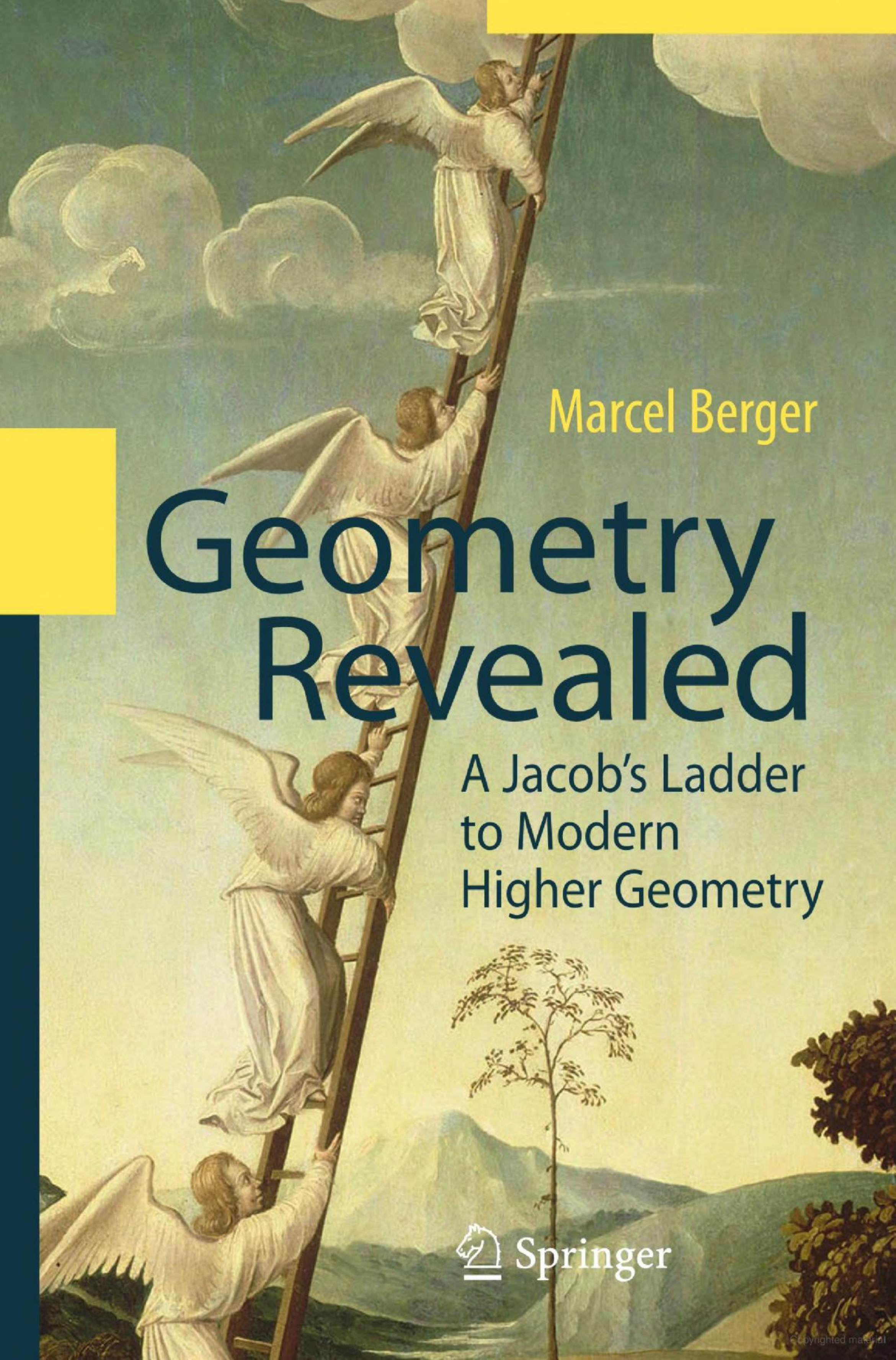


Marcel Berger

Geometry Revealed

A Jacob's Ladder to Modern Higher Geometry

Springer

MARCEL BERGER

Geometry Revealed

A Jacob's Ladder to Modern Higher Geometry

Springer

Author
Marcel Berger
Insitut des Hautes Études Scientifiques (IHES)
Bures-sur-Yvette
France

Translator
Lester J. Senechal
Professor Emeritus
Department of Mathematics
Mount Holyoke College
South Hadley, MA 10475
USA
lsenecha@mtholyoke.edu

Springer-Verlag thanks the original publishers of the figures for permission to reprint them in this book. We have made every effort to identify the copyright owners of all illustrations included in this book in order to obtain reprint permission. Some of our requests have however remained unanswered. We have inserted all sources and owners where known.

