A rigorous deductive approach to elementary Euclidean geometry

Jean-Pierre Demailly

Université Joseph Fourier Grenoble I

http://www-fourier.ujf-grenoble.fr/~demailly/documents.html

Résumé :

L'objectif de ce texte est de proposer une piste pour un enseignement logiquement rigoureux et cependant assez simple de la géométrie euclidienne au collège et au lycée. La géométrie euclidienne se trouve être un domaine très privilégié des mathématiques, à l'intérieur duquel il est possible de mettre en uvre dès le départ des raisonnements riches, tout en faisant appel de manière remarquable à la vision et à l'intuition. Notre préoccupation est d'autant plus grande que l'évolution des programmes scolaires depuis 3 ou 4 décennies révèle une diminution très marquée des contenus géométriques enseignés, en même temps qu'un affaiblissement du raisonnement mathématique auquel l'enseignement de la géométrie permettait précisément de contribuer de façon essentielle. Nous espérons que ce texte sera utile aux professeurs et aux auteurs de manuels de mathématiques qui ont la possibilité de s'affranchir des contraintes et des prescriptions trop indigentes des programmes officiels. Les premières sections devraient idéalement être maîtrisées aussi par tous les professeurs d'école, car il est à l'évidence très utile d'avoir du recul sur toutes les notions que l'on doit enseigner !

Mots-clés : géométrie euclidienne

Resumen :

El objetivo de este artículo es presentar un enfoque riguroso y aún razonablemente simples para la enseñanza de la geometría euclidiana elemental a nivel de educación secundaria. La geometría euclidiana es una área privilegiada de las matemáticas, ya que permite desde un primer nivel practicar razonamientos rigurosos y ejercitar la visión y la intuición. Nuestra preocupación es que las numerosas reformas de planes de estudio en las últimas 3 décadas en Francia, y posiblemente en otros países occidentales, han llevado a una disminución preocupante de la geometría, junto con un generalizado debilitamiento del razonamiento matemático al que la geometría contribuye específicamente de manera esencial. Esperamos que este punto de vista sea de interés para los autores de libros de texto y también para los profesores que tienen la posibilidad de no seguir exactamente las prescripciones sobre los contenidos menos relevantes, cuando están por desgracia impuestos por las autoridades educativas y por los planes de estudios. El contenido de las primeras secciones, en principio, debería también ser dominado por los profesores de la escuela primaria, ya que siempre es recomendable conocer más de lo que uno tiene que enseñar, a cualquier nivel !

Palabras clave : Geometría euclidiana

0. Introduction

The goal of this article is to explain a rigorous and still reasonably simple approach to teaching elementary Euclidean geometry at the secondary education levels. Euclidean geometry is a privileged area of mathematics, since it allows from an early stage to practice rigorous reasonings and to exercise vision and intuition. Our concern is that the successive reforms of curricula in the last 3 decades in France, and possibly in other western countries as well, have brought a worrying decline of geometry, along with a weakening of mathematical reasoning which geometry specifically contributed to in an essential way. We hope that these views will be of some interest to textbook authors and to teachers who have a possibility of not following too closely the prescriptions for weak contents, when they are unfortunately enforced by education authorities and curricula. The first sections should ideally also be mastered by primary school teachers, as it is always advisable to know more than what one has to teach at any given level !

Keywords : Euclidean geometry

1. On axiomatic approaches to geometry

As a formal discipline, geometry originates in Euclid's list of axioms and the work of his successors, even though substantial geometric knowledge existed before.


8

32. Celle qui a les costez egaux, mais non les angles droicts, se nomme Rhombe.
33. Mais celle qui a les costez opposez & les angles opposez egaux, mais n'a pas costez egaux ensemble ny les angles droicts, se nomme Rhomboïde.
34. Excepté celle-cy, les figures de quatre costez se nomment tablettes ou trapeses.
35. Lignes droictes paralleles sont celles, lesquelles estans en vn mesme plan & menees de part & d'autre ne concourent point ensemble.
36. Parallelogramme, est vne figure plane contenuë de quatre lignes droictes, desquelles les opposees sont paralleles entr'elles.

S'ENSVIVENT LES DEMANDES, ou positions simples.

Demande I.

D'Vn poinct à vn autre poinct mener vne ligne droicte.
2. Continuer vne ligne droicte finie tant qu'il en sera besoin.
3. Descrire vn cercle à l'entour d'vn poinct, & d'vne telle distance qu'on voudra.

9

COMMVNES SENTENCES.

Premiere.

LEs choses egales à vne, sont egales entr'elles.
2. Et si à choses egales s'adioustent choses egales, les toutes seront egales.
3. Et si de choses egales, se leuent choses egales, les restes seront egales.
4. Et si à choses inegales s'adioustent choses egales, les toutes seront inegales.
5. Et si des choses inegales se leuent choses egales, les restes seront inegales.
6. Les choses qui sont doubles à vn mesme, sont egales entr'elles.
7. Et les choses qui sont moictié d'vne mesme, sont egales entr'elles.
8. Et les choses qui conuiennent ensemble & entre elles, sont egales entr'elles.
9. Le tout est plus grand que sa partie.
10. Tous angles droicts sont egaux entr'eux.
11. Si dessus deux lignes droictes tombe vne ligne droicte, faisant les angles dedans d'vne mesme part, plus petits que deux droicts, icelles estans prolongees infiniment se rencontreront de la part en laquelle les angles sont plus petits que deux droicts.
12. Deux lignes droictes ne comprennent pas vn espace.

B

An excerpt of Euclid's book

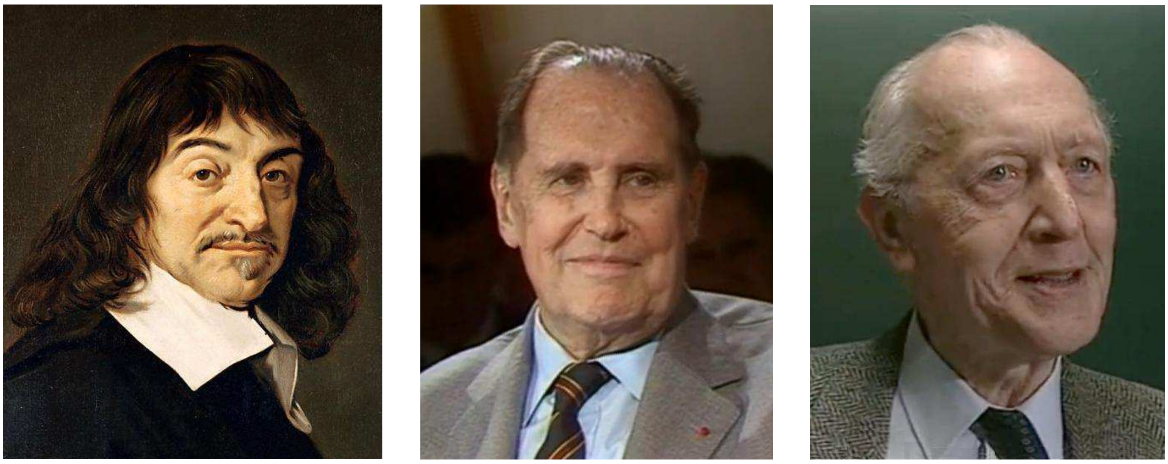
The traditional teaching of geometry that took place in France during the period 1880-1970 was directly inspired by Euclid's axioms, stating first the basic properties of geometric objects and using the "triangle isometry criteria" as the starting point of geometric reasoning. This approach had the advantage of being very effective and of quickly leading to rich contents. It also adequately reflected the intrinsic nature of geometric properties, without requiring extensive algebraic calculations. These choices echoed a mathematical tradition that was firmly rooted in the nineteenth century, aiming to develop "pure geometry", the highlight of which was the development of projective geometry by Poncelet.

Euclid's axioms, however, were neither complete nor entirely satisfactory from a logical perspective, leading mathematicians as Hilbert and Pasch to develop the system of axioms now attributed to Hilbert, that was settled in his famous memoir Grundlagen der Geometrie in 1899.

David Hilbert (1862–1943), in 1912

It should be observed, though, that the complexity of Hilbert's system of axioms makes it actually unpractical to teach geometry at an elementary level(1). The result, therefore, was that only a very partial axiomatic approach was taught, leading to a situation where a large number of properties that could have been proved formally had to be stated without proof, with the mere justification that they looked intutively true. This was not necessarily a major handicap, since pupils and their teachers may not even have noticed the logical gaps. However, such an approach, even though it was in some sense quite successful, meant that a substantial shift had to be accepted with more contemporary developments in mathematics, starting already with Descartes' introduction of analytic geometry. The drastic reforms implemented in France around 1970 (with the introduction of "modern mathematics", under the direction of André Lichnerowicz) swept away all these concerns by implementing an entirely new paradigm : according to Jean Dieudonné, one of the Bourbaki founders, geometry should be taught as a corollary of linear algebra, in a completely general and formal setting. The first step of the reform implemented this approach from "classe de seconde" (grade 10) on. A major problem, of course, is that the linear algebra viewpoint completely departs from the physical intuition of Euclidean space, where the group of invariance is the group of Euclidean motions and not the group of affine transformations.

(1) Even the improved and simplified version of Hilbert's axioms presented by Emil Artin in his famous book "Geometric Algebra" can hardly be taught before the 3rd or 4th year at university.



from Descartes (1596-1650) to Dieudonné (1906-1992) and Lichnerowicz (1915-1998)

The reform could still be followed in a quite acceptable way for about one decade, as long as pupils had a solid background in elementary geometry from their earlier grades, but became more and more unpractical when primary school and junior high school curricula were themselves (quite unfortunately) downgraded. All mathematical contents of high school were then severely axed around 1986, resulting in curricula prescriptions that in fact did not allow any more the introduction of substantial deductive activity, at least in a systematic way.

We believe however that it is necessary to introduce the basic language of mathematics, e.g. the basic concepts of sets, inclusion, intersection, etc, as soon as needed, most certainly already at the beginning of junior high school. Geometry is a very appropriate groundfield for using this language in a concrete way.

2. Geometry, numbers and arithmetic operations

An important issue is the relation between geometry and numbers. Greek mathematicians already had the fundamental idea that ratios of lengths with a given unit length were in one to one correspondence with numbers : in modern terms, there is a natural distance preserving bijection between points of a line and the set of real numbers. This viewpoint is of course not at all in contradiction with elementary education since measuring lengths in integer (and then decimal) values with a ruler is one of of the first important facts taught at primary school. However, at least in France, several reforms have put forward the extremely toxic idea that the emergence of electronic calculators would somehow free pupils from learning elementary arithmetic algorithms for addition, subtraction, multiplication and division, and that mastering magnitude orders and the "meaning" of arithmetic operations would be more than enough to understand society and even to pursue in science. The fact is that one cannot conceptually separate numbers from the operations that can be performed on them, and that mastering algorithms mentally and in written form is instrumental to realizing magnitude orders and the relation of numbers with physical quantities. The first contact that pupils will have with "elementary physics", again at primary school level, is probably through measuring lengths, areas, volumes, weights, densities, etc. Understanding the link with arithmetic operations is the basic knowledge that will be involved later to connect physics with mathematics. The idea of a real number as a possibly infinite decimal expansion then comes in a natural way when measuring a given physical quantity with greater and greater accuracy. Square roots are forced upon us by Pythagoras' theorem, and computing their numerical values is also a very good introduction to the concept of real number. I would certainly recommend to (re)introduce from the very start of junior high school (not later than grade 6 and 7), the observation that fractions of integers produce periodic decimal expansions, e.g. $1/7 = 0.\underline{142857}142857\ldots$, while no visible period appears when computing the square root of 2. In order to understand this (and before any formal proof can be given, as they are conceptually harder to grasp), it is again useful to learn here the hand and paper algorithm for computing square roots, which is only slightly more involved than the division algorithm and makes it immediately clear that there is no reason the result has to be periodic – unfortunately, this algorithm is no longer taught in France since a long time. When all this work is correctly done, it becomes really possible to give a precise meaning to the concept of real number at junior high school – of course many more details have to be explained, such as the identification of proper and improper decimal expansions, e.g. $1 = 0.99999\ldots$, the natural order relation on such expansions, decimal approximations with at given accuracy, etc. In what follows, we propose an approach to geometry based on the assumption that pupils have a reasonable understanding of numbers, arithmetic operations and physical quantities from their primary school years – with consolidation about things such infinite decimal expansions and square roots in the first two years of junior high school ; this was certainly the situation that prevailed in France before 1970, but things have unfortunately changed for the worse since then. Our small experimental network of schools SLECC ("Savoir Lire Ecrire Compter Calculer"), which accepts random pupils and operates in random parts of the country, shows that such knowledge can still be reached today for a very large majority of pupils, provided appropriate curricula are enforced. For the others, our views will probably remain a bit utopistic, or will have to be delayed and postponed at a later stage.

3. First steps of the introduction of Euclidean geometry

3.1. Fundamental concepts

The primitive concepts we are going to use freely are :

- real numbers, with their properties already discussed above ;
- points and geometric objects as sets of points : a point should be thought of as a geometric object with no extension, as can be represented with a sharp pencil ; a line or a curve are infinite sets of points (at this point, this is given only for intuition, but will not be needed formally) ;
- distances between points.

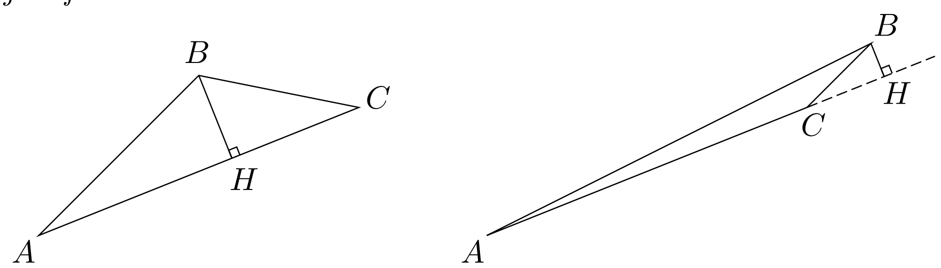
Let us mention that the language of set theory has been for more than one century the universal language of mathematicians. Although excessive abstraction should be avoided at early stages, we feel that it is appropriate to introduce at the beginning of junior high school the useful concepts of sets, of inclusion, the notation $x \in A$, operations on sets such as union, intersection and difference ; geometry and numbers already provide rich and concrete illustrations.

A geometric figure is simply an ordered finite collection of points A_j and sets S_k (vertices, segments, circles, arcs, ...)

