A PRIMER FOR LINERA ALGEBRA

PROVIDED

BY

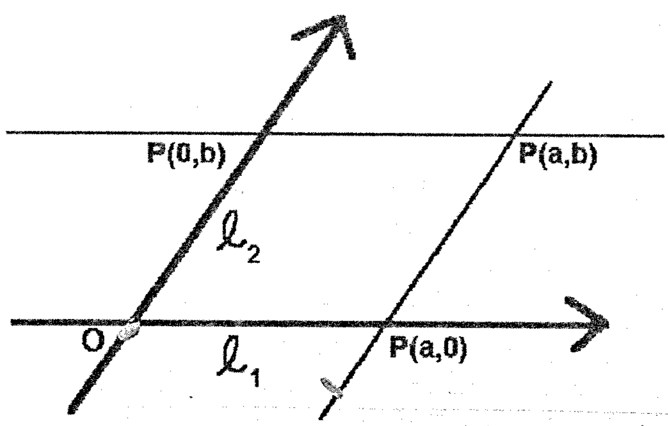
MATHLAB OF UH

These notes are meant as a "primer" for a first course on Linear Algebra. The notes deal only with points and vectors as spatial objects and no attempts have been made to present the theory in an axiomatic setting. On the other hand, we give a precise definition of what a vector is. Also, the section on the dot product is quite complete. I feel that the notes contain most of the the essentials on the algebra of vectors, which students of Calculus III (the multivariable calculus) or of a beginning physics course should be familiar with.

Klaus Kaiser

1. Points. Let ℓ be a straight line. We pick any point O on ℓ and call it the *origin.* Let U be any other point on ℓ. One customarily assumes that U is to the right of O. But this is not necessary. The half-line which contains U is called the *positive* half-line. The other half is called *negative.* We call U the *unit point* of ℓ. The segment $\overline{OU}$ determines the *unit length* for measuring distances. Let P be any point on the positive half-line. The length of $\overline{OP}$ measured in multiples of $\overline{OU}$ is called the *x-coordinate* of P. For example, U has the coordinate x = 1 while the origin O has the coordinate x = 0. The midpoint of the segment $\overline{OU}$ has coordinate x = 1/2. Points to the left of O have negative coordinates. We assume, as an axiom, that for every real number x there is exactly one point P on ℓ whose coordinate is x. The ordered pair (O,U) defines a *coordinate system* of the line ℓ.

Let ℓ_1 and ℓ_2 be two non-parallel lines in a plane π. Let O be the point of intersection. We pick any two points U_1 and U_2 as unit points of ℓ_1 and ℓ_2, respectively. The triple (O, U_1, U_2) determines an *(affine) coordinate system* for π. Let P be any point of the plane π. The line which is parallel to ℓ_1 and which goes through P intersects ℓ_2 in a point P_2. Similarly, the line which is parallel to ℓ_2 and which goes through P intersects ℓ_1 in a point P_1. Then, if a is the coordinate of P_1, and if b is the coordinate of P_2, one says that a and b are the coordinates of P = P(a,b) and that P_1 = P(0,b) and P_2 = P(a,0) are the *projections* of P along the *coordinate axes.*



A coordinate system is called *cartesian* if

(a) the coordinate axes are perpendicular to each other;

(b) the unit points have the same distance from the origin.

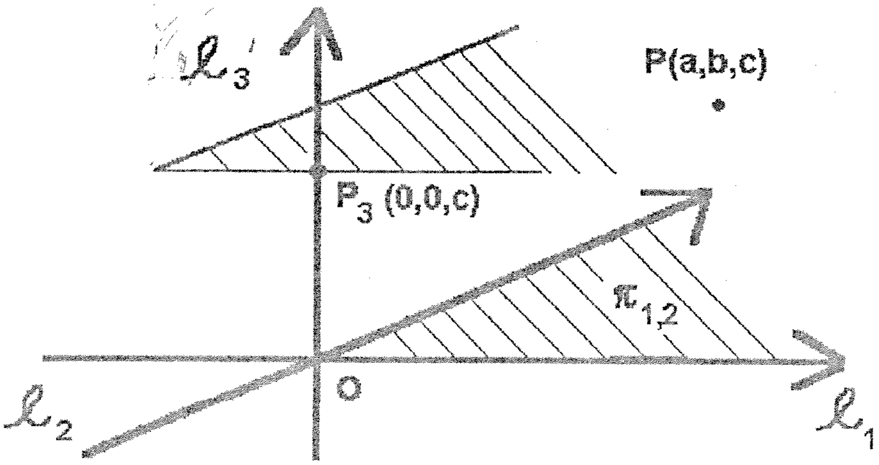
At any rate, whether the system (O, U_1, U_2) is affine or cartesian, there is a one-to-one correspondence between ordered pairs (a,b) of real numbers and points P of the plane. For example, the ordered pairs (0,0), (1,0) and (0,1) correspond to the points O, U_1 and U_2, respectively. Given (a,b) one has a unique point P_1 on ℓ_1 whose coordinate is a. Similarly, one has a unique point P_2 on ℓ_2 whose coordinate is b. All points where the first coordinate is a are on the line which goes through P_1 and which is parallel to ℓ_2. Similarly, all points with second coordinate b are on the line which goes through P_2 and which is parallel to ℓ_1. The point of intersection of these two lines is a unique point $P = P(a,b)$ with coordinates a and b.