ISBN 978-3-540-70996-1 e-ISBN 978-3-540-70997-8
DOI 10.1007/978-3-540-70997-8
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010920837

Mathematics Subject Classification: 51-01, 01-01

Cover design: Wmx design

Cover illustration: Le Songe de Jacob (detail), Nicolas Dipre (D'Ypres), École d'Avignon, Musée du Petit-Palais, Avignon © bpk, Berlin, 2009

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Table of Contents

Chapter I
Points and lines in the plane

I.1. In which setting and in which plane are we working? And right away an utterly simple problem of Sylvester about the collinearity of points

We first work in the coordinate plane, which is familiar to everyone, with its *points* and *lines*. As is usual in the “elementary” geometry of school instruction, this has to do with Euclidean geometry, where there are distances (lengths), angles, circles, etc. This will also be the setting of the next chapter, but even in this first chapter we will see that we can already do many subtle and difficult things – and even find open questions – with only the so-called “affine plane”. Affine geometry is a weaker structure than Euclidean geometry. Simply put: we won’t be working with anything but points and lines; the mathematical definition is given in Sect. I.XYZ at the end of the chapter. Here we need only recall: two distinct points uniquely determine a line that contains them, along with a segment that joins them; two distinct lines intersect in a single point, with the sole exception of parallel lines. Regarding these, through each point exterior to a given line there passes a unique parallel to that line. Finally, there is a supplementary *affine notion*, more subtle than the merely set-theoretic ones of point and of line, which is the *affine invariant* attached to three collinear points: if a, b, c are collinear, there exists a real number (and one only) denoted $\frac{\overline{ab}}{\overline{ac}}$. It indicates a ratio (that can be negative, although negative numbers had been long forbidden in geometry, even by Poncelet and d’Alembert, until Chasles actually gave them rights of citizenship), a ratio obtainable by parameterizing the line considered, but which does not depend on this parameterization. We can thus speak about the midpoint of a segment, the third-way point, etc. See the necessary details in Sects. I.XYZ and I.3 below. The precise mathematical language is that of the *real affine plane*. If we adjoin a metric – which we permit ourselves occasionally, even in this chapter – we then speak of a *Euclidean plane*. An important remark about language: we can speak of “**the**”, rather than “**an**”, affine plane. For in fact any two affine planes are necessarily isomorphic, just as are two real vector spaces of dimension 2. The same remark applies to Euclidean spaces of any dimension.

But in this introductory chapter we will see that it is practically impossible to remain in the affine setting: to comprehend and unify certain things, by Sect. I.4 we will need to climb the ladder, know how “to go to infinity” and not interject the Euclidean plane but – more subtly – define the “projective plane”. The degree of subtlety can be seen historically: projective geometry wasn’t defined until Desargues in the 1650s and then only heuristically. The sound algebraic construction, following

M. Berger, *Geometry Revealed*, DOI 10.1007/978-3-540-70997-8_1,

the synthetic attempts of Poncelet and Chasles in the years 1820–1840, was made in the 1850s by the German school: Plücker, von Staudt and Grassmann, whereas Euclid dates from 300 B.C.

♦

In 1893 Sylvester posed the following problem:

(I.1.1) *Let* E *be a finite set of points in the plane that has the following property: for an arbitrary pair of distinct points of* E *there exists, on the line joining them, a third point of* E*. Show that this is impossible, with the obvious exception of the case where* E *consists of collinear points on a single line.*

Some readers may prefer the equivalent formulation:

(I.1.2) *If* E *is a finite set of points in the plane not composed of points belonging to a single line, then there exists at least one line that contains only two of its points.*

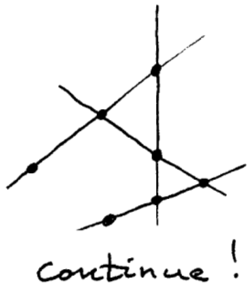


Fig. I.1.1.

