

Vectors and Matrices

ADAM KELLY

Michaelmas 2020, Updated October 15, 2020

None of the notes here have been reviewed at all, and are just exactly what was taken down live in the lectures. I would turn around now and come back in a few days, when I have gone back, cleaned things up, fixed explanations and added some structure.

This set of notes is a work-in-progress account of the course ‘Vectors and Matrices’, originally lectured by Dr. Jonathan Evans in Michaelmas 2020 at Cambridge. These notes are not a transcription of the lectures, but they do roughly follow what was lectured (in content and in structure).

These notes are my own view of what was taught, and should be somewhat of a superset of what was actually taught. I frequently provide different explanations, proofs, examples, and so on in areas where I feel they are helpful. Because of this, this work is likely to contain errors, which you may assume are my own. If you spot any or have any other feedback, I can be contacted at ak2316@cam.ac.uk.

Contents

0 Introduction	2
0.1 Course Structure	2
1 Complex Numbers	3
1.1 Definitions	3
1.2 Basic Properties of Complex Numbers	4
1.3 Exponential and Trigonometric Functions	5
1.4 Roots of Unity	6
1.5 Logarithms and Complex Powers	7
1.6 Transformations, Lines and Circles	8
2 Vectors in Three Dimensions	10
2.1 Vector Addition and Scalar Multiplication	10
2.2 The Dot Product	12

§0 Introduction

Vectors and Matrices covers topics in both algebra and geometry, and the way in which they relate to one another. The course uses approaches that are quite varied in their nature (can be abstract or more concrete, conceptual or more computational, and so on). You will need to be able to fluently switch between these approaches.

The course assumes that you are vaguely familiar with Euclidean and coordinate geometry, along with the idea of geometric transformations.

§0.1 Course Structure

This course is divided into a number of chapters.

1. *Complex Numbers*

This chapter takes the point of view of thinking of points in the plane as pairs of real numbers, and defining ‘multiplication’ on it to turn it into the complex numbers.

2. *Vectors in Three Dimensions*

Here, we will recap on the relationship between three dimensional vectors and some of their geometrical applications, and we will discuss things like the dot and cross product. Towards the end of that discussion, we will introduce the ‘index notation’, a powerful and helpful notation for dealing with vectors. We will also introduce the ‘summation convention’, which is also incredibly useful.

3. *Vectors in a General Setting*

This chapter will discuss what vectors are in general, and different ways of looking at them. We will be particularly concerned with vectors in \mathbb{R}^n and \mathbb{C}^n , that is, vectors whose entries are in \mathbb{R} and \mathbb{C} respectively.

4. *Matrices and Linear Maps*

Picking up on the idea of generalizing vectors, this chapter will consider the idea of a ‘linear map’, an abstraction of matrices.

5. *Determinants and Inverses*

This chapter will detail how to define and compute determinants of general $n \times n$ matrices. This will take two points of view, in that we need to be able to compute them but we also must understand what they mean. The relation between determinants and finding inverses of matrices will also be considered.

6. *Eigenvalues and Eigenvectors*

This chapter also involves both geometry and algebra. The core question of this chapter is: given a linear map or matrix, what does it act on in a very straightforward way?

7. *Changing Basis, Canonical Forms and Symmetries*

In this final chapter, we will consider a set of far reaching results by trying to describing an arbitrary linear map. These ideas are far reaching, and immensely useful.

§1 Complex Numbers

§1.1 Definitions

We construct the complex numbers \mathbb{C} from the real numbers \mathbb{R} by adjoining an element i with $i^2 = -1$. Any complex number $z \in \mathbb{C}$ has the form (by definition)

$$z = x + iy,$$

with $x, y \in \mathbb{R}$. $x = \operatorname{Re}(z)$ is called the **real part**, and $y = \operatorname{Im}(z)$ is called the **imaginary part**.

With this definition of complex numbers, we can define how arithmetic works with them.