Let O be any point of our physical space. We take three lines ℓ_i through O which are not in a plane. On each of the lines ℓ_i we pick a point U_i different from O. Then (O, U_1, U_2, U_3) is a space coordinate system. The origin O and any two of the unit points determine a *coordinate plane*. For example, O and U_1 and U_2 determine the plane $\pi_{1,2}$. We have three different coordinate planes: $\pi_{1,2}$, $\pi_{1,3}$ and $\pi_{2,3}$. Any two planes in space intersect in a line. For example, $\pi_{1,2}$ and $\pi_{1,3}$ have the line ℓ_1 in common. Note:

$$\pi_{1,2} \cap \pi_{1,3} = \ell_1 \ , \ \pi_{1,2} \cap \pi_{2,3} = \ell_2 \ , \ \pi_{1,3} \cap \pi_{2,3} = \ell_3$$

and

$$\pi_{1,2} \cap \pi_{1,3} \cap \pi_{2,3} = \{O\}$$



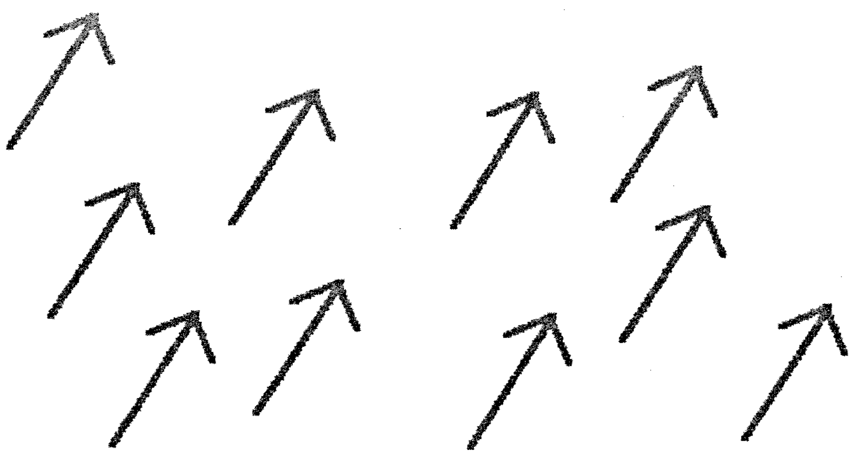
Let P be any point in space. Notice that the intersection of a line ℓ and a plane π is either empty, a point or the line ℓ. The plane which is parallel to $\pi_{1,2}$ and which goes through P intersects ℓ_3 in exactly one point P_3. Let x_3 be the coordinate of P_3 with respect to (O, U_3). The x_2-coordinate of P is determined by P_2 where P_2 is the intersection of the line ℓ_2 and the plane which goes through P and which is parallel to $\pi_{1,3}$. Finally, x_1 is the coordinate of P_1 where P_1 is the intersection of the line ℓ_1 and the plane which goes through P and which is parallel to $\pi_{2,3}$. We established a one-to-one correspondence between points P in space and triples (x_1,x_2,x_3). The origin O corresponds to (0,0,0), the point U_1 to (1,0,0), U_2 to (0,1,0) and U_3 to (0,0,1). All points whose first coordinate is x_1 are the points on the plane π_1 which goes through P_1 and which is parallel to $\pi_{2,3}$. The points for which x_2 is the second coordinate are the points on the plane π_2 which goes through P_2 and which is parallel to $\pi_{1,3}$. The intersection of the two planes π_1 and π_2 is a line ℓ. It contains all the points P for which x_1 and x_2 are the first two coordinates. If we intersect ℓ with the plane π_3 which is the plane which goes through P_3 and which is parallel to $\pi_{1,2}$, we get a unique point $P(x_1,x_2,x_3)$ whose coordinates are x_1, x_2 and x_3.

Examples.

1. Assume that we are given coordinate systems for a line ℓ, a plane π and for the space. The equation $x_1 = 2$ then determines
(a) a point of ℓ ; (b) a line of π; (c) a plane in space.

2. The equations $x_1 = 2$, $x_2 = -3$ determine a unique point of the plane π. In space this system describes a line as intersection of two planes.

2. Vectors. Let P and Q be any two points. They may be points on a line, of a plane or points in space. The *directed* line segment from the *initial* point P to the *end point* Q is called the *located* vector $\overrightarrow{PQ}$. Two located vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ are *equivalent* if the line segments $\overline{PQ}$ and $\overline{RS}$ are of the same length, are parallel and are directed the same way. Every located vector is equivalent to one where the origin O of a coordinate system is the initial point. Such located vectors are called *position* vectors. More generally, given any point R and any located vector $\overrightarrow{PQ}$ there is exactly one located vector $\overrightarrow{RS}$ which is equivalent to $\overrightarrow{PQ}$.



All these located vectors stand for one vector α

Given located vectors $\overrightarrow{PQ}$ and $\overrightarrow{QS}$ the located vector $\overrightarrow{PS}$ is called the sum:

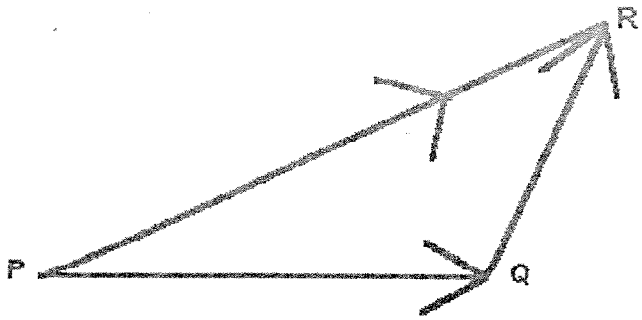
$$\overrightarrow{PQ} + \overrightarrow{QS} = \overrightarrow{PS}$$

This is called the *parallelogram* law for addition of located vectors. Notice that the end point of the first summand is the initial point of the second summand. We state a few properties for the addition of located vectors:

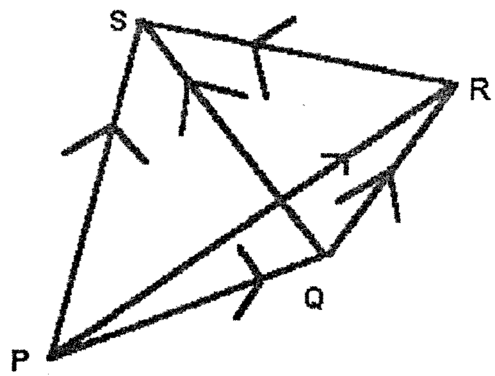
(a) $\overrightarrow{PQ} + \overrightarrow{QQ} = \overrightarrow{PQ}$,i.e., $\overrightarrow{QQ}$ is *neutral* with respect to addition.