Given two points A, B of the plane or of space, we denote by $d(A, B)$ (or simply by AB) their distance, which is in general a positive number, equal to zero when the points A and B coincide – concretely, this distance can be mesured with a ruler. A fundamental property of distances is :

3.1.1. Triangular inequality. *For any triple of points A, B, C, their mutual distances always satisfy the inequality $AC \leqslant AB + BC$, in other words the length of any side of a triangle is always at most equal to the sum of the lengths of the two other sides.*

Intuitive justification.



Let us draw the height of the triangle joining vertex B to point H on the opposite side (AC).

If H is located between A and C, we get $AC = AH + HC$; on the other hand, if the triangle is not flat (i.e. if $H \neq B$), we have $AH < AB$ and $HC < BC$ (since the hypotenuse is longer than the right-angle sides in a right-angle triangle – this will be checked formally thanks to Pythagoras' theorem). If H is located outside of the segment $[A, C]$, for instance beyond C, we already have $AC < AH \leqslant AB$, therefore $AC < AB \leqslant AB + BC$.

This justification[(2)] shows that the equality $AC = AB + BC$ holds if and only if the points A, B, C are aligned with B located between A and C (in this case, we have $H = B$ on the left part of the above figure). This leads to the following intrinsic definitions that rely on the concept of distance, and nothing more[(3)].

3.1.2. Definitions (segments, lines, half-lines).

(a) *Given two points A, B in a plane or in space, the segment $[A, B]$ of extremities A, B is the set of points M such that $AM + MB = AB$.*

---

(2) This is not a real proof since one relies on undefined concepts and on facts that have not yet been proved, for example, the concept of line, of perpendicularity, the existence of a point of intersection of a line with its perpendicular, etc ... This will actually come later (without any vicious circle, the justifications just serve to bring us to the appropriate definitions!)

(3) As far as they are concerned, these definitions are perfectly legitimate and rigorous, starting from our primitive concepts of points and their mutual distances. They would still work for other geometries such as hyperbolic geometry or general Riemannian geometry, at least when geodesic arcs are uniquely defined globally.

(b) *We say that three points A, B, C are aligned with B located between A and C if $B \in [A, C]$, and we say that they are aligned (without further specification) if one of the three points belongs to the segment determined by the two other points.*

(c) *Given two distinct points A, B, the line (AB) is the set of points M that are aligned with A and B ; the half-line $[A, B)$ of origin A containing point B is the set of points M aligned with A and B such that either M is located between A and B, or B between A and M. Two half-lines with the same origin are said to be opposite if their union is a line.*

In the definition, part (a) admits the following physical interpretration : a line segment can be realized by stretching a thin and light wire between two points A and B : when the wire is stretched, the points M located between A and B cannot "deviate", otherwise the distance AB would be shorter than the length of the wire, and the latter could still be stretched further . . .

We next discuss the notion of an axis : this is a line $\mathcal{D}$ equipped with an origin O and a direction, which one can choose by specifying one of the two points located at unit instance from O, with the abscissas $+1$ and -1 ; let us denote them respectively by I and I'. A point $M \in [O, I)$ is represented by the real value $x_M = +OM$ and a point M on the opposite half-line $[O, I')$ by the real value $x_M = -OM$. The algebraic measure of a bipoint (A, B) of the axis is defined by $\overline{AB} = x_B - x_A$, which is equal to $+AB$ or $-AB$ according to whether the ordering of A, B corresponds to the orientation or to its opposite. For any three points A, B, C of $\mathcal{D}$, we have the Chasles relation

$$\overline{AB} + \overline{BC} = \overline{AC}.$$

This relation can be derived from the equality $(x_B - x_A) + (x_C - x_B) = (x_C - x_A)$ after a simplification of the algebraic expression.

Building on the above concepts of distance, segments, lines and half-lines, we can now define rigorously what are planes, half-planes, circles, circle arcs, angles . . .[4]

3.1.3. Definitions.

(a) *Two lines $\mathcal{D}$, $\mathcal{D}'$ are said to be concurrent if their intersection consists of exactly one point.*

(b) *A plane $\mathcal{P}$ is a set of points that can be realized as the union of a family of lines (UV) such that U describes a line $\mathcal{D}$ and V a line $\mathcal{D}'$, for some concurrent lines $\mathcal{D}$ and $\mathcal{D}'$ in space. If A, B, C are 3 non aligned points, we denote by (ABC) the plane defined by the lines $\mathcal{D} = (AB)$ and $\mathcal{D}' = (AC)$ (say)*[5].

---

(4) Of course, this long series of definitions is merely intended to explain the sequence of concepts in a logical order. When teaching to pupils, it would be necessary to approach the concepts progressively, to give examples and illustrations, to let the pupils solve exercises and produce related constructions with instruments (ru2ler, compasses . . .).

(5) In a general manner, one could define by induction on n the concept of an affine subspace $\mathcal{S}_n$ of dimension n : this is the set obtained as the union of a family of lines (UV), where U describes a line $\mathcal{D}$ and V describes an affine subspace $\mathcal{S}_{n-1}$ of dimension $n-1$ intersecting $\mathcal{D}$ in exactly one point. Our definitions are valid in any dimension (even in an infinite dimensional ambient space), without taking special care !

(c) *Two lines $\mathcal{D}$ and $\mathcal{D}'$ are said to be parallel if they coincide, or if they are both contained in a certain plane $\mathcal{P}$ and do not intersect.*

(d) *A salient angle $\widehat{BAC}$ (or a salient angular sector) defined by two non opposite half-lines $[A,B)$, $[A,C)$ with the same origin is the set obtained as the union of the family of segments $[U,V]$ with $\in [A,B)$ and $V \in [A,C)$.*

(e) *A reflex angle (or a reflex angular sector) $\widecheck{BAC}$ is the complement of the corresponding salient angle $\widehat{BAC}$ in the plane (ABC), in which we agree to include the half-lines $[A,B)$ and $[A,C)$ in the boundary.*

(f) *Given a line $\mathcal{D}$ and a point M outside $\mathcal{D}$, the half-plane bounded by $\mathcal{D}$ containing M is the union of the two angular sectors $\widehat{BAM}$ and $\widehat{CAM}$ obtained by expressing $\mathcal{D}$ as the union of two opposite half-lines $[A,B)$ and $[A,C)$; this is the union of all segments $[U,V]$ such that $U \in \mathcal{D}$ and $V \in [A,M)$. The opposite half-plane is the one associated with the half-line $[A,M')$ opposite to $[A,M)$. In that situation, we also say that we have flat angles of vertex A.*

(g) *In a given plane $\mathcal{P}$, a circle of center A and radius $R > 0$ is the set of points M in the plane $\mathcal{P}$ such that $d(A,M) = AM = R$.*

(h) *A circular arc is the intersection of a circle with an angular sector, the vertex of which is the center of the circle.*

(i) *The measure of an angle (in degrees) is proportional to the length of the circular arc that it intercepts on a circle whose center coincides with the vertex of the angle, in such a way that the full circle corresponds to $360°$. A flat angle (cut by a half-plane bounded by a diameter of the circle) corresponds to an arc formed by a half-circle and has measure $180°$. A right angle is one half of a flat angle, that is, an angle corresponding to the quarter of a circle, in other words, an angle of measure equal to $90°$.*

(j) *Two half-lines with the same origin are said to be perpendicular if they form a right angle.*(5)

The usual properties of parallel lines and of angles intercepted by such lines ("corresponding angles" vs "alternate angles") easily leads to establishing the value of the sum of angles in a triangle (and, from there, in a quadrilateral).

Definition (i) requires of course a few comments. The first and most obvious comment is that one needs to define what is the length of a circular arc, or more generally of a curvilign arc : this is the limit (or the upper bound) of the lengths of polygonal line inscribed in the curve, when the curve is divided into smaller and smaller portions (cf. 2.2)(5). The second one is that the measure of an angle is independent of the radius R of the circle used to evaluate arc lengths; this follows from the fact that arc lengths are proportional to the radius R, which itself follows from Thales' theorem (see below).

(5) The concepts of right and flat angles, as well as the notion of half angle are already primary school concerns. At this level, the best way to address these issues is probably to let pupils practice paper folding (the notion of horizontality and verticality are relative concepts, it is better to avoid them when introducing perpendicularity, so as to avoid any potential confusion).

(5) The definition and existence of limits are difficult issues that cannot be addressed before high school, but it seems appropriate to introduce this idea at least intuitively.

Moreover, a proportionality argument yields the formula for the length of a circular arc located on a circle of radius R : a full arc (360°) has length $2\pi R$, hence the length of an arc of 1° is 360 times smaller, that is $2\pi R/360 = \pi R/180$, and an arc of measure a (in degrees) has length

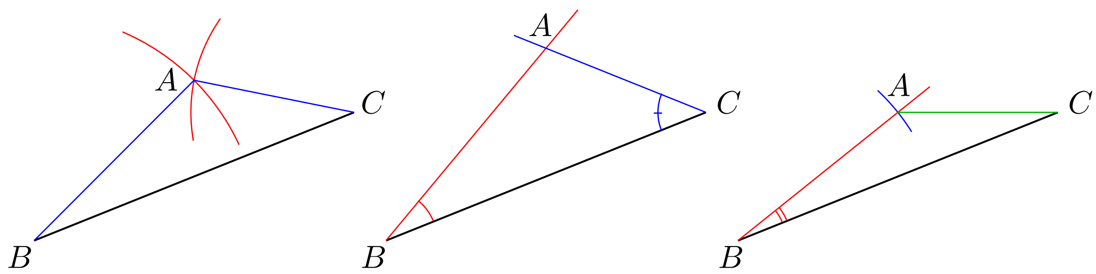
$$\ell = (\pi R/180) \times a = R \times a \times \pi/180.$$

3.2. Construction with instruments and isometry criteria for triangles

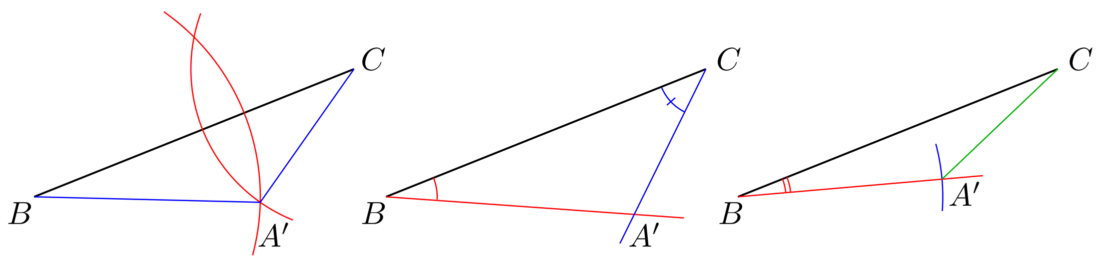
As soon as they are introduced, it is extremely important to illustrate geometric concepts with figures and construction activities with instruments. Basic constructions with ruler and compasses, such as midpoints, medians, bissectors, are of an elementary level and should be already taught at primary school. The step that follows immediately next consists of constructing perpendiculars and parallel lines passing through a given point.

At the beginning of junior high school, it becomes possible to consider conceptually more advanced matters, e.g. the problem of constructing a triangle ABC with a given base BC and two other elements, for instance :

(3.2.1) the lengths of sides AB and AC,
(3.2.2) the measures of angles $\widehat{ABC}$ and $\widehat{ACB}$,
(3.2.3) the length of AB and the measure of angle $\widehat{ABC}$.



In the first case, the solution is obtained by constructing circles of centers B, C and radii equal to the given lengths AB and AC, in the second case a protractor is used to draw two angular sectors with respective vertices B and C, in the third case one draws an angular sector of vertex B and a circle of center B. In each case it can be seen that there are exactly two solutions, the second solution being obtained as a triangle $A'BC$ that is symmetric of ABC with respect to line (BC) :



One sees that the triangles ABC and $A'BC$ have in each case sides with the same lengths. This leads to the important concept of *isometric figures*.

3.2.4. Definition.

(a) *One says that two triangles are isometric if the sides that are in correspondence have the same lengths, in such a way that if the first triangle has vertices* A, B, C *and the corresponding vertices of the second one are* A', B', C', *then* $A'B' = AB$, $B'C' = BC$, $C'A' = CA$.

(b) *More generally, one says that two figures in a plane or in space are isometric, the first one being defined by points* A_1, A_2, A_3, A_4 ... *and the second one by corresponding points* A'_1, A'_2, A'_3, A'_4 ... *if all mutual distances coincide.*

The concept of isometric figures is related to the physical concept of *solid body* : a body is said to be a solid if the mutual distances of its constituents (molecules, atoms) do not vary while the object is moved; after such a mo

ve, atoms which occupied certain positions A_i occupy new positions A'_i and we have $A'_iA'_j = A_iA_j$. This leads to a rigorous definition of solid displacements, that have a meaning from the viewpoints of mathematics and physics as well.

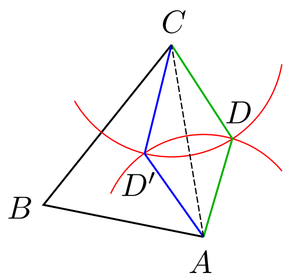
3.2.5. Definition. *Given a geometric figure (or a solid body in space) defined by characteristic points* A_1, A_2, A_3, A_4 ..., *a solid move is a continuous succession of positions* $A_i(t)$ *of these points with respect to the time* t, *in such a way that all distances* $A_i(t)A_j(t)$ *are constant. If the points* A_i *were the initial positions and the points* A'_i *are the final positions, we say that the figure* $(A'_1A'_2A'_3A'_4 \ldots)$ *is obtained by a displacement of figure* $(A_1A_2A_3A_4 \ldots)$.[(6)]

Beyond displacements, another way of producing isometric figures is to use a reflection (with respect to a line in a plane, or with respect to a plane in space, as obtained by taking the image of an object through reflection in a mirror)[(6)]. This fact is already observed with triangles, the use of transparent graph paper is then a good way of visualizing isometric triangles that cannot be superimposed by a displacement without "getting things out of the plane" ; in a similar way, it can be useful to construct elementary solid shapes (e.g. non regular tetrahedra) that cannot be superimposed by a solid move.

3.2.6. Exercise. In order to ensure that two quadrilaterals $ABCD$ and $A'B'C'D'$ are isometric, it is not sufficient to check that the four sides $A'B' = AB$, $B'C' = BC$, $C'D' = CD$, $D'A' = DA$ possess equal lengths, one must also check that the two diagonals $A'C' = AC$ and $B'D' = BD$ be equal; equaling only one diagonal is not enough as shown by the following construction :

(6) The concept of continuity that we use is the standard continuity property for functions of one real variable - one can of course introuduce this only intuitively at the junior high school level. One can further show that an isometry between two figures or solids extends an affine isometry of the whole space, and that a solid move is represented by a positive affine isometry, see Section 10. The formal proof is not very hard, but certainly cannot be given before the end of high school (this would have been possible with the rather strong French curricula as they were 50 years ago in the grade 12 science class, but doing so would be nowadays completely impossible).