Definition 1.1 (Arithmetic on Complex Numbers)

For $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, we define

- (i) *Addition.* $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$.
- (ii) *Multiplication.* $z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$.
- (iii) *Complex conjugation.* $\bar{z} = z^* = x - iy$.
- (iv) *Modulus.* $r = |z|$ is a non-negative real given by $r^2 = |z|^2 = z\bar{z} = x^2 + y^2$.
- (v) *Argument.* For nonzero z , $\theta = \arg(z)$ is a real defined^a such that $z = r(\cos \theta + i \sin \theta)$. This is the **polar form** of z .

With this we have that

$$\cos \theta = \frac{x}{(x^2 + y^2)^{1/2}} \quad \sin \theta = \frac{y}{(x^2 + y^2)^{1/2}} \quad \implies \quad \tan \theta = \frac{y}{x}.$$

In order for θ to be unique, we can restrict the range of θ . For example, if $-\pi < \theta \leq \pi$, we call this the **principle value**. However other choices of range are equally valid.

^a $\arg(z)$ is determined modulo 2π , which follows from $\tan \theta = x/y$.

Proposition 1.2 (Properties of Complex Conjugation)

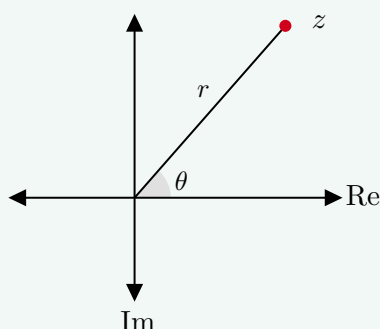
We have

- (i) $\operatorname{Re}(z) = \frac{1}{2}(z + \bar{z})$.
- (ii) $\operatorname{Im}(z) = \frac{1}{2i}(z - \bar{z})$.
- (iii) $\overline{\bar{z}} = z$.
- (iv) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$.
- (v) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

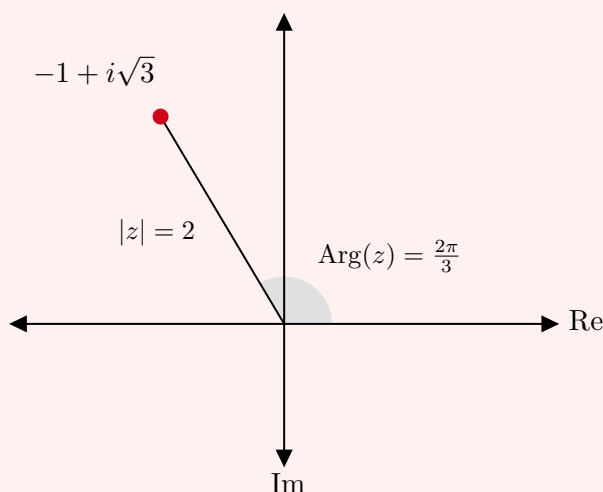
Proof Sketch. A direct check from the definitions. □

Definition 1.3 (The Argand Diagram/Complex Plane)

Plot $\operatorname{Re}(z)$ and $\operatorname{Im}(z)$ on orthogonal axes, then $r = |z|$ and $\arg(z) = \theta$ are the length and angle as shown.

**Example 1.4**

For $z = -1 + i\sqrt{3}$, we have $|z| = 2$ and $\arg(z) = \frac{2\pi}{3} + 2n\pi$ for $n \in \mathbb{Z}$. This is shown on the complex plane below.



Note that $\tan \theta = -\sqrt{3}$, so $\theta = \frac{2\pi}{3} + 2n\pi$, or $\theta = -\frac{\pi}{3} + 2n\pi$, which is $\arg(-z)$. This is notable because it is common to write $\arg(z) = \arctan(y/x)$, but that is not quite correct.

§1.2 Basic Properties of Complex Numbers

In this section, we will consider the consequences of the definitions given in the previous section. Let us begin by defining the notion of a field.

Definition 1.5

A **field** is a thing yeah.

