Lecture 1

1 Introduction:

This is the second course in mathematics for you in IIT. In the first course itself you have studied some basic properties such as continuity of functions of several variables. The central theme of this course is to study thoroughly geometric objects in higher dimensions and also functions. As you know the simplest geometric objects are lines and planes. Correspondingly, the simplest of functions which have sufficiently interesting properties (not constants!) are linear functions. Therefore, we shall start our study with that of Linear Algebra. To begin with Linear Algebra brings unified approach to certain topics which are familiar to you such as coordinate geometry and vector algebra. Apart from playing a very crucial role in the basic understanding of the calculus of several variables, Linear Algebra has its own importance with applications in almost all scientific studies. So, in the first part of the course, let us lay a proper foundation of Linear Algebra with basic definitions of vector spaces, linear transformations, linear dependence, dimension, matrices, determinants etc. Some illustrations of applications to geometric problems will be given with the study of quadratic forms. We shall also give an elementary proof of fundamental theorem of algebra using linear algebra that you learn in this course, thus achieving a kind of poetic justice since Fundamental theorem of algebra has to be used in a non trivial fashion in Linear algebra.

1.1 The Cartesian Coordinate Space:

Today let us get familiar with our 'working space'. Thanks to DesCarte, of the fame *'cogito ergo sum' (I think therefore I am)*, we can refer to each and every corner of our space, in a very precise manner on our finger tips! The n-dimensional Cartesian coordinate space $\mathbb{R}^n$ is the totality of all ordered n-tuples $\mathbf{x} := (x_1, x_2, \ldots, x_n)$ of real numbers $x_i \in \mathbb{R}$. It is denoted by $\mathbb{R} \times \cdots \times \mathbb{R}$ (n factors) or, in short, by $\mathbb{R}^n$. Elements of this space are often referred to as 'n-vectors'. For $n = 1$ you see that $\mathbb{R}^1$ is nothing but the real line and a real number can also be referred to as a 1-vector. Of course, in this case, we shall not use the absurdly complicated notation (x_1) but merely write x_1. Similarly for $n = 2$ and 3, we can avail of the familiar notation $(x, y) \in \mathbb{R}^2$ and $(x, y, z) \in \mathbb{R}^3$ etc., respectively. For a point $\mathbf{x} = (x_1, \ldots, x_n)$ we refer to x_i as the i^{th} *coordinate* of $\mathbf{x}$. This then, is the first example of a function from $\mathbb{R}^n$ to $\mathbb{R}$. For each $i = 1, 2, \ldots, n$, let us denote this function by $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $\pi_i((x_1, \ldots, x_n)) = x_i$. It is called the *$i^{th}$ coordinate function* or the *i^{th} coordinate projection.*

Given any function $f : A \longrightarrow \mathbb{R}^n$, where A is any set, we can compose it with the i^{th} coordinate projection to get n functions $f_i := \pi_i \circ f$. An important observation is that f is completely determined by these functions f_i called the i^{th} *component function of* f, viz., $f(a) = (f_1(a), \ldots, f_n(a))$, $\forall\ a \in A$. Thus one can expects that the study of functions $f : A \longrightarrow \mathbb{R}^n$ can be effectively reduced to the study of functions $f_i : A \longrightarrow \mathbb{R}$, n of them in number, and perhaps to the study of their inter-relations. Soon we shall see that this is in fact so.

For $n < m$ we can think of $\mathbb{R}^n$ as a subset of $\mathbb{R}^m$ in various ways. One such standard way is to identify it with the set of points $\mathbf{y}$ of $\mathbb{R}^m$ whose last $m - n$ coordinates are

identically zero. Indeed, the *'inclusion map'* $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, 0, \ldots 0)$ defines such an identification. Observe that, more generally, we could also have chosen any other $m - n$ coordinates to vanish identically, to get other inclusion maps. In particular, if $n = 1$, then we get the m *coordinate axes* of $\mathbb{R}^m$ for $i = 1, 2, \ldots, m$.