(6) Conversely, an important theorem - which we will show later (see section 10) says that isometric figures can be deduced from each other either by a solid move or by a solid move preceded (or followed) by a reflection.



The construction problems considered above for triangles lead us to state the following fundamental isometry criteria.

3.2.7. Isometry criteria for triangles(7). *In order that two triangles be isometric, it is necessary and sufficient to check one of the following cases*

(a) *that the three sides be respectively equal* (*this is just the definition*), *or*

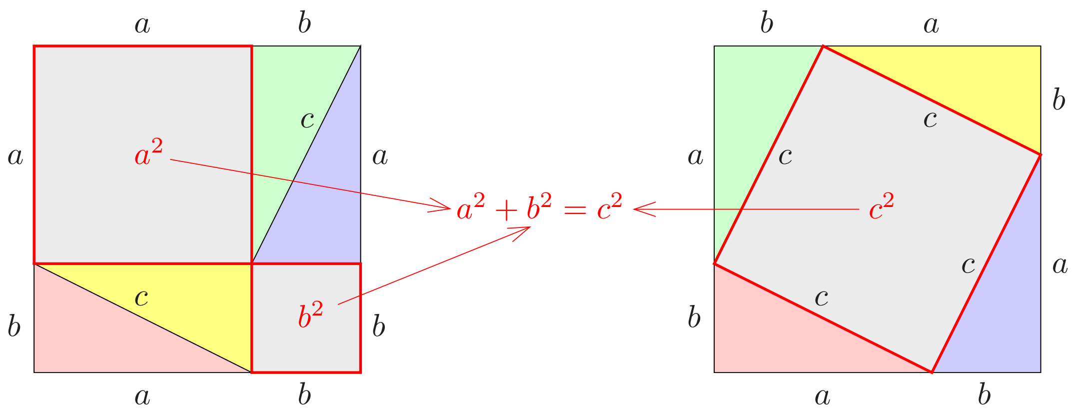
(b) *that they possess one angle with the same value and its adjacent sides equal, or*

(c) *that they possess one side with the same length and its adjacent angles of equal values.*

One should observe that conditions (b) and (c) are not sufficient if the adjacency specification is omitted - and it would be good to introduce (or to let pupils perform) constructions demonstrating this fact. A use of isometry criteria in conjunction with properties of alternate or corresponding angles leads to the various usual characterizations of quadrilaterals - parallelograms, lozenges, rectangles, squares . . .

3.3. Pythagoras' theorem

We first give the classical "Chinese" proof of Pythagoras' theorem, which is derived by a simple area argument based on moving four triangles (represented here in green, blue, yellow and light red). Its main advantage is to be visual and convincing(7).



(7) A rigorous formal proof of of these 3 isometry criteria will be given later, cf. Section 8.

(7) Again, in our context, the argument that will be described here is a justification rather than a formal proof. In fact, it would be needed to prove that the quadrilateral central figure on the right hand side is a square - this could certainly be checked with isometry properties of triangles - but one should not forget that they are not yet really proven at this stage. More seriously, the argument uses the concept of area, and it would be needed tp prove the existence of an area measure in the plane with all the desired properties: additivity by disjoint unions, translation invariance . . .

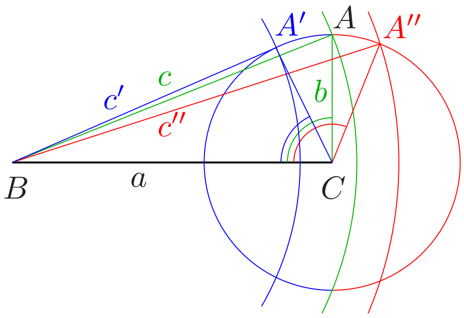
The point is to compare, in the left hand and right hand figures, the remaining grey area, which is the difference of the area of the square of side a + b with the area of the four rectangle triangles of sides a, b, c. The equality of the grey areas implies $a^2 + b^2 = c^2$.

Complement. *Let (ABC) be a triangle and a, b, c the lengths of the sides that are opposite to vertices A, B, C.*

(i) *If the angle $\widehat{C}$ is smaller than a right angle, we have $c^2 < a^2 + b^2$,*

(ii) *If the angle $\widehat{C}$ is larger than a right angle, we have $c^2 > a^2 + b^2$.*

Proof. First consider the case where (ABC) is rectangle: we have $c^2 = a^2 + b^2$ and the angle is equal to 90°.



If angle $\widehat{C}$ is $< 90°$, we have $c' < c$.

If angle $\widehat{C}$ is $> 90°$, we have $c'' > c$.

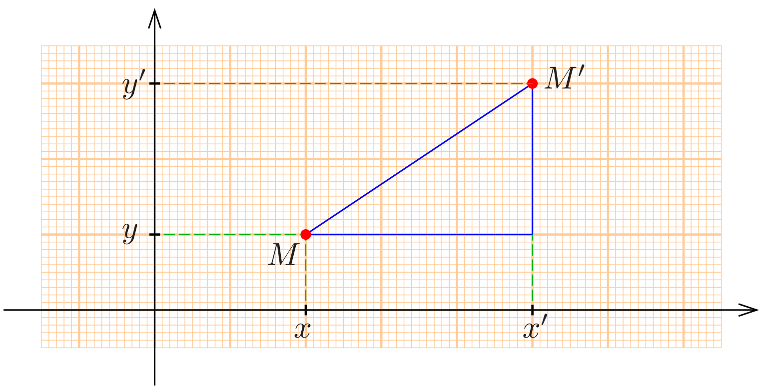
We argue by either increasing or decreasing the angle: if angle $\widehat{C}$ is $< 90°$, we have $c' < c$; if angle $\widehat{C}$ is $> 90°$, we have $c'' > c$. By this reasoning, we conclude :

Converse of Pythagoras' theorem. *With the above notation, if $c^2 = a^2 + b^2$, then angle $\widehat{C}$ must be a right angle, hence the given triangle is rectangle in C.*

4. Cartesian coordinates in the plane

The next fundamental step of our approach is the introduction of cartesian coordinates and their use to give *formal proofs of properties that had previously been taken for granted* (or given with a partial justification only). This is done by working in orthonormal frames.

4.1. Expression of Euclidean distance



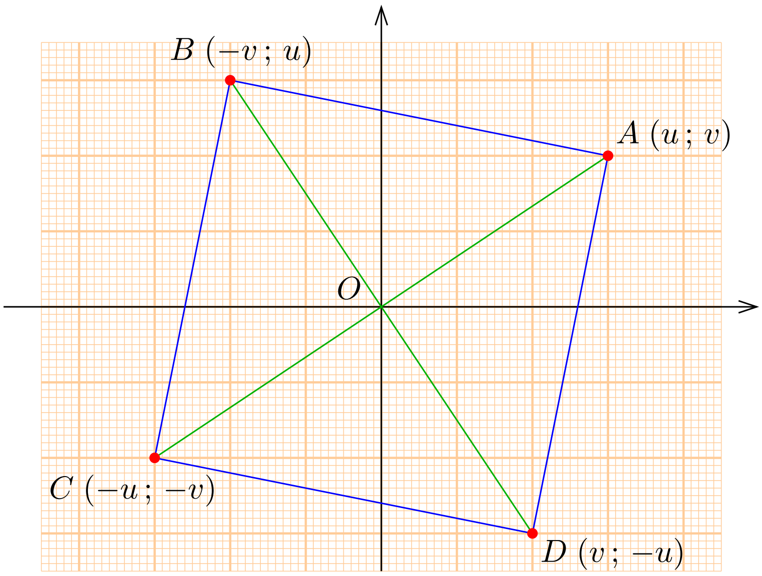
Pythagoras' theorem shows that the length MM' of the hypotenuse is given by the formula $MM'^2 = (x'-x)^2 + (y'-y)^2$, as the two sides of the right angle are $x'-x$ and $y'-y$ (up to sign). The distance from M to M' is therefore equal to

$$d(M, M') = MM' = \sqrt{(x'-x)^2 + (y'-y)^2}. \tag{4.1.1}$$

(It is of course advisable to first present the argument with simple numerical values).

4.2. Squares

Let us consider the figure formed by points $A\,(u\,;\,v)$, $B\,(-v\,;\,u)$, $C\,(-u\,;\,-v)$, $D\,(v\,;\,-u)$.



Formula (4.1.1) yields

$$AB^2 = BC^2 = CD^2 = DA^2 = (u+v)^2 + (u-v)^2 = 2(u^2+v^2),$$

hence the four sides have the same length, equal to $\sqrt{2}\sqrt{u^2+v^2}$. Similarly, we find

$$OA = OB = OC = OD = \sqrt{u^2+v^2},$$

therefore the 4 isoceles triangles OAB, OBC, OCD and ODA are isometric, and as a consequence we have $\widehat{OAB} = \widehat{OBC} = \widehat{OCD} = \widehat{ODA} = 90°$ and the other angles are equal to $45°$. Hence $\widehat{DAB} = \widehat{ABC} = \widehat{BCD} = \widehat{CDA} = 90°$, and we have proved that our figure is a square.

4.3. "Horizontal and vertical" lines

The set $\mathcal{D}$ of points $M(x\,;\,y)$ such that $y = c$ (where c is a given numerical value) is a "horizontal" line. In fact, given any three points M, M', M'' of abscissas $x < x' < x''$ we have

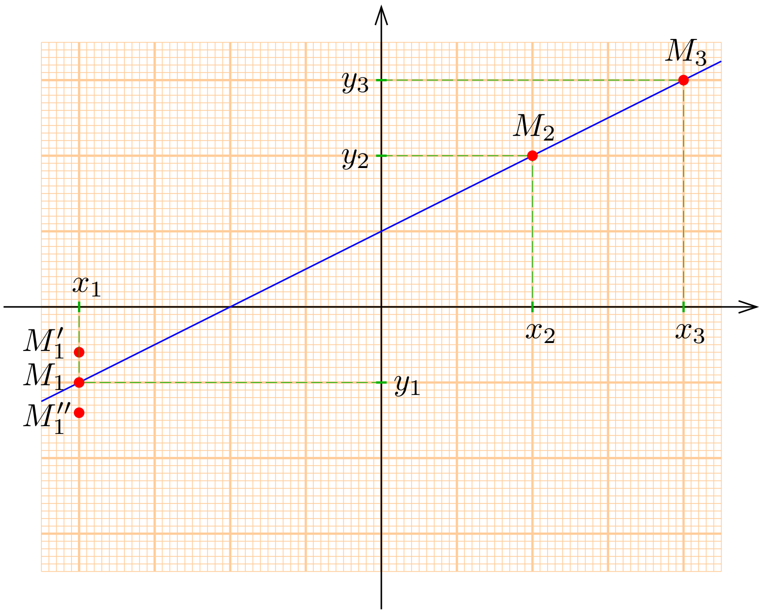
$$MM' = x' - x, \qquad M'M'' = x'' - x', \qquad MM'' = x'' - x$$

and therefore $MM' + M'M'' = MM''$. This implies by definition that our points M, M', M''. If we consider the line $\mathcal{D}_1$ given by the equation $y = c_1$ with $c_1 \neq c$, this is another horizontal line, and we have clearly $\mathcal{D} \cap \mathcal{D}_1 = \emptyset$, therefore our lines $\mathcal{D}$ and $\mathcal{D}_1$ are parallel.

Similarly, the set $\mathcal{D}$ of points $M(x\,;\,y)$ such that $x = c$ is a "vertical line" and the lines $\mathcal{D} : x = c$, $\mathcal{D}_1 : x = c_1$ are parallel.

4.4. Line defined by an equation $y = ax + b$

We start right away with the general case $y = ax + b$ to avoid any unnecessary repetitions, but with pupils it would be of course more appropriate to treat first the linear case $y = ax$.



Consider three points $M_1\ (x_1\,;\,y_1)$, $M_2\ (x_2\,;\,y_2)$, $M_3\ (x_3\,;\,y_3)$ satisfying the relations $y_1 = ax_1 + b$, $y_2 = ax_2 + b$ and $y_3 = ax_3 + b$, with $x_1 < x_2 < x_3$, say. As $y_2 - y_1 = a(x_2 - x_1)$, we find

$$M_1M_2 = \sqrt{(x_2 - x_1)^2 + a^2(x_2 - x_1)^2} = \sqrt{(x_2 - x_1)^2(1 + a^2)} = (x_2 - x_1)\sqrt{1 + a^2},$$

and likewise $M_2M_3 = (x_3 - x_2)\sqrt{1 + a^2}$, $M_1M_3 = (x_3 - x_1)\sqrt{1 + a^2}$. This shows that $M_1M_2 + M_2M_3 = M_1M_3$, hence our points M_1, M_2, M_3 are aligned. Moreover[(8)], we see that for any point $M_1'\ (x, y_1')$ with $y_1' > ax_1 + b$, then this point is not aligned with M_2 and M_3, and similarly for $M_1''\ (x, y_1'')$ such that $y_1'' < ax_1 + b$.

Consequence. *The set $\mathcal{D}$ of points $M\ (x\,;\,y)$ such that $y = ax + b$ is a line.*

The slope of line $\mathcal{D}$ is the ratio between the "vertical variation" and the "horizontal

(8) A rigorous formal proof would of course be possible by using a distance calculation, but this is much less obvious thanwhat we have done until now. One could however argue as in § 5.2 and use a new coordinate frame to reduce the situation to the case of the horizontal line $Y = 0$, in which case the proof is much easier.

variation", that is, for two points M_1 $(x_1\,;\,y_1)$, M_2 $(x_2\,;\,y_2)$ of $\mathcal{D}$ the ratio

$$\frac{y_2 - y_1}{x_2 - x_1} = a.$$

A horizontal line is a line of slope $a = 0$. When the slope a becomes very large, the inclination of the line $\mathcal{D}$ becomess intuitively close to being vertical. We therefore agree that a vertical line has infinite slope. Such an infinite value will be denoted by the symbol ∞ (without sign).

Consider two distinct points M_1 (x_1, y_1), M_2 (x_2, y_2). If $x_1 \neq x_2$, we see that there exists a unique line $\mathcal{D} : y = ax + b$ passing through M_1 and M_2 : its slope is given by $a = \frac{y_2 - y_1}{x_2 - x_1}$ and we infer $b = y_1 - ax_1 = y_2 - ax_2$. If $x_1 = x_2$, the unique line $\mathcal{D}$ passing through M_1, M_2 is the vertical line of equation $x = x_1$.