As an exercise we can attempt to convince ourselves of Sylvester's conjecture by making some sketches: we quickly see that we are forced into constructing sets having an infinite number of points. But in spite of this easily won insight, it was not until 1932 that there was a proof of this conjecture, found by Gallai. We owe to Kelly in (Kelly, 1948) a proof that uses the following Euclidean argument: if the points are not all collinear, there is a triple of non-collinear points a, b, c of E forming a true triangle such that the distance from a to the line bc is minimum among all such triples. We already have a contradiction if b and c are of the same side of the altitude from a, for then the distance from b to ac, or else that of c to ab, is less than that of a to bc. So b and c must be on opposite sides of the base of the altitude. But there exists by hypothesis a third point d of E on bc, and we are led to a new contradiction by considering either the triangle abd or the triangle acd.

But this proof leaves us with a bad taste if we are at all purist: the problem is strictly affine and we should be able to prove the conjecture in a purely affine manner, without the aid of Euclidean geometry. The purely affine proof of Gallai

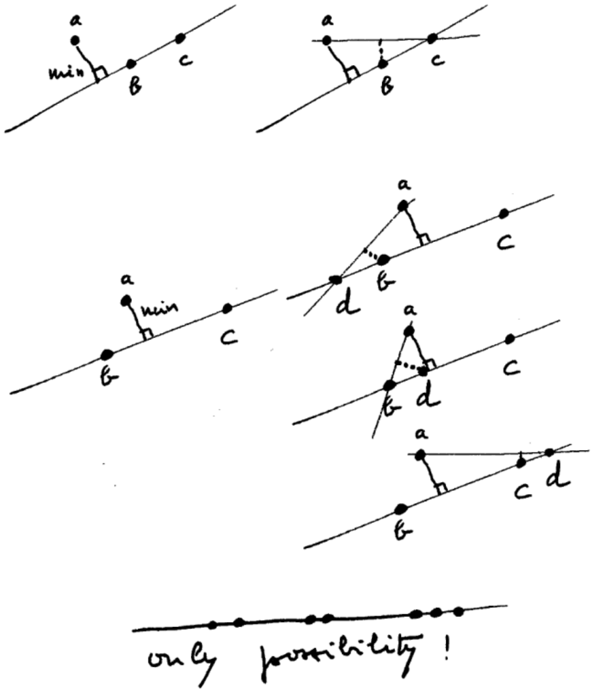


Fig. I.1.2.

is found on p. 181 of Coxeter (1989) . Courageous readers may attempt to find one of their own; but it is important to note that none of the proofs cited so far is *combinatorial* in the sense that in a combinatorial proof we compute the number of points on this or that line, how many lines joining two points of the set pass through a given point, etc., hoping to find relations that contradict the initial hypothesis. On the contrary, Gallai's proof uses the fact that a line in the plane divides it into two distinct *connected* regions; we can't pass from one to the other without intersecting the line. Apart from that, Gallai's proof doesn't introduce any new concept. Where then is Jacob's ladder? We will climb it in two different ways, but reluctant readers may skip immediately to the next section and the second problem of Sylvester.

The first way of ascending the ladder provides a conceptual and combinatorial proof of Sylvester's conjecture, due to Melchior in 1940; details can be found in Chaps. 8 and 10 of Aigner and Ziegler (1998).

We now use some tools whose motivation will be given subsequently: we extend (see Sects. I.7 and I.XYZ) the real affine plane under consideration to a real projective plane $\mathscr{P}$. There we consider not the finite set E of points satisfying Sylvester's condition, but its dual, i.e. the (necessarily finite) set of lines D dual to E under a duality of $\mathscr{P}$.

A duality consists of two mappings: the first associates with each point a of $\mathcal{P}$ a line denoted by a^*; the second associates with each line d a point denoted by d^*. The fundamental properties of a duality are the following:

- the mappings $a \to a^*$ and $d \to d^*$ are inverse to each other;
- if the line d passes through the point a, then the line a^* passes through the point d^*.

For example, corresponding to two points a and b lying on d, there are the two lines a^* and b^* intersecting in the point d^*. For a complete definition, see Sect. I.7.

