list of numbers (that is as a vector in $\mathbb{R}^n$)? What's the hedge here? One can also think of a polynomial as defining a function. If we do so, are the notions of scalar multiplication and vector addition defined for polynomials compatible with the corresponding notions for functions in $\mathcal{C}(\mathbb{R})$?*

Let me offer one more example, just to show that the notion of a vector space is not restricted to purely mathematical situations.

EXAMPLE 1.8. Let $\mathcal{S}$ denote the set of possible vibrational modes of a fixed string. Define scalar muliplication of a vibrational mode m by a non-negative real number λ as the mode that sounds the same as m but whose amplitude has changed by a factor λ, and define scalar multiplication of a vibrational mode by a negative number λ as the mode obtained by changing the amplitude of m by a factor $|\lambda|$ and then changing the phase by 180°. Define the vector sum of two vibrational modes m_1 and m_2 as the superposition of the two modes (that is the vibrational mode for which the displacement of the string at any given point x and time t is just the sum of the displacements of m_1 and m_2. Then with the notions of operations of scalar multiplication and vector addition thus defined, the vibration modes of a fixed string enjoy properties 1 - 8 of Theorem 1.2.

In this course, we follow the same axiomatic approach to vector spaces; except that we take one step further back; we don't necessarily work over vector spaces defined over the real numbers. Indeed, in many physical applications, quantum mechanics in particular, one works over the complex numbers. But there are, in fact, many other types of "numbers" besides $\mathbb{C}$ and $\mathbb{R}$. Moreover, it turns outs that the choice of the set we use for scalar multiplication is sort of immaterial to the general theorems one can prove; so why specialize scalar multiplication to the real numbers and obtain only a special case of a general situation?

What we shall use in the place of the set of real numbers is the notion of a **field**.

DEFINITION 1.9. *A **field** is a set $\mathbb{F}$ with two operations defined; **addition** and **multiplication**, which we shall denote, respectively,by $+_{\mathbb{F}}$ and $*_{\mathbb{F}}$ (just so you resist the temptation to think purely in terms of the real numbers). These two operations are required to satisfy*

(1) $\alpha +_{\mathbb{F}} \beta = \beta +_{\mathbb{F}} \alpha$ *for all* $\alpha, \beta \in \mathbb{F}$ *(commutativity of addition);*
(2) $\alpha +_{\mathbb{F}} (\beta +_{\mathbb{F}} \gamma) = (\alpha +_{\mathbb{F}} \beta) +_{\mathbb{F}} \gamma$ *for all* $\alpha, \beta, \gamma \in \mathbb{F}$ *(associativity of addition);*
(3) $\alpha *_{\mathbb{F}} \beta = \beta *_{\mathbb{F}} \alpha$ *for all* $\alpha, \beta \in \mathbb{F}$ *(commutativity of multiplication)*
(4) $\alpha *_{\mathbb{F}} (\beta *_{\mathbb{F}} \gamma) = (\alpha *_{\mathbb{F}} \beta) *_{\mathbb{F}} \gamma$ *for all* $\alpha, \beta, \gamma \in \mathbb{F}$ *(associativity of multiplication);*
(5) $\alpha *_{\mathbb{F}} (\beta +_{\mathbb{F}} \gamma) = (\alpha *_{\mathbb{F}} \beta) +_{\mathbb{F}} (\alpha *_{\mathbb{F}} \gamma)$ *for all* $\alpha, \beta, \gamma \in \mathbb{F}$ *(distributivity of multiplication over addition);*
(6) *There exists an element* 0_F *of* F *such that* $\alpha + 0_F = \alpha$ *for all* $\alpha \in \mathbb{F}$ *(additive identity element);*
(7) *For each element* $\alpha \in \mathbb{F}$ *there is an element* $-\alpha \in \mathbb{F}$ *such that* $\alpha +_{\mathbb{F}} (-\alpha) = 0_{\mathbb{F}}$ *(existence of additive inverses);*
(8) *There exists an element* $1_{\mathbb{F}} \in \mathbb{F}$ *such that* $1 *_{\mathbb{F}} \alpha = \alpha$ *for all* $\alpha \in \mathbb{F}$ *(multiplicative identity element);*
(9) *For each* $\alpha \neq 0_F$ *in* $\mathbb{F}$ *there is an element* $\alpha^{-1} \in \mathbb{F}$ *such that* $\alpha *_{\mathbb{F}} \alpha^{-1} = 1_{\mathbb{F}}$.