4.5. Intersection of two lines defined by their equations

Consider two lines $\mathcal{D} : y = ax+b$ and $\mathcal{D}' : y = a'x+b'$. In order to find the intersection $\mathcal{D} \cap \mathcal{D}'$ we write $y = ax + b = a'x + b'$, and get in this way $(a' - a)x = -(b' - b)$. Therefore, if $a \neq a'$, there is a unique intersection point $M(x\,;\,y)$ such that

$$x = -\frac{b' - b}{a' - a}, \qquad y = ax + b = \frac{-a(b' - b) + b(a' - a)}{a' - a} = \frac{ba' - ab'}{a' - a}.$$

The intersection of $\mathcal{D}$ with a vertical line $\mathcal{D}' : x = c$ is still unique, as we immediately find the solution $x = c$, $y = ac + b$. From this discussion, we can conclude:

Theorem. *Two lines $\mathcal{D}$ and $\mathcal{D}'$ possessing distinct slopes a, a' have a unique intersection point: we say that they are concurrent lines.*

On the contrary, if $a = a'$ and moreover $b \neq b'$, there is no possible solution, hence $\mathcal{D} \cap \mathcal{D}' = \emptyset$, our lines are distinct parallel lines. If $a = a'$ and $b = b'$, the lines $\mathcal{D}$ and $\mathcal{D}'$ are equal, and they are still considered as being parallel.

Consequence 1. *Consequence 1. Two lines $\mathcal{D}$ and $\mathcal{D}'$ of slopes a, a' are parallel if and only if their slopes are equal (finite or infinite).*

Consequence 2. *If $\mathcal{D}$ is parallel to $\mathcal{D}'$ and if $\mathcal{D}'$ is parallel to $\mathcal{D}''$, then $\mathcal{D}$ is parallel to $\mathcal{D}''$.*

Proof. In fact, if $a = a'$ and $a' = a''$, then $a = a''$. □

We can finally prove "Euclid's parallel postulate" (in our approach, this is indeed a rather obvious theorem, and not a postulate !).

Consequence 3. *Given a line $\mathcal{D}$ and a point M_0, there is a unique line $\mathcal{D}'$ parallel to $\mathcal{D}$ that passes through M_0.*

Proof. In fact, if $\mathcal{D}$ has a slope a and if $M_0(x_0\,;\,y_0)$, we see that

- for $a = \infty$, the unique possible line is the line $\mathcal{D}'$ of equation $x = x_0$;
- for $a \neq \infty$, the line $\mathcal{D}'$ has an equation $y = ax + b$ with $b = y_0 - ax_0$, therefore $\mathcal{D}'$ is the line that is uniquely defined by the equation $\mathcal{D}' : y - y_0 = a(x - x_0)$.

4.6. Orthogonality condition for two lines

Let us consider a line passing through the origin $\mathcal{D} : y = ax$. Select a point $M(u\,;v)$ located on $\mathcal{D}$, $M \neq O$, that is $u \neq 0$. Then $a = \frac{v}{u}$. We know that the point $M'\,(u'\,;v') = (-v\,;u)$ is such that the lines $\mathcal{D} = (OM)$ and (OM') are perpendicular, thanks to the construction of squares presented in section 4.2. Therefore, the slope of the line $\mathcal{D}' = (OM')$ perpendicular to $\mathcal{D}$ is given by

$$a' = \frac{v'}{u'} = \frac{u}{-v} = -\frac{u}{v} = -\frac{1}{a}$$

if $a \neq 0$. If $a = 0$, the line $\mathcal{D}$ coincides with the horizontal axis, its perpencular through O is the vertical axis of infinite slope. The formula $a' = -\frac{1}{a}$ is still true in that case if we agree that $\frac{1}{0} = \infty$ (let us repeat again that here ∞ means an infinite non signed value).

Consequence 1. *Two lines $\mathcal{D}$ and $\mathcal{D}'$ of slopes a, a' are perpendicular if and only if their slopes satisfy the condition $a' = -\frac{1}{a} \Leftrightarrow a = -\frac{1}{a'}$ (agreeing that $\frac{1}{\infty} = 0$ and $\frac{1}{0} = \infty$)).*

Consequence 2. *If $\mathcal{D} \perp \mathcal{D}'$ and $\mathcal{D}' \perp \mathcal{D}''$ then $\mathcal{D}$ and $\mathcal{D}''$ are parallel.*

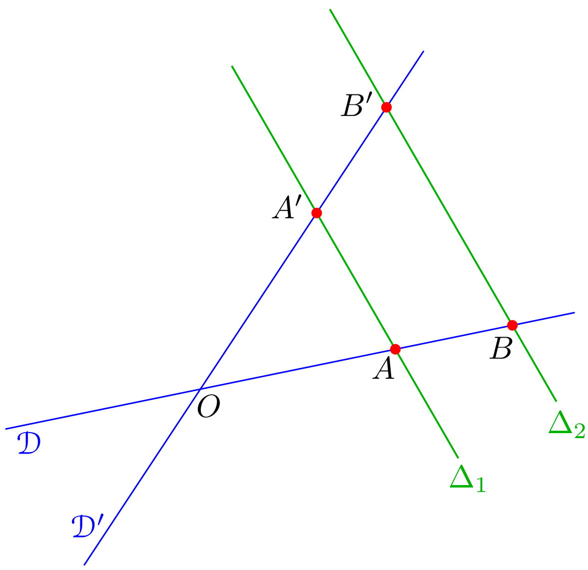
Proof. In fact, the slopes satisfy $a = -\frac{1}{a'}$ and $a'' = -\frac{1}{a'}$, hence $a'' = a$. □

4.7. Thales' theorem

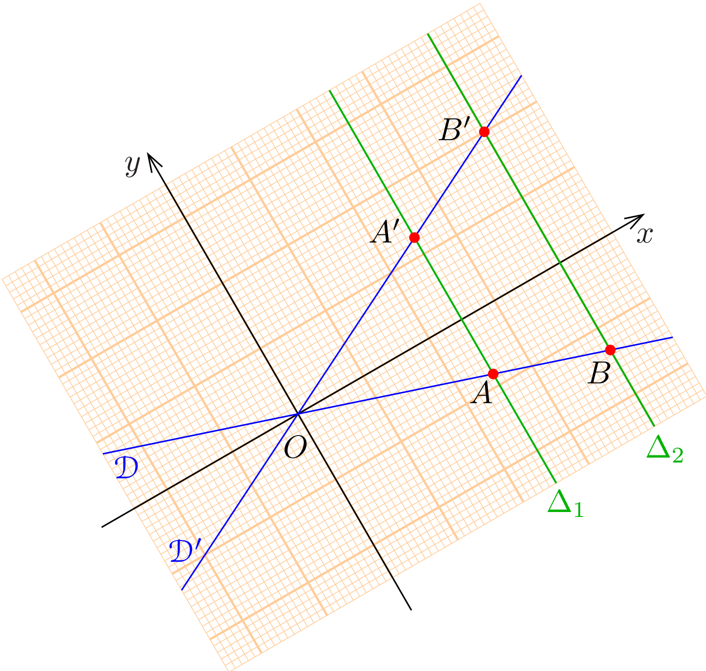
We start by stating a "Euclidean version" of the theorem, involving ratios of distances rather than ratios of algebraic measures.

Thales' theorem. *Consider two concurrent lines $\mathcal{D}$, $\mathcal{D}'$ intersecting in a point O, and two parallel lines Δ_1, Δ_2 that intersect $\mathcal{D}$ in points A, B, and $\mathcal{D}'$ in points A', B' ; we assume that A, B, A', B' are different from O. Then the length ratios satisfy*

$$\frac{OB}{OA} = \frac{OB'}{OA'} = \frac{BB'}{AA'}.$$



Proof. We argue by means of a coordinate calculation, in an orthonormal frame Oxy such that Ox is perpendicular to lines Δ_1, Δ_2, and Oy is parallel to lines Δ_1, Δ_2.



In these coordinates, lines Δ_1, Δ_2 are "vertical" lines of respective equations

$$\Delta_1 : x = c_1, \qquad \Delta_2 : x = c_2$$

with c_1, $c_2 \neq 0$, and our lines $\mathcal{D}$, $\mathcal{D}'$ admit respective equations $\mathcal{D} : y = ax$, $\mathcal{D}' : y = a'x$. Therefore

$$A\ (c_1, ac_1), \quad B\ (c_2, ac_2), \quad A'\ (c_1, a'c_1), \quad B'\ (c_2, a'c_2).$$

By Pythagoras' theorem we infer (after taking absolute values) :

$$OA = |c_1|\sqrt{1+a^2}, \;\; OB = |c_2|\sqrt{1+a^2}, \;\; OA' = |c_1|\sqrt{1+a'^2}, \;\; OB' = |c_2|\sqrt{1+a'^2},$$
$$AA' = |(a'-a)c_1|, \;\; BB' = |(a'-a)c_2|.$$

We have $a' \neq a$ since $\mathcal{D}$ and cD' are concurrent by our assumption, hence $a' - a \neq 0$, and we then conclude easily that

$$\frac{OB}{OA} = \frac{OB'}{OA'} = \frac{BB'}{AA'} = \frac{|c_2|}{|c_1|}. \qquad \square$$

In a more precise manner, if we choose orientations on $\mathcal{D}$, $\mathcal{D}'$ so as to turn them into axes, and also an orientation on Δ_1 and Δ_2, we see that in fact we have an equality of algebraic measures

$$\frac{\overline{OB}}{\overline{OA}} = \frac{\overline{OB'}}{\overline{OA'}} = \frac{\overline{BB'}}{\overline{AA'}}.$$

Converse of Thales' theorem. *Let $\mathcal{D}$, $\mathcal{D}'$ be concurrent lines intersecting in O. If Δ_1 intersects $\mathcal{D}$, $\mathcal{D}'$ in distinct points A, A', and Δ_2 intersects $\mathcal{D}$, $\mathcal{D}'$ in distinct points B, B' and if*

$$\frac{\overline{OB}}{\overline{OA}} = \frac{\overline{OB'}}{\overline{OA'}}$$

then Δ_1 *and* Δ_2 *are parallel.*

Proof. It is easily obtained by considering the line δ_2 parallel to Δ_1 that passes through B, and its intersection point β' with $\mathcal{D}'$. We then see that $\overline{O\beta'} = \overline{OB'}$, hence $\beta' = B'$ and $\delta_2 = \Delta_2$, and as a consequence $\Delta_2 = \delta_2 \mathbin{//} \Delta_1$. □

4.8. Consequences of Thales and Pythagoras theorems

The conjunction of isometry criteria for triangles and Thales and Pythagoras theorems already allows (in a very classical way !) to establish many basic theorems of elementary geometry. An important concept in this respect is the concept of similitude.

Definition. *Two figures* $(A_1A_2A_3A_4\ldots)$ *and* $(A'_1A'_2A'_3A'_4\ldots)$ *are said to be similar in the ratio* k $(k>0)$ *if we have* $A'_iA'_j/A_iA_j = k$ *for all segments* $[A_i, A_j]$ *and* $[A'_i, A'_j]$ *that are in correspondance.*

An important case where similar figures are obtained is by applying a homothety with a given center, say point O : if O is chosen as the origin of coordinates and if to each point $M(x\,;y)$ we associate the point $M'(x'\,;y')$ such that $x' = kx$, $y' = ky$, then formula (4.1.1) shows that we indeed have $A'B' = |k|\,AB$, hence by assigning to each point A_i the corresponding point A'_i we obtain similar figures in the ratio $|k|$; this situation is described by saying that we have homothetic figures in the ratio k ; this ratio can be positive or negative (for instance, if $k=-1$, this is a central symmetry with respect to O). The isometry criteria for triangles immediately extend into criteria for similarity.

Similarity criteria for triangles. *In order to conclude that two triangles are similar, ii is necessary and sufficient that one of the following conditions is met*:

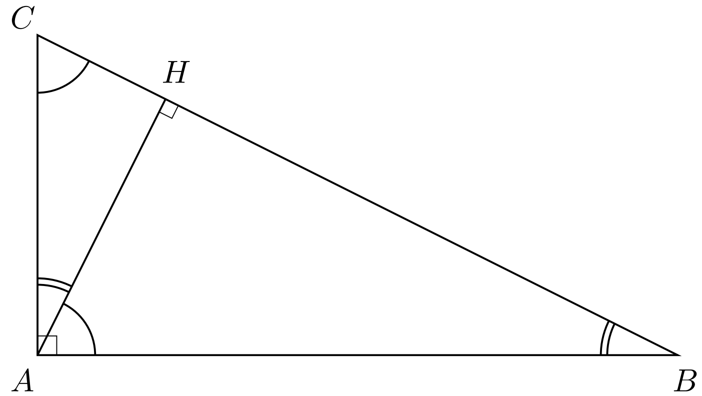
(a) *the corresponding three sides are proportional in a certain ratio* $k>0$ (*this is the definition*);

(b) *the triangles have a corresponding equal angle and the adjacent sides are proportional*;

(c) *the triangles have two equal angles in correspondence.*

An interesting application of the similarity criteria consists in stating and proving the basic metric relations in rectangle triangles: if the triangle ABC is rectangle in A and if H is the foot of the altitude drawn from vertex A, we have the basic relations

$$AB^2 = BH\cdot BC,\quad AC^2 = CH\cdot CB,\quad AH^2 = BH\cdot CH,\quad AB\cdot AC = AH\cdot BC.$$



In fact (for example) the similarity of rectangle triangles ABH and ABC leads to the equality of ratios

$$\frac{AB}{BC} = \frac{BH}{AB} \quad\Longrightarrow\quad AB^2 = BH \cdot BC.$$

One is also led in a natural way to the definition of sine, cosine and tangent of an acute angle in a rectangle triangle.

Definition. *Consider a triangle ABC that is rectangle in A. One defines*

$$\cos\widehat{ABC} = \frac{AB}{BC}, \qquad \sin\widehat{ABC} = \frac{AC}{BC}, \qquad \tan\widehat{ABC} = \frac{AC}{AB}.$$

In fact, the ratios only depend on the angle $\widehat{ABC}$ (which also determines uniquely the complementary angle $\widehat{ACB} = 90° - \widehat{ABC}$), since rectangle triangles that share a common angle else than their right angle are always similar by criterion (c). Pythagoras' theorem then quickly leads to computing the values of cos, sin, tan for angles with "remarkable values" 0°, 30°, 45°, 60°, 90°.