Proposition 1.6

\mathbb{C} with addition and multiplication is a field.

Proof. Fill in details here. □

The next consequence is the ‘Fundamental Theorem of Algebra’. This is something that’s important to know about, but we will not prove it.

Theorem 1.7 (Fundamental Theorem of Algebra)

A polynomial $p(z)$ of degree n with coefficients in \mathbb{C} can be written as a product of n linear factors. That is, if we have a polynomial

$$p(z) = c_n z^n + c_{n-1} z^{n-1} + \cdots + c_1 z + c_0$$

with $c_i \in \mathbb{C}$ and $c_n \neq 0$, then we can write

$$p(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n),$$

for some $\alpha_i \in \mathbb{C}$ not necessarily distinct.

Proof Omitted. See the IB course ‘Complex Methods/Analysis’. □

§1.3 Exponential and Trigonometric Functions

We define the functions \exp , \cos and \sin on the complex numbers by

$$\begin{aligned} \exp(z) &= e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n \\ \cos(z) &= \frac{1}{2}(e^{iz} + e^{-iz}) = 1 - \frac{1}{2!} z^2 + \frac{1}{4!} z^4 - \cdots \\ \sin(z) &= \frac{1}{2i}(e^{iz} - e^{-iz}) = z - \frac{1}{3!} z^3 + \frac{1}{5!} z^5 - \cdots \end{aligned}$$

These series converge for all $z \in \mathbb{C}$ (this is discussed rigorously in the course ‘Analysis I’).

Proposition 1.8

For all $z, w \in \mathbb{C}$, we have $e^z e^w = e^{z+w}$.

Proof. Multiplying the series on the left, we consider the terms of degree n :

$$\sum_{r=0}^n \frac{1}{r!} z^r \frac{1}{(n-r)!} w^{n-r} = \frac{1}{n!} (z+w)^n,$$

by the Binomial Theorem, and the result follows. □

Corollary 1.9

For any $z \in \mathbb{C}$, $e^z e^{-z} = e^0 = 1$, and $(e^z)^n = e^{nz}$, for $n \in \mathbb{Z}$.

When $z \in \mathbb{R}$, we note that these definitions reduce to those that we are familiar with. Differentiating the series we have

$$\begin{aligned}\frac{d}{dx}(\exp x) &= \exp x & \exp(0) &= 1 \\ \frac{d}{dx}(\cos x) &= -\sin x & \cos(0) &= 1 \\ \frac{d}{dx}(\sin x) &= \cos x & \sin(0) &= 0\end{aligned}$$

Indeed these properties characterize these functions of for a real z (in the sense that they are the unique solutions to the differential equations). Of course, these series definitions have a different behavior for when $z \in \mathbb{C}$.

Lemma 1.10 (Properties of exp)

For $z = x + iy$:

- (i) $e^z = e^x(\cos y + i \sin y)$.
- (ii) \exp on \mathbb{C} takes on all non-zero complex values.
- (iii) $e^z = 1$ occurs if and only if $z = 2n\pi i$ for $n \in \mathbb{Z}$.

Proof.

- (i) $e^{x+iy} = e^x e^{iy}$, and $e^{iy} = \cos y + i \sin y$ by De Moivre's theorem. The result follows.
- (ii) $|e^z| = e^x$ which takes on all real, positive values, and $\arg(e^z) = y$, which also takes on all real values.
- (iii) Using the result from (i), $e^z = 1$ occurs if and only if $e^x = 1$, $\cos y = 1$ and $\sin y = 0$, which implies that $x = 0$ and $y = 2n\pi$, as required.

Note that using Lemma 1.10, we obtain a new way to express a complex number in polar form by writing

$$z = r(\cos \theta + i \sin \theta) = re^{i\theta},$$

where $r = |z|$ and $\theta = \arg(z)$. We also find that De Moivre's theorem now follows directly.

§1.4 Roots of Unity

A root of unity can be thought of as an answer to the question of 'what is $1^{\frac{1}{n}}$ '.