REMARK 1.10. My apologies, but you will almost never see addition and multiplication in a field denoted by, respectively, $+_{\mathbb{F}}$ and $*_{\mathbb{F}}$. My purpose in using this elaborate notation is just to emphasize that the set $\mathbb{F}$ is not necessarily a set of numbers and so the operations we call addition and multiplication in $\mathbb{F}$, are not necessarily arithmetic operations on numbers.

EXAMPLES 1.11. $\mathbb{R}$, the set of real numbers, is a field and, in fact, $\mathbb{R}$ is the basic prototype that this definition tries to generalize.

The set $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ of integers is close to being a field; however, it does not satisfy property (8). For example, there is no integer z such that $z \cdot 2 = 1$. On the other hand, if one enlarges the set of integers to the set of rational numbers

$$\mathbb{Q} = \{p/q \mid p, q \in \mathbb{Z} \quad \text{and} \quad q \neq 0\}$$

then all 7 properties of a field are satisfied and so $\mathbb{Q}$ is a field.

Let i be the (actually a) square root of -1; viz; $i^2 = -1$. Set

$$\mathbb{C} = \left\{x + iy \mid x, y \in \mathbb{Z} \quad , \quad i^2 = -1\right\}$$

Define addition and multiplication in $\mathbb{C}$ by

$$\begin{aligned} (x_1 + iy_1) + (x_2 + iy_2) &= (x_1 + x_2) + i\,(y_1 + y_2) \\ (x_1 + iy_1) \cdot (x_2 + iy_2) &= (x_1x_2 - y_1y_2) + i\,(x_1y_2 + x_2y_1) \end{aligned}$$

Then with these two operators so defined $\mathbb{C}$ is a field. (I note that even property (8) is satisfied by taking $(x + iy)^{-1} = \frac{x}{x^2+y^2} - i\frac{y}{x^2+y^2}$.)

As a final example, that field elements do not necessarily have to be numbers. They could for example be families of numbers. Recall first that the integers themselves are not a field. Consider then the simple rules for combining even and odd integers:

$$\begin{aligned} (\text{an even integer}) + (\text{an even integer}) &= \text{an even integer} \\ (\text{an even integer}) + (\text{an odd integer}) &= \text{an odd integer} \\ (\text{an odd integer}) + (\text{an odd integer}) &= \text{an even integer} \end{aligned}$$

$$\begin{aligned} (\text{an even integer}) * (\text{an even integer}) &= \text{an even integer} \\ (\text{an even integer}) * (\text{an odd integer}) &= \text{an even integer} \\ (\text{an odd integer}) * (\text{an odd integer}) &= \text{an odd integer} \end{aligned}$$

So if we use e and o to symbolize, respectively, the set of even, respectively odd, integers, then we have the following addition and mulitplication tables:

$+$	e	o
e	e	o
0	o	e

,

$*$	e	o
e	e	e
0	e	o

If we define addition and multiplication for the set $\mathbb{Z}_2 := \{e, o\}$ in this way and take $0_{\mathbb{Z}_2} := e$ and $1_{\mathbb{Z}_2:} = o$, then $\mathbb{Z}_2$ is a field.

Okay, we can now define the arena in which this course on Advanced Linear Algebra takes place.

DEFINITION 1.12. *Let $\mathbb{F}$ be a field, and let V be a set upon which two operations are defined*

(i) ***vector addition**: a rule for combining two elements of V to get another element of V;*
(ii) ***scalar multiplication:** a rule for taking an element of $\mathbb{F}$ and an element of V and producing an element of V.*

*V is a **vector space over** $\mathbb{F}$ if the following 8 properties are satisfied:*

(1) *$u + v = v + u$ for all elements $u, v \in \mathbb{V}$ (commutativity of vector addition);*
(2) *$(u + v) + w = u + (v + w)$ for all elements $u, v, w \in V$ (associativity of vector addition);*
(3) *There exists a vector 0_V such that $v + 0_V = v$ for all $v \in V$;*
(4) *For each vector v, there exists a vector $-v$ with the property that $v + (-v) = 0_V$;*
(5) *$\alpha\,(\beta v) = (\alpha\beta) \cdot v$ for all $\alpha, \beta \in \mathbb{F}$ and all $v \in V$ (associativity of scalar multiplication)*
(6) *$(\alpha + \beta)\,v = (\alpha v) + (\beta v)$ for all $\alpha, \beta \in \mathbb{F}$ and all $v \in V$ (distributivity of scalar addition w.r.t. scalar multiplication)*
(7) *$\alpha\,(u + v) = (\alpha u + \beta v)$ for all $\alpha \in \mathbb{F}$ and for all $u, v \in V$ (distributivity of vector addition w.r.t. scalar multiplications);*