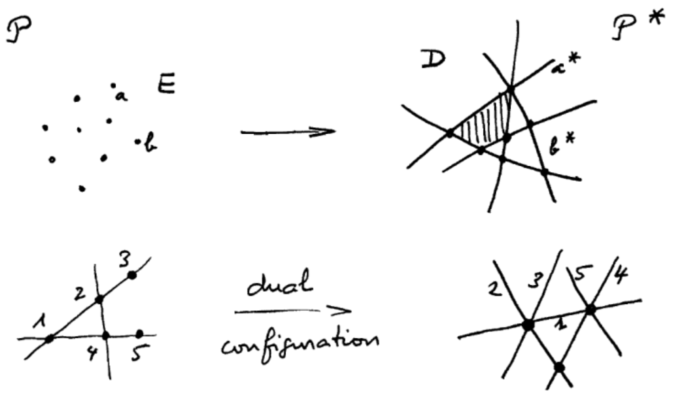


Fig. I.1.3. Duality between points and lines in the projective plane

For the configuration of points and lines provided by D we then obtain combinatorial relations between the two sequences of integers $\{p_r\}$ and $\{t_r\}$ defined as follows: p_r is the number of polygons of r sides that are found in the cellular decomposition of $\mathcal{P}$ that D defines, while t_r is the number of points that lie on r lines of D. We have the following relations, where f_0, f_1, f_2 denote the respective numbers of vertices, edges and polygons of the cellular decomposition: $f_0 = \sum t_r$, $f_2 = \sum p_r$, $f_1 = \sum r t_r = \frac{1}{2} \sum r p_r$. But algebraic topology (see the combinatorics of polyhedra in Sect. VIII.4) tells us that, for the surface $\mathcal{P}$, the Euler-Poincaré characteristic $f_0 - f_1 + f_2$ equals 1. To prove Sylvester's conjecture, we need to prove that $t_2 \geqslant 1$ (which implies that in the configuration defined by D there is one point lying on two lines and, in that defined by E, one line that contains only two points). Suppose to the contrary that we only have $t_r > 0$ when $r \geqslant 3$. The Euler-Poincaré formula yields on the one hand

$$\sum t_r + \sum p_r = 1 + \sum r t_r \geqslant 1 + 3 \sum t_r$$

and, on the other,

$$\sum t_r + \sum p_r = 1 + \frac{1}{2} \sum r p_r \geqslant 1 + \frac{3}{2} \sum p_r .$$

Upon multiplying the first relation by $\frac{1}{3}$, the second by $\frac{2}{3}$ and adding, we obtain a contradiction.

In Aigner and Ziegler (1998) or Aigner and Ziegler (2003) there is a variant of the proof by central projection, due to Steenrod, using graph theory and spherical geometry:

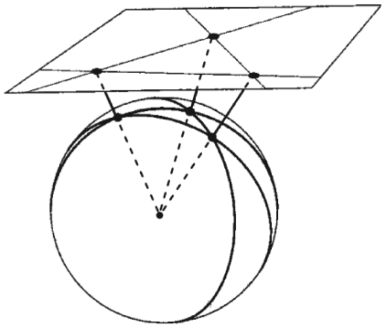


Fig. I.1.4. Aigner, Ziegler (1998) © G. M. Ziegler

♦

The second ascent entails introducing concepts of the *complex* affine plane and the *planar cubic*. We will see in Sect. V.14 that a generic cubic in the complex plane possesses nine distinct inflection points and, most importantly, that each two inflection points have the property that the line that joins them intersects the cubic once again in an inflection point. Sylvester's conjecture is thus false in the complex plane. This isn't so surprising, for we can't apply reasoning "a la Gallai" for the reason that a line in the complex plane only determines a single connected region. With some planar algebraic geometry, as in Sect. V.14, it is also easy to see that each complex planar configuration of nine points satisfying Sylvester's condition is equivalent to the one described above. However, there does exist an extension of Sylvester's result to complex affine geometry, necessarily of dimension higher than two, as will be seen in Sect. I.8, that requires a very high ascent on Jacob's ladder.