4.9. Computing areas and volumes

It is possible – and therefore probably desirable – to justify many basic formulas concerning areas and volumes of usual shapes and solid bodies (cylinders, pyramids, cones, spheres), just by using Thales and Pythagoras theorems, combined with elementary geometric arguments(9). We give here some indication on such techniques, in the case of cones and spheres. The arguments are close to those developed by Archimedes more than two centuries BC (except that we take here the liberty of reformulating them in modern algebraic notations).

Volume of a cone

The volume of a cone with an arbitrary plane base of area A and altitude h is given by

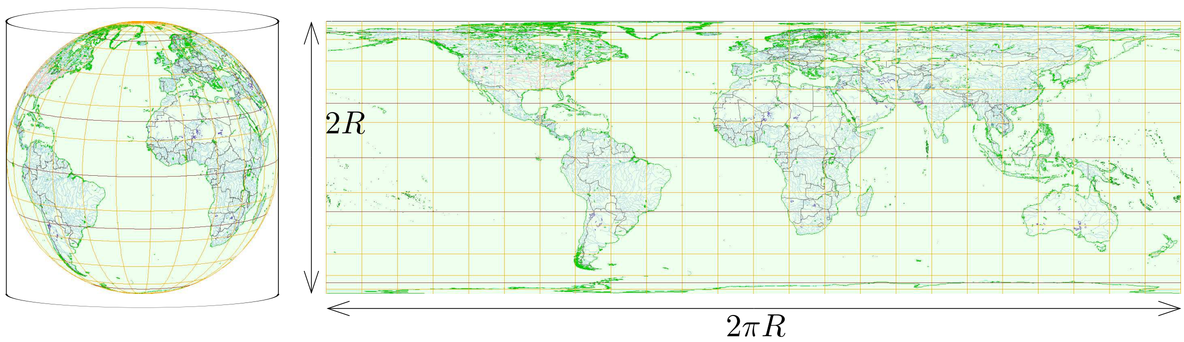
$$W = \frac{1}{3}\,Ah \tag{4.9.1}$$

One can indeed argue by a dilation argument that the volume V is proportional to h, and one also shows that is is proportional to A by approximating the base with a union of small squares. The proof is then reduced to the case of an oblique pyramid (i.e. to the case when the base is a rectangle). The coefficient $\frac{1}{3}$ is justified by observing that a cube can be divided in three identical oblique pyramids, whose summit is one of the vertices of the cube and the bases are the 3 adjacent opposite faces. The altitude of these pyramids is equal to the side of the cube, and their volume is thus $\frac{1}{3}$ of the volume of the cube.

(9) We are using here the word "justify" rather than "prove" because the necessary theoretical foundations (e.g. measure theory) are missing – and will probably be missing for 5-6 years or more. But in reality, one can see that these justifications can be made perfectly rigorous once the foundations considered here as intuitive are rigorously established. The concept of Hausdorff measure, as briefly explained in (11.3), can be used e.g. to give a rigorous definition of the p-dimensional measure of any object in a metric space, even when p is not an integer.

Archimedes formula for the area of a sphere of radius R

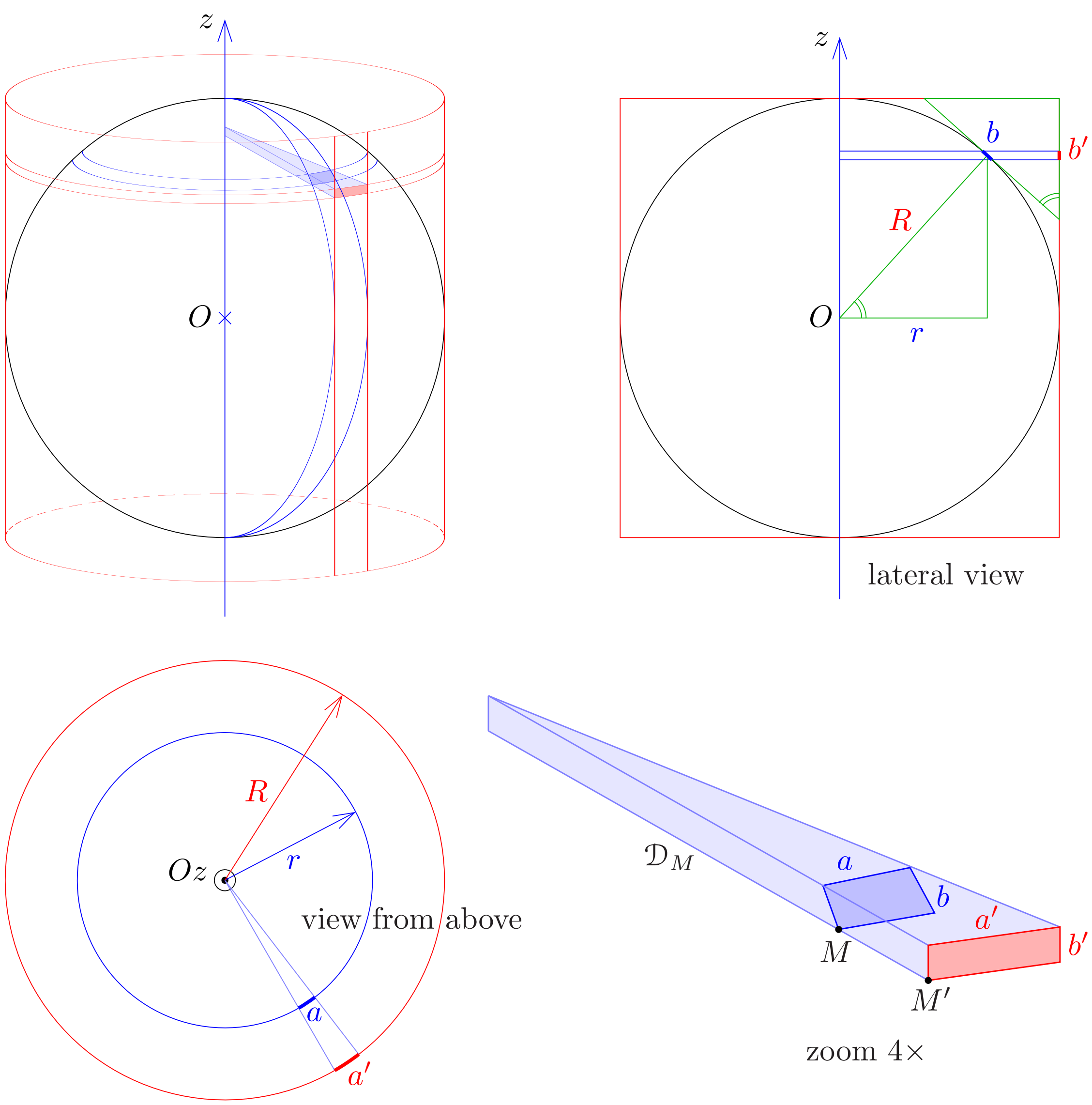
Since any two spheres of the same radius are isometric, their area depends only on the radius R. Let us take the center O of the sphere as the origin, and consider the "vertical" cylinder of radius R tangent to the sphere along the equator, and more precisely, the portion of cylinder located between the "horizontal" planes $z = -R$ and $z = R$. We use a "projection" of the sphere to the cylinder: for each point M of the sphere, we consider the point M' on the cylinder which is the intersection of the cylinder with the horizontal line $\mathcal{D}_M$ passing by M and intersecting the Oz axis. This projection is actually one of the simplest possible cartographic representations of the Earth. After cutting the cylinder along a meridian (say the meridian of longitude 180°), and unrolling the cylinder into a rectangle, we obtain the following cartographic map.



We are going to check that the cylindrical projection preserves areas, hence that the area of the sphere is equal to that of the corresponding rectangular map of sides $2R$ and $2\pi R$:

$$A = 2R \times 2\pi R = 4\pi R^2. \tag{4.9.2}$$

In order to check that the areas are equal, we consider a "rectangular field" delimited by parallel and meridian lines, of very small size with respect to the sphere, in such a way that it can be seen as a planar surface, i.e. to a rectangle (for instance, on Earth, one certainly does not realize the rotundity of the globe when the size of the field does not exceed a few hundred meters).



Let a, b be the side lengths of our "rectangular field", respectively along parallel lines direction and meridian lines direction, and a', b' the side lengths of the corresponding rectangle projected on the tangent cylinder.

In the view from above, Thales' theorem immediately implies

$$\frac{a'}{a} = \frac{R}{r}.$$

In the lateral view, the two triangles represented in green are homothetic (they share a common angle, as the adjacent sides are perpendicular to each other). If we apply again Thales' theorem to the tangent triangle and more specifically to the sides adjacent to the common angle, we get

$$\frac{b'}{b} = \frac{\text{adjacent small side}}{\text{hypotenuse}} = \frac{r}{R}.$$

The product of these equalities yields

$$\frac{a' \times b'}{a \times b} = \frac{a'}{a} \times \frac{b'}{b} = \frac{R}{r} \times \frac{r}{R} = 1.$$

We conclude from there that the rectangle areas $a \times b$ and $a' \times b'$ are equal. This implies that the cylindrical projection preserves areas, and formula (4.9.2) follows.

5. An axiomatic approach to Euclidean geometry

Although we have been able to follow a deductive presentation when it is compared to some of the more traditional approaches – almost all of the statements were "proven" from the definitions – it should nevertheless be observed that some proofs relied merely on intuitive facts – this was for instance the case of the "proof" of Pythagoras' Theorem. The only way to break the vicious circle is to take some of the facts that we feel necessary to use as "axioms", that is to say, to consider them as assumptions from which we first deduct all other properties by logical deduction ; a choice of other assumptions as our initial premises leads to non-Euclidean geometries (see section 10).

As we shall see, the notion of a Euclidean plane can be defined using a single axiom, essentially equivalent to the conjunction of Pythagoras' Theorem - which was only partially justified - and the existence of Cartesian coordinates - which we had not discussed either. In case the idea of using an axiomatic approach would look frightening, we want to stress that this section may be omitted altogether – provided pupils are in some way brought to the idea that the coordinate systems can be changed (translated, rotated, etc.) according to the needs.

5.1. The "Pythagoras/Descartes" model

In our vision, plane Euclidean geometry is based on the following "axiomatic definition".

Definition. *What we will call a Euclidean plane is a set of points denoted $\mathcal{P}$, for which mutual distances of points are supposed to be known, i.e. there is a predefined function*

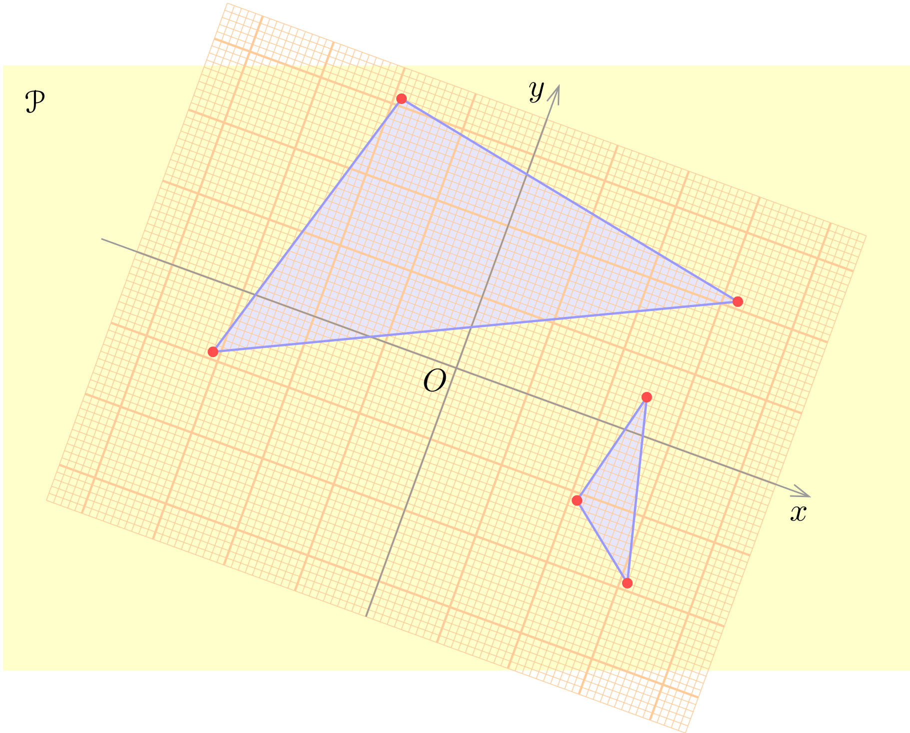
$$d : \mathcal{P} \times \mathcal{P} \longrightarrow \mathbb{R}_+, \qquad (M, M') \longmapsto d(M, M') = MM' \geqslant 0,$$

and we assume that there exist "orthonormal coordinate systems": to each point one can assign a pair of coordinates, by means of a one-to-one correspondence $M \mapsto (x\,;\,y)$ satisfying the axiom[10]

$$\text{(Pythagoras/Descartes)} \qquad d(M, M') = \sqrt{(x' - x)^2 + (y' - y)^2}$$

for all points $M\ (x\,;y)$ and $M'\ (x'\,;\,y')$.

It is certainly a good practice to represent the choice of an orthonomal coordinate system by using a transparent sheet of graph paper and placing it over the paper sheet that contains the working area of the Euclidean plane (here that area contains two triangles depicted in blue, above which the transparent sheet of graph paper has been placed).



This already shows (at an intuitive level only at this point) that there is an infinite number of possible choices for the coordinate systems. We now investigate this in more detail.