(b) $\overrightarrow{PQ} + \overrightarrow{QP} = \overrightarrow{PP}$,i.e., $\overrightarrow{QP}$ is *inverse* to $\overrightarrow{PQ}$.

(c) $(\overrightarrow{PQ} + \overrightarrow{QR}) + \overrightarrow{RS} = \overrightarrow{PR} + \overrightarrow{RS} = \overrightarrow{PS}$, $\overrightarrow{PQ} + (\overrightarrow{QR} + \overrightarrow{RS}) = \overrightarrow{PQ} + \overrightarrow{QS} = \overrightarrow{PS}$. That is, the addition is *associative.*



$\overrightarrow{PR}$ is the sum of $\overrightarrow{PQ}$ and $\overrightarrow{QR}$



Associativity of addition

Any class α of equivalent located vectors is called a *vector*. If $\overrightarrow{PQ}$ is a located vector in space and if $P = P(a_1,a_2,a_3)$ and $Q = Q(b_1,b_2,b_3)$ then $\overrightarrow{PQ}$ is equivalent to the located vector $\overrightarrow{OX}$ where $X = (b_1 - a_1, b_2 - a_2, b_3 - a_3)$. One says that $c_i = b_i - a_i$ are the *components* of the vector α.

Given a coordinate system we may define operations for points according to the following rules:

"*The sum of points $P(a_1,...)$ and $P(b_1,...)$ is the point $P(a_1+b_1,...)$.*
$-P(a_1,...)$ is the point $P(-a_1,...)$.

With this convention we have can write:

$$\overrightarrow{PQ} \sim \overrightarrow{O(Q-P)} \, .$$

Notice that P + Q depends on the chosen coordinate system. If P is the origin for a coordinate system, then the sum is Q. If we choose Q as origin, the sum is P. If P is the origin then –P = P. in other words, the **sum of points has no absolute geometric meaning.** However, **Q – P determines a unique vector, the equivalence class of $\overrightarrow{PQ}$, regardless of the chosen coordinate system.**

A vector α stands for a whole equivalence class of located vectors. For any point P, e.g. P = O, there is exactly one located vector in α which has P as initial point and

$$\overrightarrow{PQ} \sim \overrightarrow{RS} \quad \text{iff} \quad Q - P = S - R$$

Using coordinates, one writes $\alpha = Q - P = (c_1,c_2,c_3)$.
Let α and β be vectors. We wish to define $\alpha + \beta$. Let $\overrightarrow{PQ} \in \alpha$ and $\overrightarrow{QS} \in \beta$. Then $\alpha + \beta$ is the class of $\overrightarrow{PS}$. It doesn't matter which located vector $\overrightarrow{PQ}$ one picks from α. It is easy to see that $\alpha + \beta = \beta + \alpha$.

Let $\overrightarrow{OX} \in \alpha$ and $\overrightarrow{OY} \in \beta$. We have $\overrightarrow{OY} \sim \overrightarrow{X(X+Y)}$. Hence:

$$\overrightarrow{OX} + \overrightarrow{X(X+Y)} = \overrightarrow{O(X+Y)} \text{ where } \overrightarrow{O(X+Y)} \in \alpha + \beta.$$

That is:

$$\text{If } \overrightarrow{OX} \in \alpha \text{ and } \overrightarrow{OY} \in \beta \text{ then } \overrightarrow{O(X+Y)} \in \alpha + \beta.$$

Using coordinates this says that the components c_i of $\alpha + \beta$ are $a_i + b_i$.
The class of $\overrightarrow{PP}$ is called the *zero vector* **o**. If $\overrightarrow{PQ} \in \alpha$ then the class of $\overrightarrow{QP}$ is called the additive inverse $(-\alpha)$ of α. Vectors form with respect to addition a **commutative group.** That is:

(a) $\alpha + o = \alpha$.
(b) $\alpha + (-\alpha) = o$.
(c) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
(d) $\alpha + \beta = \beta + \alpha$.

Let $c \in \mathbb{R}$ and $\overrightarrow{PQ}$ be a located vector. Assume $P \neq Q$. We define $c.\overrightarrow{PQ} = \overrightarrow{PR}$ where R is the point on the line which goes through P and Q and has with respect to the coordinate system (P,Q) the coordinate c. That is, $c.\overrightarrow{PQ}$ has initial point P, the same direction as $\overrightarrow{PQ}$ if $c > 0$, but pointing in the opposite direction for $c < 0$. The length of $c.\overrightarrow{PQ}$ is c-times the length of $\overrightarrow{PQ}$. We define $c.\overrightarrow{PP} = \overrightarrow{PP}$. Notice:

$$\text{If} \quad \overrightarrow{PQ} \sim \overrightarrow{RS} \quad \text{then} \quad c.\overrightarrow{PQ} \sim c.\overrightarrow{RS}\ .$$

In order to define $c.\alpha$, we may pick any $\overrightarrow{PQ} \in \alpha$ and define as $c.\alpha$ the class of $c.\overrightarrow{PQ}$.

We define, with respect to a given coordinate system, a multiplication of real numbers and points.

Let c be a real number and $P(a_1,\ldots)$ be a point. Then $c.P(a_1,\ldots)$ is the point $P(c\cdot a_1,\ldots)$.

Again, the result c.P depends on the chosen coordinate system. If P is the origin O the $c.P = O$ holds for every $c \in \mathbb{R}$.

Let $\overrightarrow{OX} \in \alpha$. Then $c.\overrightarrow{OX} = \overrightarrow{O(c.X)}$. This is quite easy to see. The following rules have very easy proofs. Together with rules (a)-(d) they constitute the *axioms of a vector space.*

(e) $c.(\alpha + \beta) = c.\alpha + c.\beta$.
(f) $(c + d).\alpha = c.\alpha + d.\alpha$.
(g) $(c\cdot d).\alpha = c.d.\alpha$.
(h) $1.\alpha = \alpha$.