We presume that you are familiar with basic set theory, composition of functions, inverses etc.. A good reference is the I volume of Apostol's Calculus.

Example 1
(1) As a refresher, let us consider the following example: Let

$$f(x, y) = (x^2y, x + y);\ g(x, y) = (\cos(x + y), \sin(x/y)).$$

Then obviously the domain of definition of f is the whole of $\mathbb{R}^2$, whereas that of g is $\mathbb{R}^2 \setminus \{(x, 0) : x \in \mathbb{R}\}$. Therefore, neither $f \circ g$ nor $g \circ f$ is defined on $\mathbb{R}^2$. But if we restrict the domain of f to $\mathbb{R}^2 \setminus \{(x, y) : x + y = 0\}$, then $g \circ f$ defined. In that case we have, $g \circ f(x, y) = (\cos(x^2y + x + y), \sin(x^2y/x + y))$. Can you determine the natural domain for $f \circ g$ and write down the expression for it?
(2) Let $N = \{1, 2, \ldots, n\}$ be the set of first n natural numbers. Consider the cartesian product of n copies of N :

$$N^n = N \times \cdots \times N$$

How many elements does N^n has? Next consider the set $S(n)$ of all sequences of length n with values in $\{1, 2, \ldots, n\}$. Can you see any relation between N^n and $S(n)$? Further look at the set $F(N, N)$ of all functions from N to N. Is there any relation between this set and the earlier two? Indeed, there are natural ways of getting one-to-one mappings of any one of these three sets into another. Let $\Sigma(n)$ denote the subset of $F(N, N)$ which are one-to-one. What does it correspond to in N^n and $S(n)$?

Exercise:
(i) Let $\psi(x, y) = (x + y^2, y + x^2 + 2xy^2 + y^4)$. Is ψ a bijective map?
(ii) (Here is an optional exercise which you can take up as a challenge. Do not consult the tutors for the solution, since they have been told not to help you with this problem. If you have solved it you may directly contact your Instructor.) Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a continuous map. Then there exist $a, b \in \mathbb{R}$ such that for all $r \in \mathbb{R} \setminus \{a, b\}$, $f^{-1}(r)$ is either $\emptyset$ or infinite. Hence or otherwise prove that
(a) for every n there exist at least n points which are mapped to the same point by f;
(b) if f is surjective then $f^{-1}(r)$ is infinite for all $r \in \mathbb{R}$.
(c) Finally construct a continuous map $\mathbb{R}^2 \longrightarrow \mathbb{R}$ to illustrate that improvement on this statement is not possible. Explicitly, find a continuous function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ such that $f^{-1}(-1) = \{-1\}$ and $f^{-1}(1) = \{1\}$.

1.2 Algebraic Structure of $\mathbb{R}^n$

Some of the algebraic structures of the real numbers induce similar structure on $\mathbb{R}^n$. The most important amongst them is the *'sum' or 'addition.'* Given two n-vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ we define their sum $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$. It is straightforward to verify that this addition obeys all the usual laws that the addition of real numbers obey.

In particular, observe that the zero element for this sum is the zero vector $\mathbf{0} = (0, \ldots, 0)$ and the negative of $(x_1, \ldots, x_n)$ is $(-x_1, \ldots, -x_n)$.

We can of course try to define a 'multiplication' in $\mathbb{R}^n$ using the multiplication of real numbers, as above, viz., $\mathbf{xy} := (x_1y_1, \ldots, x_ny_n)$. However, what we are presently looking for is something much simpler than this viz., the so called *scalar multiplication;* for each real number α and a vector $\mathbf{x}$, we define $\alpha\mathbf{x} = (\alpha x_1, \ldots, \alpha x_n)$. It is

1. *associative:* $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$;
2. *distributes with the sum:* $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$, and $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$;
3. and has *an identity* element $1\mathbf{x} = \mathbf{x}$.