5.1.1. Rotating the sheet of graph paper around O by $180°$

A rotation of $180°$ of the graph paper around O has the effect of just changing the orientation of axes. The new coordinates $(X\,;Y)$ are given with respect to the old ones by

$$X = -x, \qquad Y = -y.$$

Since $(-u)^2 = u^2$ for every real number u, we see that the formula

$$(*) \qquad d(M, M') = \sqrt{(X' - X)^2 + (Y' - Y)^2}$$

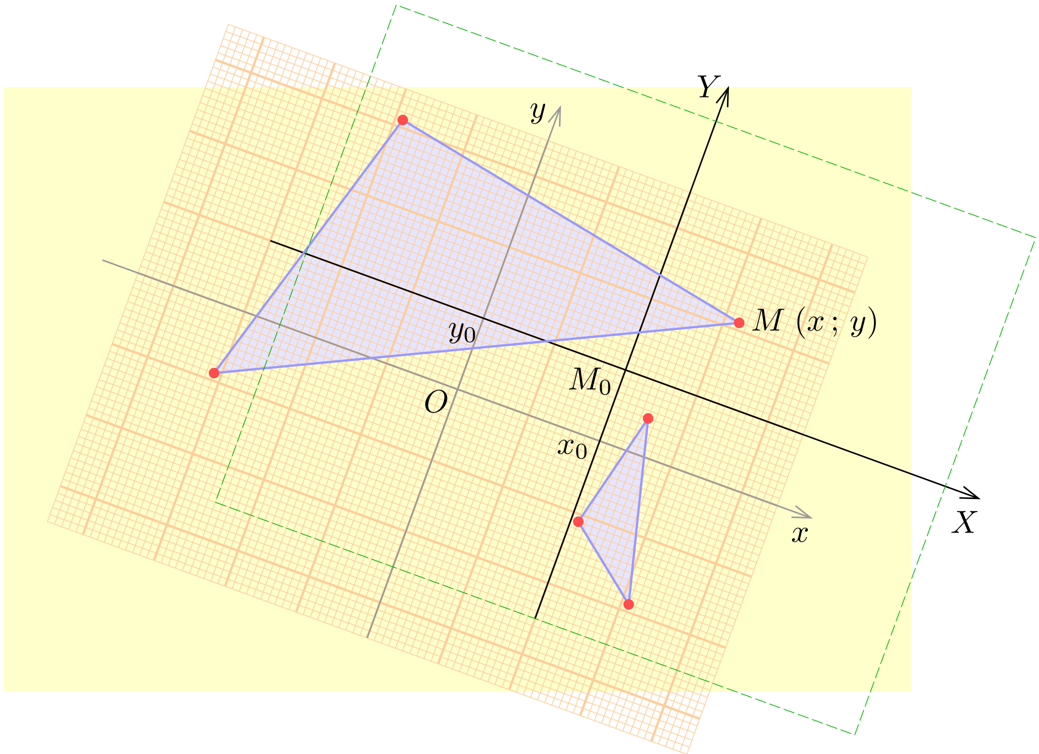
is still valid in the new coordinates, assuming it was valid in the original coordinates $(x\,;y)$.

5.1.2. Reversing the sheet of graph paper along one axis

If we reverse along Ox, we get $X = -x$, $Y = y$ and formula $(*)$ is still true. The argument is similar when reversing the sheet along Oy, we get the change of coordinates $X = x$, $Y = -y$ in that case.

5.1.3. Change of origin

Here we replace the origin O by an arbitrary point $M_0\ (x_0\,;y_0)$.



The new coordinates of point $M\ (x\,;\,y)$ are given by

$$X = x - x_0, \qquad Y = y - y_0.$$

For any two points M, M', we get in this situation

$$X' - X = (x' - x_0) - (x - x_0) = x' - x, \qquad Y' - Y = (y' - y_0) - (y - y_0) = y' - y$$

and we see that formula $(*)$ is still unchanged.

5.1.4. Rotation of axes

We will show that when the origin O is chosen, one can get the half-line Ox to pass through an aribrary point $M_1\ (x_1\,;\,y_1)$ distinct from O. This is intuitively obvious by "rotating" the sheet of graph paper around point O, but requires a formal proof relying on our "Pythagoras/Descartes" axiom. This proof is substantially more involved than what we have done yet, and can probably be jumped over at first – we give it here to show that there is no logical flaw in our approach. We start from the algebraic equality called *Lagrange's identity*

$$(au + bv)^2 + (-bu + av)^2 = a^2u^2 + b^2v^2 + b^2u^2 + a^2v^2 = (a^2 + b^2)(u^2 + v^2),$$

which is valid for all real numbers a, b, u, v. It can be obtained by developping the squares on the left and observing that the double products annihilate. As a consequence, if a and b satisfy $a^2 + b^2 = 1$ (such an example is $a = 3/5$, $b = 4/5$) and if we perform the change of coordinates

$$X = ax + by, \qquad Y = -bx + ay$$

we get, for any two points M, M' in the plane

$$X' - X = a(x' - x) + b(y' - y), \qquad Y' - Y = -b(x' - x) + a(y' - y),$$
$$(X' - X)^2 + (Y' - Y)^2 = (x' - x)^2 + (y' - y)^2$$

by Lagrange's identity with $u = x' - x$, $v = y' - y$. On the other hand, it is easy to check that

$$aX - bY = x, \qquad bX + aY = y,$$

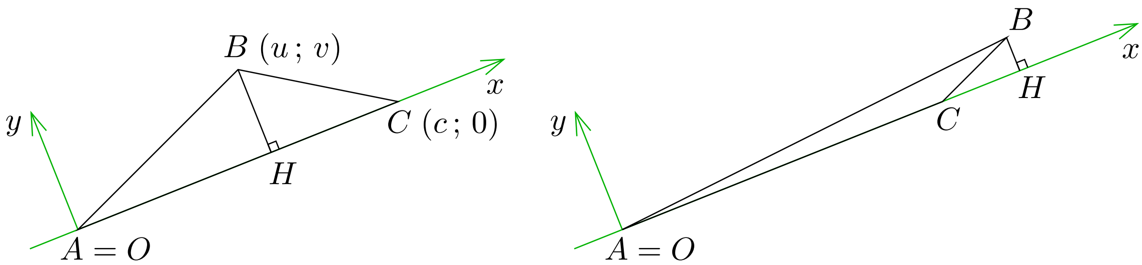
hence the assignment $(x\,;\,y) \mapsto (X\,;\,Y)$ is one-to-one. We infer from there that in the sense of our definition, $(X\,;\,Y)$ is indeed an orthonormal coordinate system. If we now choose $a = kx_1$, $b = ky_1$, the coordinates of point M_1 $(x_1\,;\,y_1)$ are transformed into

$$X_1 = ax_1 + by_1 = k(x_1^2 + y_1^2), \qquad Y_1 = -bx_1 + ay_1 = k(-y_1x_1 + x_1y_1) = 0,$$

and the condition $a^2 + b^2 = k^2(x_1^2 + y_1^2) = 1$ is satisfied by taking $k = 1/\sqrt{x_1^2 + y_1^2}$. Since $X_1 = \sqrt{x_1^2 + y_1^2} > 0$ and $Y_1 = 0$, the point M_1 is actually located on the half-line OX in the new coordinate system.

5.2. Revisiting the triangular inequality

The proof given in 3.1.1, which relied on facts that were not entirely settled, can now be made completely rigorous.



Given three distinct points A, B, C distincts, we select $O = A$ as the origin and the half line $[A, C)$ as the Ox axis. Our three points then have coordinates

$$A\ (0\,;\,0), \qquad B\ (u\,;\,v), \qquad C\ (c\,;\,0), \qquad c > 0,$$

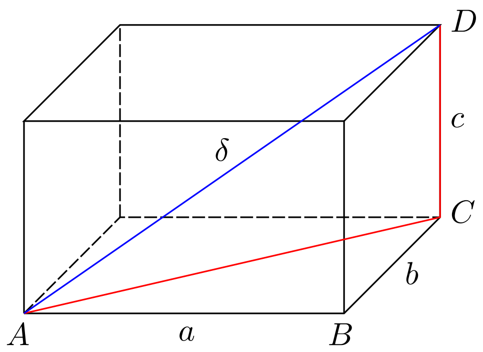
and the foot H of the altitude starting at B is H $(u\,;\,0)$. We find $AC = c$ and

$$AB = \sqrt{u^2 + v^2} \geqslant AH = |u| \geqslant u, \quad BC = \sqrt{(c - u)^2 + v^2} \geqslant HC = |c - u| \geqslant c - u.$$

Therefore $AC = c = u + (c - u) \leqslant AB + BC$ in all cases. The equality only holds when we have at the same time $v = 0$, $u \geqslant 0$ and $c - u \geqslant 0$, i.e. $u \in [0, c]$ and $v = 0$, in other words when B is located on the segment $[A, C]$ of the Ox axis.

5.3. Axioms of higher dimensional affine spaces

The approach that we have described is also appropriate for the introduction of Euclidean geometry in any dimension, especially in dimension 3. The starting point is the calculation of the diagonal δ of a rectangular parallelepiped with sides a, b, c:



As the triangles ACD and ABC are rectangle in C and B respectively, we have

$$AD^2 = AC^2 + CD^2 \quad \text{and} \quad AC^2 = AB^2 + BC^2$$

hence the "great diagonal" of our rectangle parallelepiped is given by

$$\delta^2 = AD^2 = AB^2 + BC^2 + CD^2 = a^2 + b^2 + c^2 \quad \Rightarrow \quad \delta = \sqrt{a^2 + b^2 + c^2}.$$

This leads to the expression of the distance function in dimension 3

$$d(M, M') = \sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2}$$

and we can just adopt the latter formula as the 3-dimensional Pythagoras/Descartes axiom.

6. Foundations of vector calculus

We will work here in the plane to simplify the exposition, but the only change in higher dimension would be the appearance of additional coordinates.

6.1. Median formula

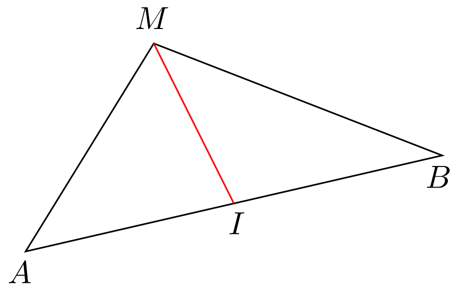
Consider points A, B with coordinates $(x_A\,;y_A)$, $(x_B\,;y_B)$ in an orthonormal frame Oxy.The point I of coordinates

$$x_I = \frac{x_A + x_B}{2}, \qquad y_I = \frac{y_A + y_B}{2}$$

satisfies $IA = IB = \frac{1}{2}AB$: this is the midpoint of segment $[A, B]$.

Median formula. *For every point M $(x\,;y)$, one has*

$$MA^2 + MB^2 = 2\,MI^2 + \frac{1}{2}AB^2 = 2\,MI^2 + 2\,IA^2.$$



Proof. In fact, by expanding the squares, we get

$$(x - x_A)^2 + (x - x_B)^2 = 2x^2 - 2(x_A + x_B)x + x_A^2 + x_B^2,$$

while

$$\begin{aligned} 2(x - x_I)^2 + \frac{1}{2}(x_B - x_A)^2 &= 2(x^2 - 2x_I x + x_I^2) + \frac{1}{2}(x_B - x_A)^2 \\ &= 2\Big(x^2 - (x_A + x_B)x + \frac{1}{4}(x_A + x_B)^2\Big) + \frac{1}{2}(x_B - x_A)^2 \\ &= 2x^2 - 2(x_A + x_B)x + x_A^2 + x_B^2. \end{aligned}$$

Therefor we get

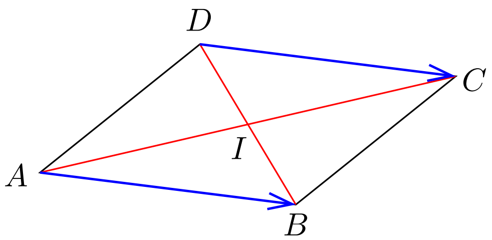
$$(x - x_A)^2 + (x - x_B)^2 = 2(x - x_I)^2 + \frac{1}{2}(x_B - x_A)^2.$$

The median formula is obtained by adding the analogous equality for coordinates y and applying Pythagoras' theorem. □

It follows from the median formula that there is a unique point M such that $MA = MB = \frac{1}{2}AB$, in fact we then find $MI^2 = 0$, hence $M = I$. The coordinate formulas that we initially gave to define midpoints are therefore independent of the choice of coordinates.

6.2. Parallelograms

A quadrilateral $ABCD$ is a parallelogram if and only if its diagonals $[A, C]$ and $[B, D]$ intersect at their midpoint :



In this way, we find the necessary and sufficient condition

$$x_I = \frac{1}{2}(x_B + x_D) = \frac{1}{2}(x_A + x_C), \qquad y_I = \frac{1}{2}(y_B + y_D) = \frac{1}{2}(y_A + y_C),$$

which is equivalent to

$$x_B + x_D = x_A + x_C, \qquad y_B + y_D = y_A + y_C$$

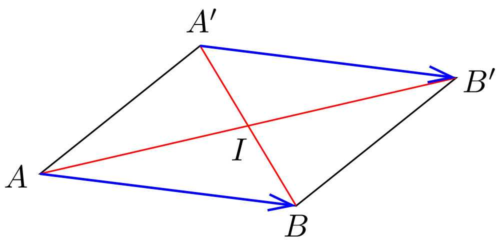
or, alternatively, t

$$x_B - x_A = x_C - x_D, \qquad y_B - y_A = y_C - y_D,$$

in other words, the variation of coordinates involved in getting from A to B is the same as the one involved in getting from D to C.

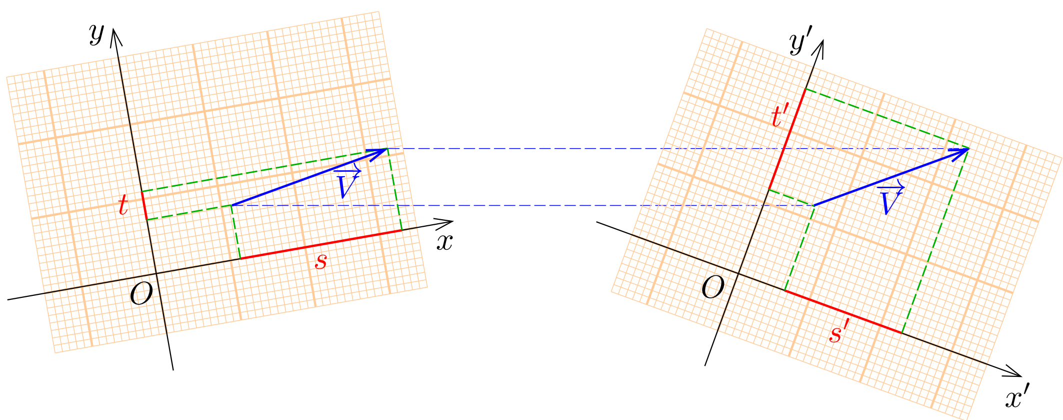
6.3. Vectors

A *bipoint* is an ordered pair (A, B) of points ; we say that A is the *origin* and that B is the *extremity* of the bipoint. The bipoints (A, B) and (A', B') are said to be *equipollent* if the quadrilateral $ABB'A'$ is a parallelogram (which can possibly be a "flat" parallelogram in case the four points are aligned).