Examples.
1. Find the X such that $\overrightarrow{OX} = \overrightarrow{(1,2,-2)(4,3,1)}$. Answer:
$X = (4,3,1) - (1,2,-2) = (3,1,3)$.
2. Find the X such that $\overrightarrow{(2,1,-5)X} = \overrightarrow{(1,2,-2)(4,3,1)}$. Answer:
$X - (2,1,-5) = (3,1,3)$, $X = (2,1,-5) + (3,1,3) = (5,2,-2)$.
3. Find the X such that $\overrightarrow{X(2,1,-5)} = \overrightarrow{(1,2,-2)(4,3,1)}$. Answer:

$(2,1,-5) - X = (3,1,3)$, $X = (2,1,-5) - (3,1,3) = (-1,0,-2)$.

Instead of $\overrightarrow{PQ} \in \alpha$ one often writes $\overrightarrow{PQ} = \alpha$. That is, one identifies a single representative, i.e., a located vector, with its equivalence class.

Let P be a point and α be a vector. One defines:

$$P + \alpha = Q, \quad \text{where } \overrightarrow{PQ} = \alpha$$

With respect to any coordinate system one has that the coordinates q_i of Q are given by $p_i + a_i$ where p_i are the coordinates of P and a_i are the components of α.

Let (O, U_1, U_2) be a coordinate system. The classes for the located vectors $\overrightarrow{OU_i}$ are called the *unit vectors* ε_i. If P is any point then

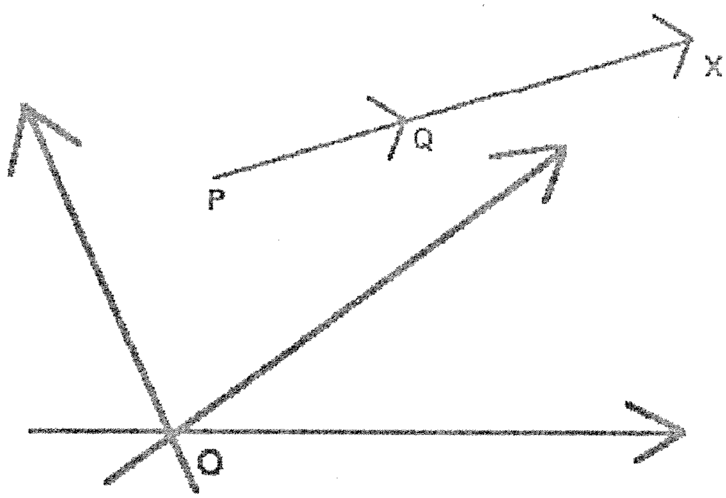
$$\overrightarrow{OP} = \overrightarrow{O(x_1,x_2,x_3)} = x_1.\varepsilon_1 + x_2.\varepsilon_2 + x_3.\varepsilon_3 = \overrightarrow{OP_1} + \overrightarrow{OP_2} + \overrightarrow{OP_3}$$

That is, every vector is the sum of its projections along the coordinate axes.

Let α and β be two non-zero vectors. We say that α is *parallel* to β if there is some $c \in \mathbb{R}$ such that $\alpha = c.\beta$. This defines obviously an equivalence relation among non-zero vectors.

3. Lines and Planes. Two different points P and Q determine a line. The set of points which are on the line through P and Q are the points X such that

$$\overrightarrow{PX} = t.\overrightarrow{PQ} \ , \ t \in \mathbb{R}.$$



$\overrightarrow{PX} = s.\overrightarrow{PQ}$

With respect to a coordinate system this reads as $X - P = t.(Q - P)$, i.e.,

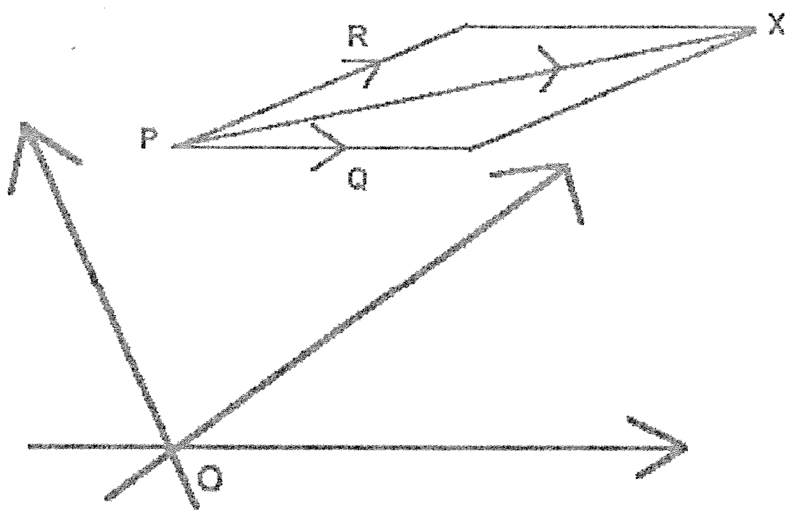
$$X = P + t.(Q - P), \ t \in \mathbb{R}.$$

Let P, Q and R be three different points where S is not on the line through P and Q. Then (P,Q,R) constitutes an affine coordinate system $\mathfrak{S}$ of a plane π where P plays the role of the origin and Q and R are the unit points. Let $\overrightarrow{PQ} = \alpha$ and $\overrightarrow{PR} = \beta$. That is, α and β play the roles of unit vectors for $\mathfrak{S}$. Let X be a point of π. Then

$$\overrightarrow{PX} = s.\alpha + t.\beta \ , \ s,t \in \mathbb{R}.$$

In coordinates, this is $X - P = s.(Q - P) + t.(R - P)$. Thus:

$$X = P + s.(Q - P) + t.(R - P) \ , \ s,t \in \mathbb{R}.$$



$\overrightarrow{PX} = s.\overrightarrow{PQ} + t.\overrightarrow{PR}$

Example. Show that the three medians of a triangle intersect in one point.
Answer: Let A,B,C be the vertices of a triangle. Let N, L and M be the midpoints of the sides a,b and c, respectively.