(8) $1_{\mathbb{F}} \cdot v = v$ *for all vectors* $v \in V$ *(scalar mulitplication by* 1 *is trivial).*

EXAMPLE 1.13. Here is the natural generalization of the vector space $\mathbb{R}^n$. Let $\mathbb{F}$ be a field and let $\mathbb{F}^n$ be the set of ordered lists of n elements of $\mathbb{F}$

$$\mathbb{F}^n = \{[\alpha_1, \alpha_2, \ldots, \alpha_n] \mid \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{F}\}$$

endowed with the following rule for scalar multiplication by elements of $\mathbb{F}$

$$\beta * [\alpha_1, \alpha_2, \ldots, \alpha_n] = [\beta \otimes \alpha_1, \beta \otimes \alpha_2, \ldots, \beta \otimes \alpha_n]$$

and the following rule for adding elements of $\mathbb{F}^n$

$$[\alpha_1, \alpha_2, \ldots, \alpha_n] + [\beta_1, \beta_2, \ldots, \beta_n] = [\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n] \quad .$$

Then $\mathbb{F}^n$ is a vector space over $\mathbb{F}$.

Below is a simple proposition and then two simple identities we'll use constantly latter on.

PROPOSITION 1.14. *The zero vector* $\mathbf{0}_{\mathbb{V}}$ *of a vector space is unique.*

Proof. Suppose we had two vectors $\mathbf{0}_V$ and $\mathbf{0}'_V$ with the property that when added to an arbitrary vector $v \in V$, the vector v is reproduced. Then, in particular, we could add $\mathbf{0}_{\mathbb{F}}$ to $\mathbf{0}'_{\mathbb{F}}$ to get

$$\mathbf{0}_{\mathbb{F}} + \mathbf{0}'_{\mathbb{F}} = \begin{cases} \mathbf{0}'_{\mathbb{F}} & \text{if we use the fact that } \mathbf{0}_{\mathbb{F}} \text{ acts like an additive identity in } V \\ \mathbf{0}_{\mathbb{F}} & \text{if we use the fact that } \mathbf{0}'_{\mathbb{F}} \text{ acts like an additive identity in } V \end{cases}$$

So

$$\mathbf{0}_{\mathbb{F}} = \mathbf{0}'_{\mathbb{F}} \quad .$$

□

PROPOSITION 1.15. *Let* V *be a vector space over a field* $\mathbb{F}$. *Then* $0_F \cdot v = \mathbf{0}_V$ *for all* $v \in V$.

Proof. Let $v \in V$ be arbitary. We have

$$v = (0_{\mathbb{F}} + 1_{\mathbb{F}})\, v = 0_{\mathbb{F}} \cdot v + v$$

Adding $-v$ to both sides yields

$$\mathbf{0}_V = v + (-v) = 0_F \cdot v + v + (-v) = 0_F \cdot v$$

Thus, $0_F \cdot v = \mathbf{0}_V$ as desired. □

PROPOSITION 1.16. *Let* V *be a vector space over a field.* $\mathbb{F}$. *Then* $\alpha \cdot \mathbf{0}_V = \mathbf{0}_V$ *for all* $\alpha \in \mathbb{F}$.

Proof. For any $v \in V$ we have

$$v + \mathbf{0}_V = v \quad .$$

Multiplying both sides by $\alpha \in \mathbb{F}$ we have

$$\alpha \cdot v + \alpha \cdot \mathbf{0}_V = \alpha \cdot v$$

then adding $-\alpha \cdot v$ to both sides yields

$$\alpha \cdot \mathbf{0}_V = -\alpha \cdot v + \alpha \cdot v + \alpha \cdot \mathbf{0}_V = -\alpha \cdot v + \alpha \cdot v = \mathbf{0}_V$$

and so the conclusion follows. □