Definition 1.11 (Roots of Unity)

We say that $z \in \mathbb{C}$ is an **n th root of unity** if $z^n = 1$.

To find all solutions to the equation $z^n = 1$, we write $z = re^{i\theta}$, then

$$\begin{aligned}z^n &= r^n e^{in\theta} = 1 \\ \implies r^n &= 1 \quad \text{and} \quad \theta = \frac{2k\pi}{n}\end{aligned}$$

for $k \in \mathbb{Z}$. This gives n distinct solutions,

$$z = e^{2k\pi i/n} = \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} = \omega^k,$$

where $\omega = e^{2\pi i/n}$. These solutions lie on the unit circle at the vertices of a regular n -gon.

§1.5 Logarithms and Complex Powers

We define the logarithm for complex numbers as follows.

Definition 1.12 (Complex Logarithm)

We define the complex logarithm $w = \log z$ such that $e^w = e^{\log z} = z$, for non-negative $z \in \mathbb{C}$.

This defines \log to be the inverse of \exp , but \exp is many-to-one so \log is *multi-valued*.

Using this definition we have

$$z = re^{i\theta} = e^{\log r} e^{i\theta} = e^{\log r + i\theta},$$

so

$$\log z = \log r + i\theta = \log |z| + i \arg(z).$$

This implies that the multi-valued behavior of \arg and \log are related:

$$\theta \rightarrow \theta + 2n\pi \implies \log z \rightarrow \log z + 2n\pi i, \quad \text{for } n \in \mathbb{Z}.$$

To make both of these functions single valued, we need to restrict them in some way, for example $0 \leq \theta < 2\pi$ or $-\pi < \theta \leq \pi$ (which we previously called the principle value).

Similarly, we can define complex powers.

Definition 1.13

We define complex powers by $z^\alpha = e^{\alpha \log z}$, for $z \in \mathbb{C}$, $z \neq 0$ and $\alpha \in \mathbb{C}$.

This is also multi-valued since

$$\arg z \rightarrow \arg z + 2n\pi \implies z^\alpha \rightarrow z^\alpha e^{2\pi i n \alpha}.$$

Again we can restrict \arg to make this single valued, and using $-\pi < \theta \leq \pi$ gives the principal value of z^α . This multi-valuedness does not occur in every case. Importantly, we have the two following cases (which have no or a finite amount of ambiguity).

- If $\alpha = p \in \mathbb{Z}$, then $z^\alpha = z^p$ is unique.
- If $\alpha = \frac{p}{q} \in \mathbb{Q}$, then $z^\alpha = z^{p/q}$ takes on finitely many values.

In general though, we do have infinitely many values.

Example 1.14

To compute i^i we write $i = e^{i\pi/2}$. We have $\arg(i) = \frac{\pi}{2} + 2n\pi$ and $\log(i) = i(\frac{\pi}{2} + 2n\pi)$.

Then

$$i^i = e^{i \log i} = e^{-\left(\frac{\pi}{2} + 2n\pi\right)},$$

for $n \in \mathbb{Z}$.

Remark. We can immediately extend the standard quadratic formula to quadratic equations with complex coefficients.

Remark. For $x \in \mathbb{R}$ and $x > 0$ we have a *unique* real value of $\log x$ and a unique $x^\alpha = e^{\alpha \log x}$ for any $\alpha \in \mathbb{R}$. Then $x^\alpha x^\beta = x^{\alpha+\beta}$ and $(x^\alpha)^\beta = x^{\alpha\beta}$.

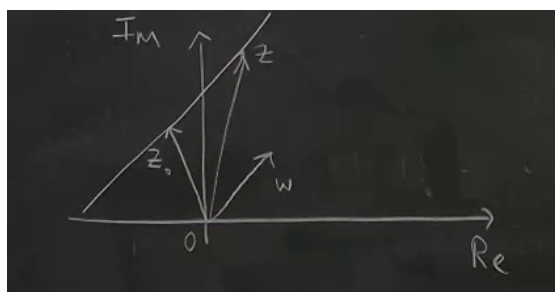
For complex variables however, we must beware as $z^\alpha z^\beta$ and $z^{\alpha+\beta}$ each have infinitely many possible values, but each has sets of values which may not match.