We first calculate the intersection S of the medians $\overline{AN}$ and $\overline{CM}$:

$M = A + \frac{1}{2}(B - A)$, $N = B + \frac{1}{2}(C - B)$. There are numbers s and t such that:

$$S = A + t.(N - A) = C + s.(M - C)$$

$$A + t.(B + \frac{1}{2}(C - B) - A) = C + s.(A + \frac{1}{2}(B - A) - C)$$

$$(1 - t - \frac{s}{2}).A - (1 - s - \frac{t}{2}).C - \frac{1}{2}.(s - t).B = 0$$

$$(1 - t - \frac{s}{2}).A - (1 - t - \frac{s}{2}).C + \frac{s}{2}C - \frac{t}{2}C - \frac{1}{2}.(s - t).B = 0$$

$$(1 - t - \frac{s}{2}).A - (1 - t - \frac{s}{2}).C = \frac{1}{2}.(s - t).B - \frac{1}{2}.(s - t).C$$

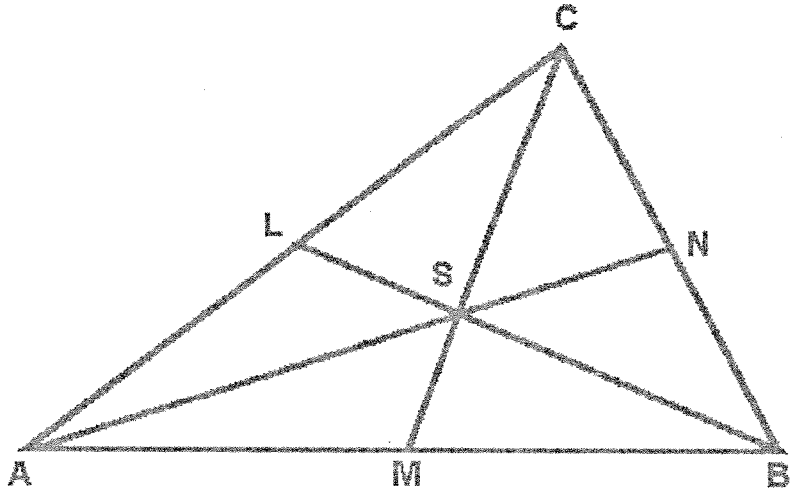
$$(1 - t - \frac{s}{2}).\overrightarrow{CA} = \frac{1}{2}.(s - t).\overrightarrow{CB}$$

But $\overrightarrow{CA}$ and $\overrightarrow{CB}$ are not parallel. Hence,

$$(1 - t - \frac{s}{2}) = 0, \frac{1}{2}.(s - t) = 0$$

This yields

$$s = t \text{ and } s = t = \frac{2}{3}.$$



$$S = A + \frac{2}{3}(N - A) = C + \frac{2}{3}(M - C)$$

We got

$$S = A + \frac{2}{3}\cdot(N - A) = C + \frac{2}{3}\cdot(M - C)$$

as point of intersection for $\overrightarrow{AN}$ and $\overrightarrow{CM}$. By symmetry, the intersection T of $\overrightarrow{AN}$ and $\overrightarrow{BL}$ is

$$T = A + \frac{2}{3}\cdot(N - A) = B + \frac{2}{3}\cdot(L - B)$$

Hence, S = T and all three medians intersect in one point.

4. **Scalar Product.** Let $\mathbb{C}$ be any cartesian coordinate system of a plane or the space. This gives us a fixed unit for measuring lengths and a frame for measuring angles. Let $\alpha = (a_1,a_2,a_3)$ and $\beta = (b_1,b_2,b_3)$. We define the *scalar product* by the number:

$$\alpha\cdot\beta = a_1\cdot b_1 + a_2\cdot b_2 + a_3\cdot b_3$$

The scalar product generalizes the product of numbers. Our main goal is to show that the scalar product, despite the fact that it has been defined in terms of components, does not depend on the chosen coordinate system $\mathbb{C}$. For any vector $\alpha = (a,b,c)$ one has

$$\alpha\cdot\alpha = a^2 + b^2 + c^2$$

That is,

$$\alpha\cdot\alpha \geq 0 \quad \text{and} \quad \alpha\cdot\alpha = 0 \text{ iff } \alpha = o$$

If $\overrightarrow{P_1P_2} \in \alpha$ where $P_1 = P_1(a_1,b_1,c_1)$ and $P_2 = P_2(a_2,b_2,c_2)$ then the components of α are $a = a_2 - a_1$, $b = b_2 - b_1$ and $c = c_2 - c_1$ and, according to the *Pythagorean Theorem,* one has for the length of the line segment

$$|P_1P_2| = \sqrt{a^2 + b^2 + c^2}$$

For any vector α we define the *norm* or *length* by

$$\|\alpha\| = \sqrt{\alpha\cdot\alpha}$$

The following rules comprise the basic properties for the scalar product. They are very easy to prove:

(a) $\alpha\cdot\beta = \beta\cdot\alpha$.

(b) $(\alpha + \beta)\cdot\gamma = \alpha\cdot\gamma + \alpha\cdot\gamma$.

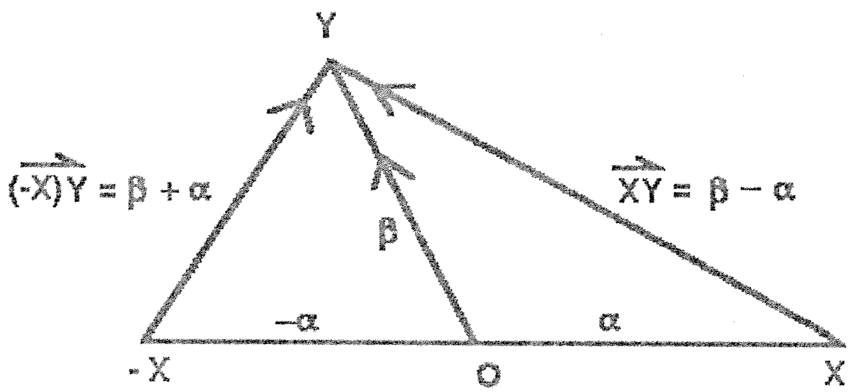
(c) $(c.\alpha)\cdot\beta = c\cdot(\alpha\cdot\beta)$.