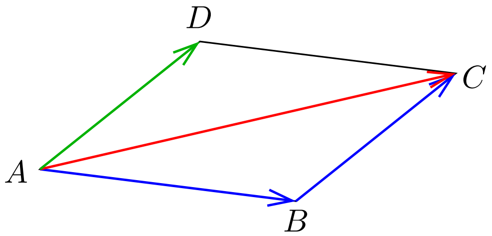


Definition. *Given two points A, B, the vector $\overrightarrow{AB}$ is the "variation of position" needed to get from A to B. Given a coordinate frame Oxy, this "variation of position" is expressed along the Ox axis by $x_B - x_A$ and along the Oy axis by $y_B - y_A$. If the bipoints (A, B) and (A', B') are equipollent, the vectors $\overrightarrow{AB}$ and $\overrightarrow{A'B'}$ are equal since the variations $x_{B'} - x_{A'} = x_B - x_A$ and $y_{B'} - y_{A'} = y_B - y_A$ are the same (this is true in any coordinate system).*

The "component" of vector $\overrightarrow{AB}$ in the coordinate system Oxy are the numbers denoted in the form of an ordered pair $(x_B - x_A \,;\, y_B - y_A)$. The components $(s\,;\,t)$ of a vector $\overrightarrow{V}$ depend of course on the choice of the coordinate frame Oxy : to a given vector $\overrightarrow{V}$ one assigns different components $(s\,;\,t)$, $(s'\,;\,t')$ in different coordinate frames Oxy, $Ox'y'$.



6.4. Addition of vectors



The addition of vectors is defined by means of *Chasles' relation*

$$\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC} \qquad (6.4.1)$$

for any three points A, B, C : when one takes the sum of the variation of position required to get from A to B, and then from B to C, one finds the variation of position to get from A to C ; actually, we have for instance

$$(x_B - x_A) + (x_C - x_B) = x_C - x_A$$

for the component along the Ox axis. Equivalently, if $ABCD$ is a parallelogram, one can also put

$$\overrightarrow{AB} + \overrightarrow{AD} = \overrightarrow{AC}. \qquad (6.4.2)$$

That (6.4.1) and (6.4.2) are equivalent follows from the fact that $\overrightarrow{AD} = \overrightarrow{BC}$ in parallelogram $ABCD$. For any choice of coordinae frame Oxy, the sum of vectors of components $(s\,;t)$, $(s'\,;t')$ has components $(s+s'\,;t+t')$.

For every point A, the vector $\overrightarrow{AA}$ has zero componaents: it will be denoted simply $\overrightarrow{0}$. Obviously, we have $\overrightarrow{V} + \overrightarrow{0} = \overrightarrow{0} + \overrightarrow{V} = \overrightarrow{V}$ for every vector $\overrightarrow{V}$. On the other hand, Chasles' relation yields

$$\overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \overrightarrow{0}$$

for all points A, B. Therefore we define

$$-\overrightarrow{AB} = \overrightarrow{BA},$$

in other words, the opposite of a vector is obtained by exchanging the origin and extremity of any corresponding bipoint.

6.5. Multiplication of a vector by a real number

Given a vector $\overrightarrow{V}$ of components $(s\,;t)$ in a coordinate frame Oxy and an arbitrary real number λ, we define $\lambda\overrightarrow{V}$ as the vector of components $(\lambda s\,;\lambda t)$.

This definition is actually independent of the coordinate frame Oxy. In fact if $\overrightarrow{V} = \overrightarrow{AB} \neq \overrightarrow{0}$ and $\lambda \geqslant 0$, we have $\lambda\overrightarrow{AB} = \overrightarrow{AC}$ where C is the unique point located on the half-line $[A,B)$ such that $AC = \lambda\, AB$. On the other hand, if $\lambda \leqslant 0$, we have $-\lambda \geqslant 0$ and

$$\lambda\overrightarrow{AB} = (-\lambda)(-\overrightarrow{AB}) = (-\lambda)\overrightarrow{BA}.$$

Finally, it is clear that $\lambda\,\overrightarrow{0} = \overrightarrow{0}$. Multiplication of vectors by a number is distributive with respect to the addition of vectors (this is a consequence of the distributivity of multiplication with respect to addition in the set of real numbers).

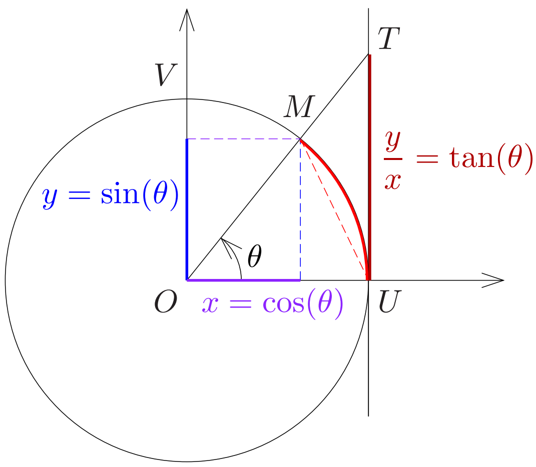
7. Cartesian equation of circles and trigonometric functions

By Pythagoras' thorem, the circle of center $A\ (a, b)$ and radius R in the plane is the set of points M satisfying the equation

$$AM = R \quad \Leftrightarrow \quad AM^2 = R^2 \quad \Leftrightarrow \quad (x-a)^2 + (y-b) = R^2,$$

which can also be put in the form $x^2 + y^2 - 2ax - 2by + c = 0$ with $c = a^2 + b^2 - R^2$. Conversely, the set of solutions of such an equation defines a circle of center $A\ (a\,;\,b)$ and of radius $R = \sqrt{a^2+b^2-c}$ if $c < a^2 + b^2$, is reduced to point A if $c = a^2 + b^2$, and is empty if $c > a^2 + b^2$.

The *trigonometric circle* $\mathcal{C}$ is defined to be the unit circle centered at the origin in an orthormal coordinate system Oxy, that is, the of points $M\ (x\,;\,y)$ such that $x^2+y^2 = 1$. Let U be the point of coordinates $(1\,;\,0)$ and V the point of coordinates $(0\,;\,1)$. The usual trigonometric functions cos, sin and tan are then defined for arbitrary angle arguments as shown on the above figure[10] :

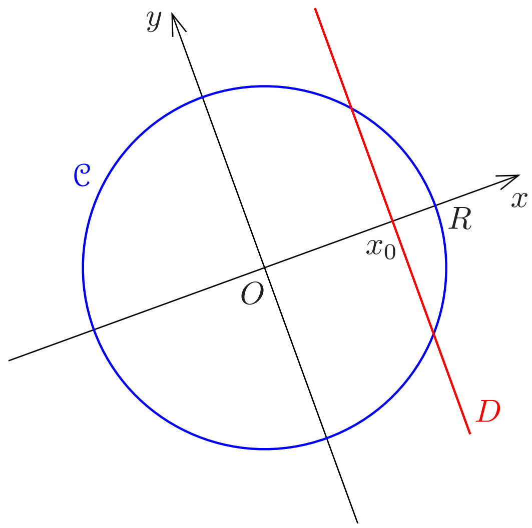


The equation of the circle implies the relation $(\cos\theta)^2 + (\sin\theta)^2 = 1$ for every θ.

8. Intersection of lines and circles

Let us begin by intersecting a circle $\mathcal{C}$ of center A and radius R with an arbitrary line $\mathcal{D}$. In order to simplify the calculation, we take $A = O$ as the origin and we take the axis Ox to be perpendicular to the line $\mathcal{D}$. The line $\mathcal{D}$ is then "vertical" in the coordinate frame Oxy. (We start here right away with the most general case, but, once again, it would be desirable to approach the question by treating first simple numerical examples ...).

(10) It seems essential at this stage that functions cos, sin, tan have already been introduced as the ad hoc ratios of sides in a right triangle, i.e. at least for the case of acute angles, and that their values for the remarkable angle values 0°, 30°, 45°, 60°, 90° are known.



This leads to equation

$$\mathcal{C} : x^2 + y^2 = R^2, \qquad D : x = x_0,$$

hence

$$y^2 = R^2 - x_0^2.$$

As a consequence, if $|x_0| < R$, we have $R^2 - x_0^2 > 0$ and there are two solutions $y = \sqrt{R^2 - x_0^2}$ and $y = -\sqrt{R^2 - x_0^2}$, corresponding to two intersection points$(x_0, \sqrt{R^2 - x_0^2})$ and $(x_0, -\sqrt{R^2 - x_0^2})$ that are symmetric with respect to the Ox axis. If $|x_0| = R$, we find a single solution $y = 0$: the line $\mathcal{D} : x = x_0$ is tangent to circle $\mathcal{C}$ at point $(x_0\,; 0)$. If $|x_0| > R$, the equation $y^2 = R^2 - x_0^2 < 0$ has no solution ; the line $\mathcal{D}$ does not intersect the circle.

Consider now the intersection of a circle $\mathcal{C}$ of center A and radius R with a circle $\mathcal{C}'$ of center A' and radius R'. Let $d = AA'$ be the distance between their centers. If $d = 0$ the circles are concentric and the discussion is easy (the circles coincide if $R = R'$, and are disjoint if $R \neq R'$). We will therefore assume that $A \neq A'$, i.e. $d > 0$. By selecting $O = A$ as the origin and $Ox = [A, A')$ as the positive x axis, we are reduces to the case where $A\,(0\,; 0)$ and $A'\,(d\,; 0)$. We then get equations

$$\mathcal{C} : x^2 + y^2 = R^2, \qquad \mathcal{C}' : (x - d)^2 + y^2 = R'^2 \Longleftrightarrow x^2 + y^2 = 2dx + R'^2 - d^2.$$

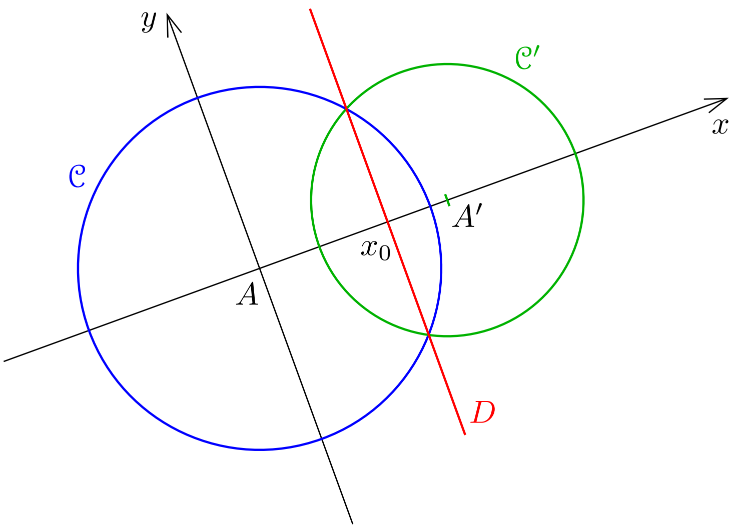
For any point M in the intersection $\mathcal{C} \cap \mathcal{C}'$, we thus get $2dx + R'^2 - d^2 = R^2$, hence

$$x = x_0 = \frac{1}{2d}(d^2 + R^2 - R'^2).$$

This shows that the intersection $\mathcal{C} \cap \mathcal{C}'$ is contained in the intersection $\mathcal{C} \cap D$ of $\mathcal{C}$ with the line $\mathcal{D} : x = x_0$. Conversely, one sees that if $x^2 + y^2 = R^2$ and $x = x_0$, then $(x\,; y)$ also satisfies the equation

$$x^2 + y^2 - 2dx = R^2 - 2dx_0 = R^2 - (d^2 + R^2 - R'^2) = R'^2 - d^2$$

which is the equation of $\mathcal{C}'$, hence $\mathcal{C} \cap D \subset \mathcal{C} \cap \mathcal{C}'$ and finally $\mathcal{C} \cap \mathcal{C}' = \mathcal{C} \cap D$.



The intersection points are thus given by $y = \pm\sqrt{R^2 - x_0^2}$. As a consequene, we have exactly two solutions that are symmetric with respect to the line (AA') as soon as $-R < x_0 < R$, or equivalently

$$\begin{aligned} -2dR < d^2 + R^2 - R'^2 < 2dR &\iff (d+R)^2 > R'^2 \ \text{ et } \ (d-R)^2 < R'^2 \\ &\iff d + R > R', \ \ d - R < R', \ \ d - R > -R', \end{aligned}$$

i.e. $|R - R'| < d < R + R'$. If one of the inequalities is an equality, we get $x_0 = \pm R$ and we thus find a single solution $y = 0$. The circles are tangent internally if $d = |R - R'|$ and tangent externally if $d = R + R'$.

Note that these results lead to a complete and rigorous proof of the isometry criteria for triangles : up to an orthonormal change of coordinates, each of the three cases entirely determines the coordinates of the triangles modulo a reflection with respect to Ox (in this argument, the origin O is chosen as one of the vertices and the axis Ox is taken to be the direction of a side of known length). The triangles specified in that way are thus isometric.

9. Scalar product

The *norm* $\|\overrightarrow{V}\|$ of a vector $\overrightarrow{V} = \overrightarrow{AB}$ is the length $AB = d(A, B)$ of an arbitrary bipoint that defines $\overrightarrow{V}$. From there, we put

$$\overrightarrow{U} \cdot \overrightarrow{V} = \frac{1}{2}\left(\|\overrightarrow{U} + \overrightarrow{V}\|^2 - \|\overrightarrow{U}\|^2 - \|\overrightarrow{V}\|^2\right) \tag{9.1}$$

in particular $\overrightarrow{U} \cdot \overrightarrow{U} = \|\overrightarrow{U}\|^2$. The real number $\overrightarrow{U} \cdot \overrightarrow{V}$ is called the *inner product* of $\overrightarrow{U}$ and $\overrightarrow{V}$, and $\overrightarrow{U} \cdot \overrightarrow{U}$ is also defined to be the inner square of $\overrightarrow{U}$, denoted $\overrightarrow{U}^2$. Consequently we obtain

$$\overrightarrow{U}^2 = \overrightarrow{U} \cdot \overrightarrow{U} = \|\overrightarrow{U}\|^2.$$

By definition (9.1), we have

$$\|\overrightarrow{U}+\overrightarrow{V}\|^2=\|\overrightarrow{U}\|^2+\|\overrightarrow{V}\|^2+2\,\overrightarrow{U}\cdot\overrightarrow{V}, \tag{9.2}$$

and this formula can also be rewritten

$$(\overrightarrow{U}+\overrightarrow{V})^2=\overrightarrow{U}^2+\overrightarrow{V}^2+2\,\overrightarrow{U}\cdot\overrightarrow{V}. \tag{9.2'}$$

This was the main motivation of the definition: that the usual identity for the square of a sum be valid for inner products. In dimension 2 and in an orthonormal frame Oxy, we find $\overrightarrow{U}^2=x^2+y^2$; if $\overrightarrow{V}$ has components $(x'\,;\,y')$, Definition (9.1) implies

$$\overrightarrow{U}\cdot\overrightarrow{V}=\frac{1}{2}\big((x+x')^2+(y+y')^2-(x^2+y^2)-(x'^2+y'^2)\big)=xx'+yy'. \tag{9.3}$$

In dimension n, we would find similarly

$$\overrightarrow{U}\cdot\overrightarrow{V}=x_1x_1'+x_2x_2'+\ldots+x_nx_n'.$$

From there, we derive that the inner product is “bilinear”, namely that

$$\begin{aligned}(k\overrightarrow{U})\cdot\overrightarrow{V}&=\overrightarrow{U}\cdot(k\overrightarrow{V})=k\,\overrightarrow{U}\cdot\overrightarrow{V},\\ (\overrightarrow{U_1}+\overrightarrow{U_2})\cdot\overrightarrow{V}&=\overrightarrow{U_1}\cdot\overrightarrow{V}+\overrightarrow{U_2}\cdot\overrightarrow{V},\\ \overrightarrow{U}\cdot(\overrightarrow{V_1}+\overrightarrow{V_1})&=\overrightarrow{U}\cdot\overrightarrow{V_1}+\overrightarrow{U}\cdot\overrightarrow{V_2}.\end{aligned}$$

if $\overrightarrow{U}$, $\overrightarrow{V}$ are two vectors, we can pick a point A and write $\overrightarrow{U}=\overrightarrow{AB}$, then $\overrightarrow{V}=\overrightarrow{BC}$, so that $\overrightarrow{U}+\overrightarrow{V}=\overrightarrow{AC}$. The triangle ABC is rectangle if and only if we have Pythagoras' relation $AC^2=AB^2+BC^2$, i.e.

$$\|\overrightarrow{U}+\overrightarrow{V}\|^2=\|\overrightarrow{U}\|^2+\|\overrightarrow{V}\|^2,$$

in other words, by (9.2), if and only if $\overrightarrow{U}\cdot\overrightarrow{V}=0$.