§1.6 Transformations, Lines and Circles

Consider the following transformations on the complex numbers

$z \mapsto z + a$	Translation by $a \in \mathbb{C}$
$z \mapsto \lambda z$	Scaling by $\lambda \in \mathbb{R}$
$z \mapsto e^{i\alpha} z$	Rotation by $\alpha \in \mathbb{R}$
$z \mapsto \bar{z}$	Reflection in the real axis
$z \mapsto \frac{1}{z}$	Inversion

We can use these ideas to find the general point on a line in \mathbb{C} through some $z_0 \in \mathbb{C}$ and parallel to $w \in \mathbb{C}$, with $w \neq 0$.



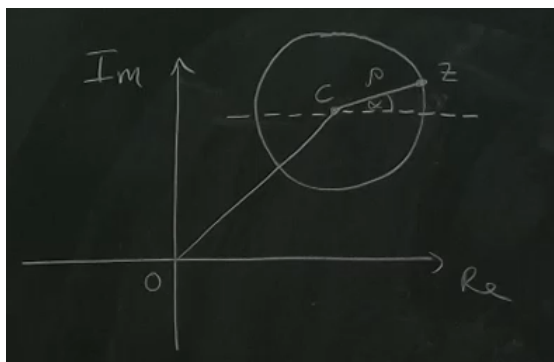
This is $z = z_0 + \lambda w$ for any $\lambda \in \mathbb{R}$. To eliminate the parameter λ , take the conjugate

$$\bar{z} = \bar{z}_0 + \lambda \bar{w}$$

and combine the equations to get

$$\bar{w}z - w\bar{z} = \bar{w}z_0 - w\bar{z}_0$$

Next, consider the circle with center $c \in \mathbb{C}$ and radius ρ .



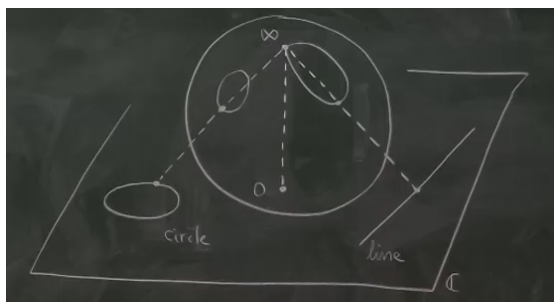
A general point on a circle is $z = c + \rho e^{i\alpha}$ for any α . Equivalently $|z - c| = \rho$. We can even expand this to get

$$|z|^2 - \bar{c}z - c\bar{z} = \rho^2 - |c|^2.$$

Aside: Inversion with Complex Numbers

We can use *stereographic projection* to understand how inversion works with complex numbers.

Identify the set $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$ with the Riemann Sphere as shown.



Imagine this sphere sitting on the plane at the origin. If we consider points on the sphere, we have a correspondence with points on the plane, created by drawing a line through the north pole (labelled ∞) on the sphere through a point on the sphere until it hits the plane (and vice versa).

Now when we look at a circle on the complex plane, it will correspond to a circle on the sphere. But also, a line in the plane also corresponds with a circle on the sphere, which goes through the north pole (∞) on the sphere.

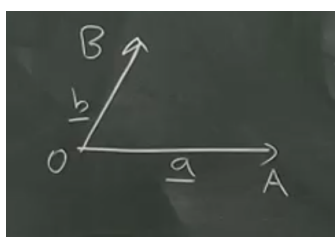
§2 Vectors in Three Dimensions

What is a vector? The traditional answer to this is that it's a quantity with both magnitude (length, size) and direction. There are many examples such quantities: force, momentum, electric and magnetic fields – but the vector nature of these quantities is essentially the same as the vector notion of position.