We have for the unit vectors of our chosen cartesian coordinate system:

$$\varepsilon_1\cdot\varepsilon_j = \delta^i_j \;,\; \delta^i_j = 1, \text{ for } i = j, \text{ and } 0 \text{ otherwise.}$$

We are going to show that the scalar product of any two non-zero vectors is zero if and only if the vectors are perpendicular to each other. But how can we express in concise mathematical terms that α and β are perpendicular? Let $\overrightarrow{OX}$ represent α and $\overrightarrow{OY}$ represent β. Then $\overrightarrow{XY}$ represents $\beta - \alpha$ and $\overrightarrow{(-X)Y}$ represents $\beta + \alpha$. Clearly, $\overrightarrow{OX}$ is perpendicular ($\perp$) to $\overrightarrow{OY}$ iff $\overrightarrow{XY}$ and $\overrightarrow{(-X)Y}$ are of the same length. That is,

$$\alpha \perp \beta \text{ iff } \|\beta - \alpha\| = \|\beta + \alpha\|$$



But $\|\beta - \alpha\| = \|\beta + \alpha\|$ is the same as $\|\beta - \alpha\|^2 = \|\beta + \alpha\|^2$. This is, $(\beta - \alpha)\cdot(\beta - \alpha) = (\beta + \alpha)\cdot(\beta + \alpha) \Leftrightarrow \beta\cdot\beta - 2(\alpha\cdot\beta) + \beta\cdot\beta = \beta\cdot\beta + 2(\alpha\cdot\beta)\ \alpha\cdot\alpha \Leftrightarrow 4(\alpha\cdot\beta) = 0 \Leftrightarrow \alpha\cdot\beta = 0$. Hence,

$$\alpha \perp \beta \text{ iff } \alpha\cdot\beta = 0$$

Assume $\alpha\cdot\beta = 0$ and let $\gamma = \alpha + \beta$. An easy calculation establishes the

Pythagorean Theorem for vectors: $\|\gamma\|^2 = \|\alpha\|^2 + \|\beta\|^2$ if $\alpha \perp \beta$

If $\alpha \neq o$, then

$$\varepsilon_\alpha = \frac{1}{\|\alpha\|}\alpha$$

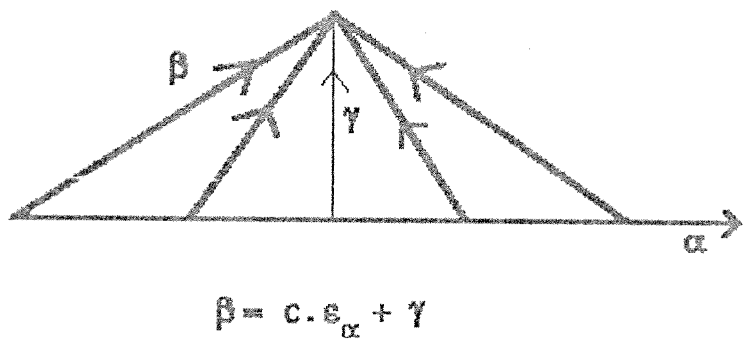
is called the unit vector in the direction of α. One has $\|\varepsilon_\alpha\| = 1$ and, of course, ε_α points in the direction of α. Unit vectors are also called *directions*. If α and β are non-zero vectors and if $c,d \neq 0$ then

$$\alpha \perp \beta \text{ iff } c.\alpha \perp d.\beta \text{ iff } \varepsilon_\alpha \perp \varepsilon_\beta$$

The zero vector o has no direction and is considered as perpendicular to any other vector. Many applications require the decomposition of a vector β (e.g. of force or velocity) into a direction given by a vector α and into a direction which is perpendicular to α.

Given the direction ε_α and a vector β we try to find some c such that

$$(*) \quad \beta = c.\varepsilon_\alpha + \gamma \text{ where } \gamma \perp \alpha$$



There is only one c such that $\gamma = \beta - c.\varepsilon_\alpha$ is perpendicular to α. γ is called the projection of β along α.

Multiplication of both sides of both sides of (*) by ε_α yields

$$c = \beta \cdot \varepsilon_\alpha.$$

The vector

$$\text{proj}_\alpha(\beta) = (\beta \cdot \varepsilon_\alpha).\varepsilon_\alpha = \frac{1}{\|\alpha\|^2} \cdot (\beta \cdot \alpha).\alpha$$

is called the *projection of β along* α. Its length is given by

$$\|\text{proj}_\alpha(\beta)\| = |\beta \cdot \varepsilon_\alpha| = \frac{1}{\|\alpha\|} \cdot |(\beta \cdot \alpha)|$$

The vector

$$\gamma = \beta - \text{proj}_\alpha(\beta)$$

is perpendicular to ε_α and

$$\beta = \text{proj}_\alpha(\beta) + (\beta - \text{proj}_\alpha(\beta))$$

is the decomposition of β into a component which points in the direction of α, and a component which is perpendicular to α. There is only one such decomposition. If δ is the angle between the vectors α and β then

$$\cos(\delta) = \frac{\frac{1}{\|\alpha\|} \cdot (\beta \cdot \alpha)}{\|\beta\|}$$

Hence:

$$(\beta \cdot \alpha) = \|\beta\| \cdot \|\alpha\| \cdot \cos(\delta)$$

The last formula shows that the value of the scalar product is independent of the chosen cartesian coordinate system. Because $|\cos(\delta)| \leq 1$, one has the

Cauchy -Schwarz inequality: $|(\beta \cdot \alpha)| \leq \|\beta\| \cdot \|\alpha\|$

There is a somewhat more elementary proof of Cauchy-Schwarz which doesn't rely on the cosine function. The orthogonal decomposition

$$\beta = \text{proj}_\alpha(\beta) + (\beta - \text{proj}_\alpha(\beta))$$

leads to:

$$\|\beta\|^2 = \|\text{proj}_\alpha(\beta)\|^2 + \|\beta - \text{proj}_\alpha(\beta)\|^2$$

Thus: $\|\beta\|^2 \geq \|\text{proj}_\alpha(\beta)\|^2$, i.e., $\|\beta\| \geq \|\text{proj}_\alpha(\beta)\| = \frac{1}{\|\alpha\|} \cdot |(\beta \cdot \alpha)|$

If we multiply the last inequality by $\|\alpha\|$, Cauchy-Schwarz follows.