Consequence. *Tw vectors $\overrightarrow{U}$ and $\overrightarrow{V}$ are perpendicular if and only if $\overrightarrow{U}\cdot\overrightarrow{V}=0$.*

More generally, if we fix an origin O and a point A such that $\overrightarrow{U}=\overrightarrow{OA}$, one can also pick a coordinate system such that A belongs to the Ox axis, that is, $A=(u\,;\,0)$. For every vector $\overrightarrow{V}=\overrightarrow{OB}$ $(v\,;\,w)$ in Oxy, we then get

$$\overrightarrow{U}\cdot\overrightarrow{V}=uv$$

whereas

$$\|\overrightarrow{U}\|=u,\qquad \|\overrightarrow{V}\|=\sqrt{v^2+w^2}.$$

As the half-line $[O,B)$ intersects the trigonometric circle at point $(kv\,;\,kw)$ with $k=1/\sqrt{v^2+w^2}$, we get by definition

$$\cos(\widehat{\overrightarrow{U},\overrightarrow{V}})=\cos(\widehat{AOB})=kv=\frac{v}{\sqrt{v^2+w^2}}.$$

This leads to the very useful formulas

$$\overrightarrow{U}\cdot\overrightarrow{V}=\|\overrightarrow{U}\|\,\|\overrightarrow{V}\|\,\cos(\widehat{\overrightarrow{U},\overrightarrow{V}}),\qquad \cos(\widehat{\overrightarrow{U},\overrightarrow{V}})=\frac{\overrightarrow{U}\cdot\overrightarrow{V}}{\|\overrightarrow{U}\|\,\|\overrightarrow{V}\|}. \tag{9.4}$$

10. More advanced material

At this point, we have all the necessary foundations, and the succession of concepts to be introduced becomes much more flexible – much of what we discuss below only concerns high school level and beyond.

One can for example study further properties of triangles and circles, and gradually introduce the main geometric transformations (in the plane to start with) : translations, homotheties, affinities, axial symmetries, projections, rotations with respect to a point ; and in space, symmetries with respect to a point, a line or a plane, orthogonal projections on a plane or on a line, rotation around an axis. Available tools allow making either intrinsic geometric reasonings (with angles, distances, similarity ratios, ...), or calculations in Cartesian coordinates. It is actually desirable that these techniques remain intimately connected, as this is common practice in contemporary mathematics (the period that we describe as "contemporary" actually going back to several centuries for mathematicians, engineers, physicists ...)

It is then time to investigate the phenomenon of linearity, independently of any distance consideration. This leads to the concepts of linear combinations of vectors, linear dependence and independence, non orthonormal frames, etc, in relation with the resolution of systems of linear equations. One is quickly led to determinants 2×2, 3×3, to equations of lines, planes, etc. The general concept of vector space provides an intrinsic vision of linear algebra, and one can introduce general affine spaces, bilinear symmetric forms, Euclidean and Hermitian geometry in arbitrary dimension. What we have done before can be deepened in various ways, especially by studying the general concept of isometry.

10.1. Definition. *Let $\mathcal{E}$ and $\mathcal{F}$ be two Euclidean spaces and let $s : \mathcal{E} \to \mathcal{F}$ be an arbitrary map between these. We say that s is an isometry from $\mathcal{E}$ to $\mathcal{F}$ if for every pair of points (M, N) of $\mathcal{E}$, we have $d(s(M), s(N)) = d(M, N)$.*

Isometries are closely tied to inner product via the following fundamental theorem.

10.2. Theorem. *If $s : \mathcal{E} \to \mathcal{F}$ is an isometry, then s is an affine transformation, and its associated linear map $\sigma : \overrightarrow{\mathcal{E}} \to \overrightarrow{\mathcal{F}}$ is an orthogonal transform of Euclidean vector spaces, namely a linear map preserving orthogonality and inner products :*

$$\sigma(\overrightarrow{V}) \cdot \sigma(\overrightarrow{W}) = \overrightarrow{V} \cdot \overrightarrow{W} \tag{10.3}$$

for all vectors $\overrightarrow{V}, \overrightarrow{W} \in \overrightarrow{E}$.

In the same vein, one can prove the following result, which provides a rigorous mathematical justification to all definitions and physical considerations appeared in section 3.2.

10.4. Theorem. *Let $(A_1A_2A_3A_4 \ldots)$ and $(A'_1A'_2A'_3A'_4 \ldots)$ be two isometric figures formed by points A_i, A'_i of a Euclidean space $\mathcal{E}$. Then there exists an isometry s of the entire Euclidean space $\mathcal{E}$ such that $A'_i = s(A_i)$ for all i.*

Non Euclidean geometries.

Bernhard Riemann (1826–1866)

In contemporary mathematics, non Euclidean geometries are best seen as a special instance of Riemannian geometry, so called in reference to Bernhard Riemann, one of the founders of modern complex analysis and differential geometry [Rie]. A Riemannian manifold is by definition a differential manifold M, namely a topological space that admits local differentiable systems of coordinates $x = (x_1\,;\, x_2\,;\, \ldots\,;\, x_n)$, equipped with an infinitesimal metric g of the form

$$ds^2 = g(x) = \sum_{1 \leqslant i,j \leqslant n} a_{ij}(x)\, dx_i dx_j. \tag{10.5}$$

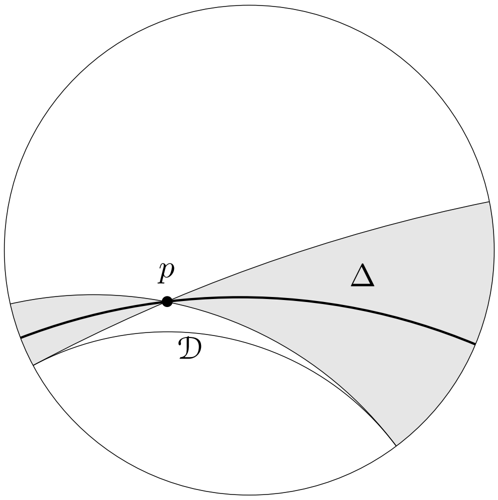
By integrating the infinitesimal metric along paths, one obtains the geodesic distance which is used as a substitute of the Euclidean distance (in physics, general relativity also arises in a similar way by considering Lorentz-like metrics of the form $ds^2 = dx_1^2 + dx_2^2 + dx_3^2 - c^2\, dt^2$). On the unit disk $\mathbb{D} = \{z \in \mathbb{C}\,;\, |z| < 1\}$ in the complex plane, denoting $z = x + iy$, one considers the so-called Poincaré metric (named after Henri Poincaré, 1854-1912, see [Poi])

$$ds = \frac{|dz|}{1 - |z|^2} \quad \Leftrightarrow \quad ds^2 = \frac{dx^2 + dy^2}{(1 - (x^2 + y^2))^2}. \tag{10.6}$$

The associated geodesic distance can be computed to be

$$d_{\mathrm{P}}(a, b) = \frac{1}{2} \ln \frac{1 + \frac{|b-a|}{|1-\overline{a}b|}}{1 - \frac{|b-a|}{|1-\overline{a}b|}}. \tag{10.7}$$

When substituting this distance to the Pythagoras/Descartes axiom, one actually obtains a non Euclidean geometry, which is a model of the hyperbolic geometry discovered by Nikolai Lobachevski (1793-1856). In this geometry, there are actually infinitely many parallel lines to a given line $\mathcal{D}$ through a given point p exterior to $\mathcal{D}$, so that Euclid's fifth postulate fails !



Lobachevski's hyperbolic geometry and the failure of Euclid's fifth postulate

On some ideas of Felix Hausdorff and Mikhail Gromov.

We first describe a few important ideas due to Felix Hausdorff (1868 - 1942), one of the founders of modern topology [Hau]. The first one is that the Lebesgue measure of $\mathbb{R}^n$ can be generalized without any reference to the vector space structure, but just by using the metric. If $(\mathcal{E}, d)$ is an arbitrary metric space, one defines the p-dimensional Hausdorff measure of a subset A of $\mathcal{E}$ as

$$(10.8)\qquad \mathcal{H}_p(A) = \lim_{\varepsilon\to 0} \mathcal{H}_{p,\varepsilon}(A), \qquad \mathcal{H}_{p,\varepsilon}(A) = \inf_{\operatorname{diam} A_i \leqslant \varepsilon} \sum_i (\operatorname{diam} A_i)^p$$

where $\mathcal{H}_{p,\varepsilon}(A)$ is the least upper bound of sums $\sum_i (\operatorname{diam} A_i)^p$ running over all countable partitions $A = \bigcup A_i$ with $\dim A_i \leqslant \varepsilon$. For $p = 1$ (resp. $p = 2$, $p = 3$) one recovers the usual concepts of length, area, volume, and the definition even works when p is not an integer (it is then extremely useful to define the dimension of fractal sets). Another important idea of Hausdorff is the existence of a natural metric structure on the set of compact subsets of a given metric space $(\mathcal{E}, d)$. If K, L are two compact subsets of $\mathcal{E}$, the *Hausdorff distance* of K and L is defined to be

$$(10.9)\qquad d_H(K, L) = \max\big\{ \max_{x\in K} \min_{y\in L} d(x, y), \max_{y\in L} \min_{x\in K} d(x, y)\big\}.$$

The study of metric structures has become today one of the most active domains in mathematics. We should mention here the work of Mikhail Gromov (Abel prize 2009) on length spaces and "moduli spaces" of Riemannian manifolds. If X and Y are two compact metric spaces, one defines their *Gromov-Hausdorff distance* $d_{\mathrm{GH}}(X, Y)$ to be the infimum of all Hausdorff distances $d_{\mathrm{H}}(f(X), f(Y))$ for all possible isometric embeddings $f : X \to \mathcal{E}$, $g : Y \to \mathcal{E}$ of X and Y in another compact metric space $\mathcal{E}$. This provides a crucial tool to study deformations and degenerations of Riemannian manifolds.

References

[**Art**] Emil Artin, *Geometric Algebra*, Interscience publishers Inc., New York, 1957, download at `http://ia700300.us.archive.org/24/items/geometricalgebra033556mbp/geometricalgebra033556mbp.pdf`

[**Des**] René Descartes, *La Géométrie,* appendix of *Discours de la méthode*, available from Project Gutenberg `http://www.gutenberg.org/files/26400/26400-pdf.pdf`

[**Die**] Jean Dieudonné, *Algèbre linéaire et géométrie élémentaire*, Hermann, 1968, 240 pages

[**Euc**] Euclid, *Euclid's elements of geometry*, English translation from Greek by Richard Fitzpatrick, `http://farside.ph.utexas.edu/euclid/Elements.pdf`

[**Gro**] Mikhail Gromov, *Structures métriques pour les variétés riemanniennes*, edited by J. Lafontaine and P. Pansu, Textes Mathématiques, no. 1, CEDIC, Paris, 1981, translated by S.M. Bates, *Metric Structures for Riemannian and Non-Riemannian Spaces*, Birkhäuser, 1999

[**Hau**] Felix Hausdorff, *Dimension und äusseres Mass*, Mathematische Annalen 79, 1-2 (1919) 157-179, and *Grundzüge der Mengenlehre*, Leipzig, 1914. viii + 476 pages, English translation *Set Theory*, American Mathematical Soc., 1957, 352 pages

[**Hil**] David Hilbert, *Grundlagen der Geometrie*, Teubner, Leipzig, 1899, English translation *The Foundations of Geometry*, by E.J. Townsend (2005), available from project Gutenberg, `http://www.gutenberg.org/files/17384/17384-pdf.pdf`

[**Lob**] Nikolai Lobachevski, *Geometrical researches on the theory of parallels*, `http://quod.lib.umich.edu/u/umhistmath/AAN2339.0001.001/1?rgn=full+text;view=pdf`

[**Pas**] Moritz Pasch, *Vorlesungen über neuere Geometrie*, Teubner, Leipzig, 1882, download at `https://ia700308.us.archive.org/32/items/vorlesungenbern00pascgoog/vorlesungenbern00pascgoog.pdf`

[**Poi**] Henri Poincaré, *uvres*, tome 2, 1916 - Fonctions fuchsiennes, et *Les géométries non euclidiennes* `http://www.univ-nancy2.fr/poincare/bhp/`

[**Pon**] Jean-Victor Poncelet, *Traité des propriétés projectives des figures*, Gauthier-Villars, Paris, 1866, download at `http://gallica.bnf.fr/ark:/12148/bpt6k5484980j`

[**Rie**] Bernhard Riemann, *Grundlagen für eine allgemeine Theorie der Funktionen einer veränderlichen complexen Grösse*, Inauguraldissertation, Göttingen, 1851, and *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*, Habilitationsschrift, 1854, Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen, 13 (1868), English translation *On the Hypotheses which lie at the Bases of Geometry*, `http://www.emis.de/classics/Riemann/`