In this section, we will take a geometrical approach to vectors, and we will consider position vectors in 3d space, using standard (Euclidean) notions of points, lines, planes, lengths and angles.

Chose some point O to be the **origin**, then points A, B, \dots are represented as displacements from O , that is, position vectors, that we write

$$\underline{a} = \overrightarrow{OB}, \quad \underline{b} = \overrightarrow{OA}, \dots$$



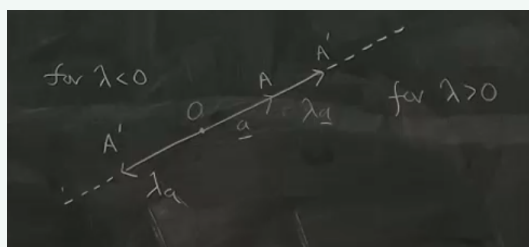
The lengths of these vectors is denoted $|\underline{a}| = |\overrightarrow{OA}|$, and we will use $\underline{0}$ for the position vector of the origin O .

§2.1 Vector Addition and Scalar Multiplication

We now need to establish what we can do with these vectors: adding them and multiplying them by scalars.

Definition 2.1 (Scalar Multiplication)

Given \underline{a} , a position vector for some point A , and a scalar $\lambda \in \mathbb{R}$, we define $\lambda \underline{a}$ to be the position vector of a new point A' on the line OA , with $|\lambda \underline{a}| = |\lambda| |\underline{a}|$ and the direction as shown (depending on the sign of λ).



Using scalar multiplication, we can develop a sensible notion of when two vectors have the same direction, that is, when they are parallel.

Definition 2.2 (Parallelism)

Define \underline{a} and \underline{b} to be **parallel**, written $\underline{a} \parallel \underline{b}$, if and only if

$$\underline{a} = \lambda \underline{b} \quad \text{or} \quad \underline{b} = \lambda \underline{a},$$

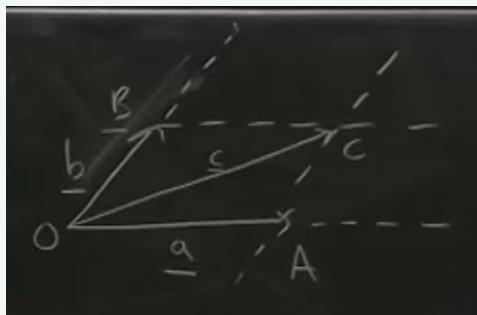
for some $\lambda \in \mathbb{R}$.

In this definition, we allow $\lambda = 0$, so $\underline{a} \parallel \underline{0}$ for any \underline{a} . This is slightly non-standard, but is convenient in this course.

We can also define the addition of vectors.

Definition 2.3 (Vector Addition)

Given $\underline{a}, \underline{b}$, the position vectors of points A, B , if $\underline{a} \not\parallel \underline{b}$, construct the parallelogram $OACB$. Then define $\underline{a} + \underline{b} = \underline{c}$, the position vector of point C as shown.



If $\underline{a} \parallel \underline{b}$, then we can write $\underline{a} = \alpha \underline{u}$ and $\underline{b} = \beta \underline{u}$ where $\alpha, \beta \in \mathbb{R}$ and $|\underline{u}| = 1$. Then $\underline{a} + \underline{b} = (\alpha + \beta)\underline{u}$.

A core idea to vectors is that given any number of them, $\underline{a}, \underline{b}, \dots, \underline{c}$, we can form a **linear combination**

$$\alpha \underline{a} + \beta \underline{b} + \dots + \gamma \underline{c},$$

for any $\alpha, \beta, \dots, \gamma \in \mathbb{R}$.

We can then define the *span* of a linear combination of vectors to be the set of points that can be reached by changing the coefficients.