In particular, we also observe that, if $n = 1$, then these two operations coincide with the usual addition and multiplication.

1.3 The Geometry of $\mathbb{R}^n$

Before starting of with any sort of calculus, we need one more important ingredient viz., the distance function on $\mathbb{R}^n$. For any two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ as above, we define the distance between $\mathbf{x}$ and $\mathbf{y}$ by

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Indeed, the distance function is closely related to the *dot product*

$$(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \cdot \mathbf{y} := \sum_i x_i y_i.$$

One defines the *norm* function by

$$\|\mathbf{x}\| := d(\mathbf{x}, \mathbf{0}) = \sqrt{\sum_{i}^{n} x_i^2} = \sqrt{\mathbf{x} \cdot \mathbf{x}}.$$

Then it follows that $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. (Such a distance function is called a *linear metric.*

You are already familiar with various properties of this function such as

(d1) *symmetry:* $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;
(d2) *triangle inequality:* $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{w}) + d(\mathbf{w}, \mathbf{y})$;
(d3) *positivity:* $d(\mathbf{x}, \mathbf{y}) \geq 0$ and $= 0$ if and only if $\mathbf{x} = \mathbf{y}$.

Moreover, we also have

(d4) *homogeneity:* $d(\alpha\mathbf{x}, \alpha\mathbf{y}) = |\alpha| d(\mathbf{x}, \mathbf{y})$.

All these properties are standard and can be verified easily. For instance, in the proof of triangle inequality, we can use the famous Cauchy-Schwarz inequality:

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\| \tag{1}$$

The space $\mathbb{R}^n$ together with the two algebraic structures and the distance function d is called the n-dimensional *euclidean space.* This is going to be our work-space for quite some time.

You are quite familiar with the case $n = 2$ and $n = 3$. Most of the concepts you have met in these two special cases have appropriate and obvious generalizations for all n. However some caution is necessary, particularly while relying on 'pictures'.

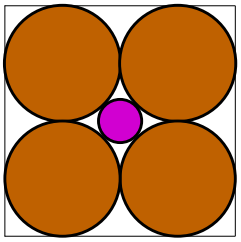
Exercises:
(i) Define

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

and

$$D(\mathbf{x}, \mathbf{y}) = max\,\{|x_i - y_i| \ : \ 1 \leq i \leq n\}.$$

Show that δ, D satisfy properties (d1), (d2), (d3) above in place of d.
(ii) On the set of 64 squares of a chess board define the distance d from one square to another to be the least number of knight moves required. Check that this distance function satisfy (d1), (d2), (d3). Determine the diameter of the chess board with respect to this distance where diameter of a space is the supremum of all values of d.
(iii) Here is an example to show that it is not advisable to rely too much on intuition when dealing with higher dimensions. Consider a square of side 4 units and place four coins of unit radius one in each corner so as to touch two of the sides. Of course each coin touches other two coins. Now place a coin at the center of the square so as to touch all the four coins. Do this inside a n-dimensional cube of side 4 units with 2^n n-dimensional balls of unit radius. For $n = 2, \ldots, 9$ the central ball which is kept touching all the balls in the corner lies inside the cube. The surprise is that for $n > 9$ the central ball cannot fit inside the cube. Prove this by showing that the radius of the central ball is $\sqrt{n} - 1$.



Lecture 2

2 Linear Maps on Euclidean Spaces and Matrices

2.1 Basic Definitions and Examples

An important class of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ are the so called *linear transformations*. Though this word makes sense even in a larger context, we shall first study them on euclidean spaces.

Definition 1 A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is said to be a *linear map* if $f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y})$, for all scalars $\alpha, \beta \in \mathbb{R}$ and all vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Equivalently, this can also be said in the following way:
1. $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ (*additivity*).
2. $f(\alpha\mathbf{x}) = \alpha f(\mathbf{x})$(*homogeneity*).