We are going to show that the projection of β along α is the unique vector $c.\alpha$ for which the function $f(c) = \|\beta - c.\alpha\|$ takes on its minimum. We have the orthogonal decomposition:

$$\beta - c.\alpha = (\beta - \text{proj}_\alpha(\beta)) + (\text{proj}_\alpha(\beta) - c.\alpha)$$

Thus: $$\|\beta - c.\alpha\|^2 = \|\beta - \text{proj}_\alpha(\beta)\|^2 + \|\text{proj}_\alpha(\beta) - c.\alpha\|^2 .$$

Hence: $\|\beta - c.\alpha\| \geq \|\beta - \text{proj}_\alpha(\beta)\|$ and equality holds iff $\text{proj}_\alpha(\beta) = c.\alpha$.

This latter property of the projection of a vector β along $\alpha \neq o$ makes it clear that the projection of β along α depends only on the line $\langle\alpha\rangle = \{c.\varepsilon_\alpha | c \in \mathbb{R}\}$ which is generated by α. We are now going to find the projection of a vector β along a plane.

Let α and β be non-zero vectors and assume that they are not parallel. Then ε_α and $\gamma = \beta - \text{proj}_\alpha(\beta)$ are non-zero and perpendicular to each other. It is easy to see that the vectors $\varepsilon_\alpha = \varepsilon_1$ and $\varepsilon_\gamma = \varepsilon_2$ produce the same "span" as α and β. That is, they span the same plane π:

$$\pi = \langle\alpha,\beta\rangle = \{c.\alpha + d.\beta\} = \{c.\varepsilon_1 + d.\varepsilon_2\} \text{ where } c,d \in \mathbb{R}$$

Using projections to ortho-normalize the spanning vectors of a plane in order to obtain a cartesian system is called the Gram-Schmidt process.

If $\gamma \in \pi$ then $\gamma = c.\varepsilon_1 + d.\varepsilon_2$. In order to calculate the first component c of γ, we multiply both sides by ε_1. It follows $c = \gamma\cdot\varepsilon_1$ and, similarly, $d = \gamma\cdot\varepsilon_2$. Hence:

$$\gamma = (\gamma.\varepsilon_1) + (\gamma.\varepsilon_2)$$

Notice that $(\gamma\cdot\varepsilon_1).\varepsilon_1$ is the projection of γ along ε_1 and $(\gamma\cdot\varepsilon_2).\varepsilon_2$ is the projection of γ along ε_2.

Let γ be any vector, not necessarily in π. We define the projection of γ along π by

$$\text{proj}_\pi(\gamma) = (\gamma\cdot\varepsilon_1)\varepsilon_1 + (\gamma\cdot\varepsilon_2).\varepsilon_2$$

We need to show that $\text{proj}_\pi(\gamma)$ depends only on π, not on the chosen cartesian coordinate system. In order to prove this, let η be any vector in π. Then $\eta = c.\varepsilon_1 + d.\varepsilon_2$ and an easy calculation shows: $(\gamma - \text{proj}_\pi(\gamma))\cdot\eta = 0$. That is, $\gamma - \text{proj}_\pi(\gamma)$ is perpendicular to any vector η in π. If x is another vector in π, then

$$\gamma - x = (\gamma - \text{proj}_\pi(\gamma)) + (\text{proj}_\pi(\gamma) - x)$$

is an orthogonal decomposition: $\eta = \text{proj}_\pi(\gamma) - x$ is as the difference of two vectors in π also a vector in π. Hence, by the Pythagorean theorem:

$$\|\gamma - x\|^2 = \|\gamma - \text{proj}_\pi(\gamma)\|^2 + \|\text{proj}_\pi(\gamma) - x\|^2$$

Thus: $$\|\gamma - x\| \geq \|\gamma - \text{proj}_\pi(\gamma)\|$$

and equality holds iff $x = \text{proj}_\alpha(\gamma)$.

We have shown that $x = \text{proj}_\pi(\gamma)$ is the only vector in π such that $\gamma - x$ is perpendicular to all vectors in π. We also showed that it is characterized as the unique vector for which $f(x) = \|\gamma - x\|$, $x \in \pi$, takes on its minimum. If ε_1, ε_2 is any *ortho-normal* system of vectors in π, then the projection of γ onto π is the sum of the projections of γ along ε_1 and ε_2, respectively.

Let α, β and γ be three vectors which are not in a plane. That is, together with a point P as origin, they form a coordinate system different from $\mathbb{C}$. We can continue the Gram-Schmidt process in order to obtain a cartesian system where the first two vectors ε_1 and ε_2 span the same plane π as α and β. The vector $\varepsilon_3' = \gamma - \text{proj}_\pi(\gamma)$ is perpendicular to ε_1 and ε_2. If we make ε_3' to a unit vector ε_3, then ε_1, ε_2 and ε_3 form a cartesian base of our three dimensional space. Actual calculations are quite tedious and should be done with the help of mathematical software, but in the following example we may use projections in order to calculate explicitly the distance of a point from a plane. A more efficient method will be provided in the next section.

Example. Find the distance $d(P,\pi)$ of the point $P(1,1,1)$ from the plane π which goes through the points $P_0(1,0,0)$, $P_1(0,1,0)$ and $P_2(0,0,1)$.