Definition 2.4 (Span of Multiple Vectors)

We define the **span** of a set of vectors to be $\text{span}\{\underline{a}, \underline{b}\} = \{\alpha \underline{a} + \beta \underline{b} : \alpha, \beta \in \mathbb{R}\}$.

If $\underline{a} \not\parallel \underline{b}$, then this is the plane through points O, A, B .

We will finish our discussion of vector addition and scalar multiplication by looking at some of the elementary properties that are implied by our definitions.

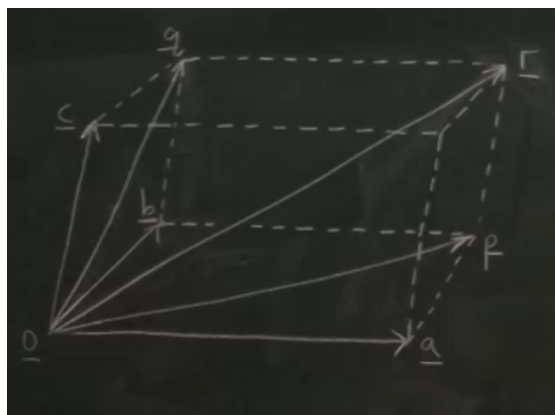
Theorem 2.5

Vectors, along with vector addition, form an abelian group.

Proof. For any vectors $\underline{a}, \underline{b}$ and \underline{c} ,

- *Identity.* We have the identity element $\underline{0}$, as $\underline{a} + \underline{0} = \underline{0} + \underline{a} = \underline{a}$.
- *Inverses.* There exists $-\underline{a}$ such that $\underline{a} + (-\underline{a}) = (-\underline{a}) + \underline{a} = \underline{0}$.
- *Associativity.* $\underline{a} + (\underline{b} + \underline{c}) = (\underline{a} + \underline{b}) + \underline{c}$.
- *Commutativity.* $\underline{a} + \underline{b} = \underline{b} + \underline{a}$. □

Remark. Given that we defined vector addition geometrically, it is not immediately clear that the associativity holds. Indeed, we really would need to consider the three dimensional construction of a parallelepiped, and check that the construction is indeed associative.

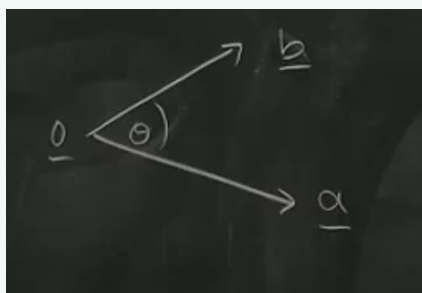


§2.2 The Dot Product

This subsection will involve length and angle in a way that our previous discussion didn't. Although we talked about length, it had more to do with proportionality, and we only touched on angles in regards to vectors being parallel (which is a rather special case).

Definition 2.6 (Dot Product)

Given vectors \underline{a} and \underline{b} , let θ be the angle between them as shown.



We define $\underline{a} \cdot \underline{b} = |\underline{a}||\underline{b}| \cos \theta$ to be the **dot product** of \underline{a} and \underline{b} . If either of $|\underline{a}| = 0$ or $|\underline{b}| = 0$, then $\underline{a} \cdot \underline{b} = 0$.

We can use the dot product to define when two vectors are perpendicular.

Definition 2.7 (Perpendicularity)

We say vectors \underline{a} and \underline{b} are **perpendicular** or **orthogonal**, written $\underline{a} \perp \underline{b}$, if and only if

$$\underline{a} \cdot \underline{b} = 0,$$

which occurs if and only if $\theta = \pm\pi/2$, when θ is defined.

As in the definition of parallelism for vectors, we allow \underline{a} or \underline{b} to be $\underline{0}$. This gives us that $\underline{0} \perp \underline{a}$ for any vector \underline{a} .

There is a geometric interpretation of the dot product. For $\underline{a} \neq \underline{0}$, $|\underline{b}| \cos \theta$ is the component of \underline{b} along \underline{a} .