Finally, a result such as (I.1.2) will not completely satisfy a mathematical intellect, requiring as it does for the set E merely the existence of at least one line that contains only two of its points. A few sketches will convince readers that we might prove a stronger result, of a sort such as this: we will say that a line associated with a finite set of points is *ordinary* if it contains but two points of the set. We *denote* by $t(n)$ the minimum number of ordinary lines of a set E of n noncollinear points. Theorem (I.1.2) states that we always have $t(n) \geqslant 1$ for each integer n, but we might suppose that $t(n)$ may be rather large with increasing n. The question isn't yet settled. Here briefly is the present state of affairs; for more details and references see Problem F12 in Croft, Falconer and Guy (1991), Chap. 8 of Aigner and Ziegler (2003), and the second part of Pach and Agarwal (1995) – which is more conceptual – and also the Introduction, p. 679, of Vol. II of Hirzebruch (1997). The exact general value of $t(n)$ is unknown; the best we know presently is that we always have $t > [n/2]$ (*integer part of* $n/2$), which is due to Hansen, but we don't have an optimal answer. Moreover the proof of this result of Hansen doesn't at the moment seem to bring with it any new concept. Nevertheless, knowledge of the combinatorics of arrangements of lines in the real plane has recently increased considerably, see the reference Pach and Agarwal (1995). Finally, for the complex case, see Sect. I.8. For their aesthetic aspect and their naturalness, the configurations called *Sylvester-Gallai* remain much studied; see for example Bokowski and Richter-Gebert (1992).

♦

The name Erdös deserves special mention. Beyond his numerous results and his innumerable lectures, he was known first for having a rather long waiting line of researchers at the end of his lectures. Each in his turn would say: "Professor Erdös, I don't know how to settle this or that question". Almost invariably the response would be: "Here's how to do it. Write the article, we'll sign it jointly". Given then the innumerable articles written jointly with him, practically every mathematician of a certain age appears as a *connected component of Erdös* and even possesses an *Erdös number* defined thus: it's the minimum number of elements in a chain of several articles which ends with an article written jointly with Erdös. Your humble author didn't escape either; his Erdös number equals 3, via Aryeh Dvoretsky (who has seven articles jointly with Erdös – if we want to compute an *Erdös valence*) – and Eugenio Calabi. Another of Erdös's striking traits was his ease in making conjectures. For many of them he actually offered compensation (which he always paid) up to five thousand dollars, and he and his purse might thank the deity that he rarely deceived himself regarding their difficulty.

I.2. Another naive problem of Sylvester, this time on the geometric probabilities of four points

The second (still purely affine) problem of Sylvester from 1865 treats the arrangement of a quadruple of points in the affine plane: only two arrangements are possible (in the *generic* case, where three points are never collinear), either they form a *convex quadrilateral* or one of the points lies in the interior of the triangle formed by the other three. Then there is the natural question:

(I.2.1) *If four points are thrown randomly at the plane, what are the probabilities for obtaining one or the other of the possible configurations?*

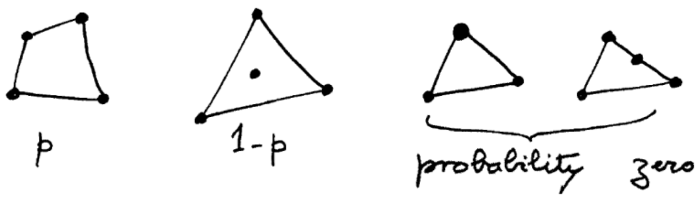


Fig. I.2.1.

There are really only two cases to consider; the degenerate ones have probability zero. But, for the question to make sense, i.e. in order to have a good notion of probability, we take as our target a planar domain D that is bounded and everywhere convex. The theoretical answer is then quite simple, the probability of obtaining four points such that one of them lies in the triangle formed by the other three is given by the triple integral

$$\text{Sylv(D)} = \frac{4}{\text{Area}^4(\text{D})} \int_{\text{D}}\int_{\text{D}}\int_{\text{D}} \text{Area}(x_1, x_2, x_3)\, dx_1 dx_2 dx_3 \qquad \text{(I.2.2)}$$

where we integrate over all triples of points of D (i.e. over all the triangles contained in D) and where Area(x_1, x_2, x_3) denotes the area of the triangle with vertices x_1, x_2, x_3. The proof is very simple: the probability that the first three points fall respectively in x_1+dx_1, x_2+dx_2, x_3+dx_3 is $dx_1 dx_2 dx_3/$ Area^3(D). Knowing this, the probability that the fourth point is in the interior of the triangle formed by the first three is Area$(x_1, x_2, x_3)/Area$(D). From this we get the formula by observing that the event considered is the union of four mutually exclusive events of equal probability.