Exercise: Show that if f is a linear map then $f(\sum_{i=1}^k \alpha_i \mathbf{x}_i) = \sum_{i=1}^k \alpha_i \mathbf{x}_i$.

Example 2 Examples of linear maps are easy to produce. The projection maps π_i, the inclusion maps etc. seen during the last lecture are the standard examples. The identity map, the zero map and multiplication by a fixed scalar α are also linear. More generally, taking dot product with a fixed vector is a linear map $\mathbb{R}^n \longrightarrow \mathbb{R}$. Indeed, we shall soon see that any linear map $\mathbb{R}^n \longrightarrow \mathbb{R}$ corresponds to taking dot product with a unique vector in $\mathbb{R}^n$. Finally, a map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is linear if and only if its component maps $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ are all linear. In this way, we will be able to determine all linear maps $\mathbb{R}^n \longrightarrow \mathbb{R}^m$.

Distance travelled is a linear function of time when velocity is constant. So is the voltage as a function of resistance when the current is constant. The logarithm of the change in concentration in any first order chemical reaction is a linear function of time.

Non examples are even easier to produce. On $\mathbb{R}$, $|x|, x^n (n \geq 2), \sin x$ etc. are easily seen to be non linear maps. Gravitational force between two bodies is not a linear function of the distance between the two bodies.

Exercises:
(i) Show that the projection on a line L passing through the origin defines a linear map of $\mathbb{R}^2$ to $\mathbb{R}^2$ and its image is equal to L.
(ii) Show that rotation through a fixed angle θ is a linear map, $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$.
(iii) By a *rigid motion* of $\mathbb{R}^n$ we mean a map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $d(f(x), f(y)) = d(x, y)$. Show that a rigid motion of $\mathbb{R}^3$ which fixes the origin is a linear map.

2.2 The Structure of Linear Maps

Let $\mathbb{L}(n, m)$ denote the totality of all the linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. The algebraic structure on $\mathbb{R}^m$ induces similar structure on $\mathbb{L}(n, m)$ also: For $\alpha \in \mathbb{R}, f, g \in \mathbb{L}(n, m)$ we can define αf and $f + g$ by the formula,

$$(\alpha f)(\mathbf{x}) = \alpha f(\mathbf{x}); \quad (f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}).$$

Moreover, given $f \in L(n,m), g \in \mathbb{L}(m,l)$ we can define the composite function $g \circ f$ and it is not difficult to verify that this is a linear map; thus $g \circ f \in \mathbb{L}(n,l)$. However, we should be careful, while considering such new operations. For instance, if $f, g \in \mathbb{L}(n,1)$ then we can define $fg : \mathbb{R}^n \longrightarrow \mathbb{R}$ by the formula $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x}),\ \forall\ \mathbf{x} \in \mathbb{R}^n$. But can we say that $fg \in \mathbb{L}(n,1)$?

The structure of linear maps is indeed, very simple. First of all consider the elements $\mathbf{e}_i = (0,\ldots,0,1,0,\ldots,0)$, wherein we have put 1 in the i^{th}-place. These elements, for $i = 1,2,\ldots,n$ are called the *standard basis* elements for $\mathbb{R}^n$. It is easily seen that every element $\mathbf{x} := (x_1,\ldots,x_n) \in \mathbb{R}^n$ can be written in a unique way, as a *linear combination* of $\mathbf{e}_1,\ldots\mathbf{e}_n$, viz.,

$$\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i. \tag{2}$$

Now given a linear map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, the linearity of f implies that

$$f(\mathbf{x}) = \sum_{i} x_i f(\mathbf{e}_i). \tag{3}$$

This means that to know the effect of f on any arbitrary element we have only to know the n values $f(\mathbf{e}_1),\ldots,f(\mathbf{e}_n) \in \mathbb{R}^m$. Conversely, if we are given any n vectors $\mathbf{v}_1,\ldots,\mathbf{v}_m \in \mathbb{R}^m$, in a particular order, then we can define a (unique) linear map f which has its values on $\mathbf{e}_i$ equal to $\mathbf{v}_i$ (taken in that order) by the formula (3).