Answer: $X = (1,0,0) + s.(-1,1,0) + t.(-1,0,1)$ is the parametric equation of π. Let $\alpha = (-1,1,0)$ and $\beta = (-1,0,1)$ and $\gamma = \overrightarrow{P_0P} = (0,1,1)$. The vectors α and β span the plane π_0 through O and $d(P,\pi)$ is $\|\gamma - \text{proj}_{\pi_0}(\gamma)\|$. In order to calculate $\text{proj}_{\pi_0}(\gamma)$ we have to replace α and β by an ortho normal system. We set $\varepsilon_1 = \varepsilon_\alpha$ and define ε_2 as the unit vector for $\beta - \text{proj}_{\varepsilon_1}(\beta)$. Thus:

$$\varepsilon_1 = \frac{1}{\sqrt{2}}.(-1,1,0),$$

$$\text{proj}_{\varepsilon_1}(-1,0,1) = \Big((-1,0,1)\cdot\frac{1}{\sqrt{2}}.(-1,1,0)\Big).\frac{1}{\sqrt{2}}.(-1,1,0) = \frac{1}{2}.(-1,1,0)$$

$$\beta - \text{proj}_{\varepsilon_1}(-1,0,1) = (-1,0,1) - \frac{1}{2}.(-1,1,0) = (-\frac{1}{2}, -\frac{1}{2}, 1),$$

$$\varepsilon_2 = \sqrt{\frac{2}{3}}.(-\frac{1}{2}, -\frac{1}{2}, 1),$$

$$\begin{aligned}\text{proj}_{\pi_0}(\gamma) &= \Big((0,1,1)\cdot\frac{1}{\sqrt{2}}.(-1,1,0)\Big).\frac{1}{\sqrt{2}}.(-1,1,0)\Big) + \\ &\quad \Big((0,1,1)\cdot\sqrt{\frac{2}{3}}.(-\frac{1}{2}, -\frac{1}{2}, 1)\Big).\sqrt{\frac{2}{3}}.(-\frac{1}{2}, -\frac{1}{2}, 1) \\ &= \frac{1}{2}.(-1,1,0) + \frac{1}{3}.(-\frac{1}{2}, -\frac{1}{2}, 1) = (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3})\end{aligned}$$

$$\gamma - \text{proj}_{\pi_0} = (0,1,1) - (-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}) = \frac{2}{3}.(1,1,1) = \frac{2\sqrt{3}}{3}.(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$$

$\nu = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ is called the *normal* vector for π_0. We have $\|\nu\| = 1$ and ν is perpendicular to π_0. We get an orthogonal decomposition:

$$\gamma = (\gamma - \text{proj}_{\pi_0}) + \text{proj}_{\pi_0}(\gamma)$$

Thus: $$(\gamma - \text{proj}_{\pi_0}(\gamma)) = \text{proj}_\nu(\gamma) = (\gamma\cdot\nu).\nu$$

and $\|\gamma - \text{proj}_{\pi_0}(\gamma)\| = \frac{2\sqrt{3}}{3}$ is the length of the projection of $\overrightarrow{P_0P}$ along the normal ν, i.e., the distance of P to the plane.

5. **The Hesse-Formula of a Plane.** Let π be the plane through P_0 and which has α and β as spanning vectors. That is, π is the set of points given by

$$X = P_0 + s.\alpha + t.\beta \quad \text{where } s,t \in \mathbb{R}$$

There is a different description of the same plane. Let ν be a vector which is perpendicular to all the vectors $\overrightarrow{P_0X}$, where X is a point of π. That is,

$$\nu \perp s.\alpha + t.\beta \quad \text{where } s,t \in \mathbb{R} \Leftrightarrow \nu \perp \overrightarrow{P_0X} \quad \text{where } X \in \pi$$

Let $\mathbb{C}$ be any cartesian coordinate system. Then $\alpha \perp \beta$ is the same as $\alpha \cdot \beta = 0$. If ν has with respect to $\mathbb{C}$ components a,b and c and if $P = P(x_0, y_0 z_0)$ then

$$\nu \perp \overrightarrow{P_0X} \Leftrightarrow \nu \cdot \overrightarrow{P_0X} = 0 \Leftrightarrow a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

The formula

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

is the equation of a plane through P_0 and which is perpendicular to the vector ν. The formula can be rewritten as

$$a \cdot x + b \cdot y + c \cdot z = a \cdot x_0 + b \cdot y_0 + c \cdot z_0 = d$$

That is, a plane is the set of all points X for which the scalar product with a vector ν is constant d. In particular,

$$a \cdot x + b \cdot y + c \cdot z = 0$$

is the equation of all points X for which $\overrightarrow{OX}$ is perpendicular to ν. In this case, X = O is a point of π.

Both sides of the equation $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ can be multiplied by any number different from zero. In particular, we may multiply by $1/\|\nu\|$. Then

$$\left(\frac{1}{\|\nu\|} . \nu\right) \cdot (X - X_0) = \varepsilon_\nu \cdot (X - X_0) = 0$$

is called the *Hesse-Equation* of the plane π. Its left-hand side is called the *Hesse-Formula.* If X is any point in space, not necessarily a point of π, then the projection of $(X - X_0) = \overrightarrow{P_0X}$ along ν is given by the formula

$$\text{proj}_\nu \overrightarrow{P_0X} = ((X - X_0)\cdot\varepsilon_\nu).\varepsilon_\nu$$

The length of the projection is $|(X - X_0)\cdot\varepsilon_\nu|$. Hence,

$$\varepsilon_\nu\cdot(X - X_0)$$

is $\pm$ the length of the projection of $\overrightarrow{P_0X}$ along the normal vector. But this number is also equal to the distance of the point X to the plane. The unit vector ε_ν is called the *normal* vector of the plane π.

Example. Find the distance $d(P,\pi)$ of the point $P(1,1,1)$ from the plane π which goes through the points $P_0(1,0,0)$, $P_1(0,1,0)$ and $P_2(0,0,1)$.

Answer. All three points satisfy the equation $x + y + z = 1$. This is,

$$(x - 1) + y + z = 0$$

The vector $\nu = (1,1,1)$ is the perpendicular to π and has length $\sqrt{3}$. Thus

$$\frac{(x - 1) + y + z}{\sqrt{3}}$$

is the Hesse formula for π. If we substitute the point $X = (1,1,1)$ we get the nunber $2/\sqrt{3}$, as we have seen before.