Example 3
(i) When $m = 1 = n$, a linear map $f : \mathbb{R} \longrightarrow \mathbb{R}$ is determined by its value on $\mathbf{e}_1 = 1$. In fact, we have $f(r) = rf(1)$. Conversely, multiplication by each fixed real number is a linear map. Therefore the only linear maps $\mathbb{R} \longrightarrow \mathbb{R}$ are those obtained by multiplication by a fixed scalar.
(ii) Consider the case $n \geq 2$, and $m = 1$. Put $\mathbf{u} = (f(\mathbf{e}_1),\ldots,f(\mathbf{e}_n))$. Then we see that $f(\mathbf{x}) = \mathbf{u} \cdot \mathbf{x}$. On the other hand, we know that taking dot product with a fixed vector is a linear map. Thus a linear map $\mathbb{R}^n \longrightarrow \mathbb{R}$ is nothing but taking dot product with a vector.
(iii) Consider the problem of solving a system of m linear equations in n variables:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \cdots & & & \cdots & & & \cdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{4}$$

A solution $(x_1,\ldots,x_n)$ can be thought of as an element of the euclidean space $\mathbb{R}^n$. Then the set of all solutions which satisfy the j^{th} equation above is nothing but some hyper plane P_j in $\mathbb{R}^n$. Solving the system means nothing but finding the intersection of all these planes viz., $P_1 \cap \cdots \cap P_m$. On the other hand we also can think of the lhs of each of these equation as a map $T_i : \mathbb{R}^n \longrightarrow \mathbb{R}$. Then each T_i is easily seen to be a linear map. Together, these m linear functions define one linear function $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that $T = (T_1,\ldots,T_m)$. Solving the system of equations is now the same as determining points $\mathbf{x} \in \mathbb{R}^n$ such that $T(\mathbf{x}) = \mathbf{b}(= (b_1,\ldots,b_m))$.

These ideas can be put in a very concrete way in terms of matrices.

Lecture 3

2.3 Matrix Representation

We have been used to write elements of $\mathbb{R}^n$ in the form $(x_1, \ldots, x_n)$. This is certainly a good practice if all the scripts are written along horizontal lines. However, what would happen if we use Chinese or Japanese? We would be writing them in columns! Certainly this does not cause any change in the mathematical aspects of the objects that are under study. We shall borrow this idea from the Chinese and for the sake of notational conformity later on, we shall now itself agree to write any element of $\mathbb{R}^n$ as a column of real numbers

$$\begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{pmatrix}$$

As we shall see soon, this convention will bring better conformity in notation rather than confusion. However, we also see that this practice will also cost us a lot of space. So, as space saving device, let us also have another convention viz., $(x_1, \ldots, x_n)^T$ (read 'x_1, dotdotdot x_n transpose') to represent the above column. The net effect is that from now onwards we will write an element of $\mathbb{R}^n$ in the form $(x_1, \ldots, x_n)^T$ and call it a *column vector,* in addition to our earlier convention of referring to it as a *row vector* as and when it suits us. Wait a moment to see why all this fuss.

Given a linear map $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ we have n column vectors (each of size m) viz., $f(\mathbf{e}_1), \ldots, f(\mathbf{e}_n)$. We place them side by side to obtain an array of real numbers arranged in m rows and n columns. For instance, if $f(\mathbf{e}_j) = (f_{1j}, f_{2j}, \ldots f_{mj})^T$, then the array we obtain looks like:

$$\mathcal{M}_f := \begin{bmatrix} f_{11} & f_{12} & \ldots & f_{1n} \\ f_{21} & f_{22} & \ldots & f_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{m1} & f_{m2} & \ldots & f_{mn} \end{bmatrix}$$

This array is called a *matrix with m rows and n columns* and with real *entries* $f_{ij} \in \mathbb{R}$ in the $(i,j)^{th}$ place. We say the matrix $\mathcal{M}_f$ is of *size* $m \times n$ or simply that it is a $m \times n$ matrix. We also use the simplified notation $((f_{ij}))$ for this matrix. Clearly, two matrices like this will be treated as equal if and only if their sizes are the same and the entries in the corresponding places are equal. In particular, if we have $m = 1$ we get *row matrices* which are the same as row vectors. Similarly, if $n = 1$ we get *column matrices* which are the same as column vectors.

From what we have seen earlier, it follows that the correspondence

$$f \mapsto \mathcal{M}_f$$

is one-one from $\mathbb{L}(n, m)$ onto $M_{m,n}$. This is called the *matrix representation* of linear maps. The correspondence $\mathcal{M} : \mathbb{L}(n, m) \longrightarrow M_{m,n}$ itself has the linearity property, viz.,

$$\mathcal{M}_{(f+g)} = \mathcal{M}_f + \mathcal{M}_g; \quad \mathcal{M}_{\alpha f} = \alpha \mathcal{M}_f.$$

Later on we shall see various properties and usefulness of this representation.

Example 4

Let us write down the matrix representation for some simple maps.

(i) Since the identity map takes $\mathbf{e}_i$ to $\mathbf{e}_i$ it follows that $\mathrm{M}_{Id} = I_n$ where $I = ((\delta_{ij}))$ is an $n \times n$ matrix with $\delta_{ij} = 1$ if $i = j$ and $= 0$ otherwise. (We shall denote it by I_n, the suffix n to indicate the size of the matrix if this is not clear from the context.)

(ii) Clearly then $\mathrm{M}_{-Id} = -I$.

(iii) The linear map $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ which interchanges the two coordinates is represented by

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{5}$$

(iv) Consider the scalar multiplication by $\alpha \in \mathbb{R}$ on $\mathbb{R}^n$. The matrix of this linear map is the diagonal matrix $D(\alpha, \cdots, \alpha)$.

(v) The rotation through an angle θ about the origin in the plane is represented by

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{6}$$

(First of all why it is a linear map? In order to see this, we must understand how the sum of two vectors in the plane is defined geometrically, viz., by the parallelogram. Now use the law of congruent triangles to see that under a rotation the entire parallelogram gets rotated. Hence the sum of the two vectors after rotating them is the same as rotating the sum of the two vectors. The same law for scaling is easier to see. Alternatively, first write down the image of $\mathbf{e}_1$ and $\mathbf{e}_2$ under the rotation, right down the matrix M and then prove that an arbitrary point $(x, y)^T$ is taken to $M(x, y)^T$ by the rotation to conclude that the rotation is after all a linear map.)

(vi) Let us right down the matrix representation for the reflection through a line L_ψ passing through the origin in the plane and making an angle ψ with the positive real axis, $0 \leq \psi < \pi$. Let us denote this by R_ψ. Clearly $\mathbf{e}_1$ is mapped to a unit vector which makes an angle 2ψ with the positive real axis whereas $\mathbf{e}_2$ is mapped to one which makes an angle $2\psi - \pi/2$. Hence the matrix of reflection R_ψ is

$$\begin{bmatrix} \cos 2\psi & \sin 2\psi \\ \sin 2\psi & -\cos 2\psi \end{bmatrix}. \tag{7}$$

Thus you can see that reflection through the line L_ψ is the composite of rotation through an angle 2ψ followed by reflection in the $x-$axis. If we denote the rotation through an angle θ by ρ_θ then we have

$$R_\psi = \rho_{2\psi} \circ R_0 \tag{8}$$