Fig. I.2.2.

The probability of having four points that form a convex quadrilateral is then simply equal to $1 -$ Sylv(D). The value of Sylv(D) depends on the "shape" of the domain D considered; we have Sylv(D) $= \frac{1}{3}$ for an arbitrary triangle and $\frac{35}{12\pi^2}$ for an arbitrary ellipse.

These results should give us much to think about. First, the value is the same for all triangles and for all ellipses. The reason is simple: in affine geometry, all triangles are "the same", all ellipses are "the same". We will return to all this amply in Sect. I.3, where we introduce notions that permit us to clarify what we mean by "the same". It will be noted that Sylvester's condition is purely affine. We now observe that, in Euclidean geometry, we clearly no longer have such equivalences for similarly shaped domains.

The two values above show in any case that the probability of having a quadrilateral is significantly lower for triangles than for ellipses. This is intuitive enough: when we take three points, each of which is close to a vertex of the triangle, there remains but little space for the fourth point outside the new triangle thus formed. In contrast, near the boundary of a round domain, we have more space. It is important to go further, since up to this point we know nothing about other domains. This problem was settled by Blaschke in 1917:

We always have $\quad \dfrac{35}{12\pi^2} \leqslant \mathrm{Sylv(D)} \leqslant \dfrac{1}{3} \quad$ *for any domain whatsoever.*

And surely our curiosity won't be completely satisfied until we know that Blaschke also showed that equality isn't attained for the lower and upper bounds except by triangles and ellipses, respectively: a nice characterization of triangles and ellipses! See Note I.4.5 of Santalo (1976) and Klee (1969) and Sect. 5.2 of Gruber and Wills (1993). We give here the two-fold idea of Blaschke. For the left inequality we use the *Steiner symmetrization* which we will encounter several times in Chap. VII (beginning in Sect. VII.5.A), but why not make quick use of it right off? It is described on the diagram: with each convex domain D and each linear direction Δ is associated the symmetrization $\mathrm{sym}_\Delta(\mathrm{D})$ of D for the direction Δ.

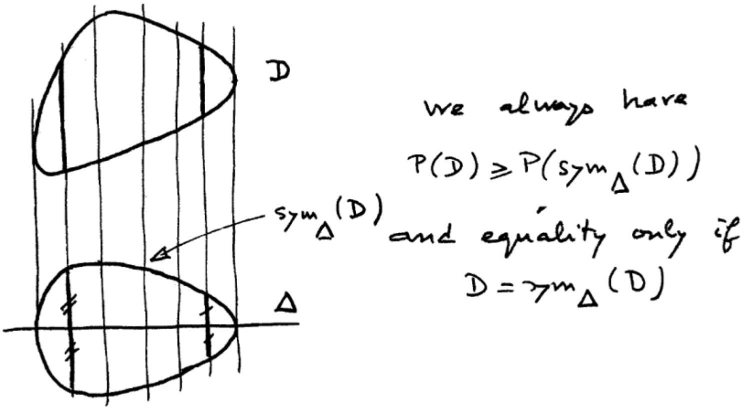


Fig. I.2.3.

Blaschke shows that each symmetrization can only diminish the integral (I.2.2). This is easy enough to perceive intuitively, for a few sketches quickly convince us that the symmetrization of a triangle interior to D often becomes a quadrilateral in $\mathrm{sym}_\Delta(\mathrm{D})$. Furthermore the diminution is strict provided that the convex set is not symmetric with respect to the direction considered. Knowing this, we effect some symmetrizations about well chosen lines (the directions alone matter), for example, by taking lines with inclinations that are irrational multiples of π.

For inequality in the reverse sense, Blaschke introduces the notion of *cosymmetrization*, which we haven't encountered anywhere else in the geometric literature.

It is easy to see that cosymmetrization, conversely, strictly increases the integral (I.2.2). We approximate D by polygons and thus obtains a reduction to the polygonal case. For the polygons for which one direction is orthogonal to a line

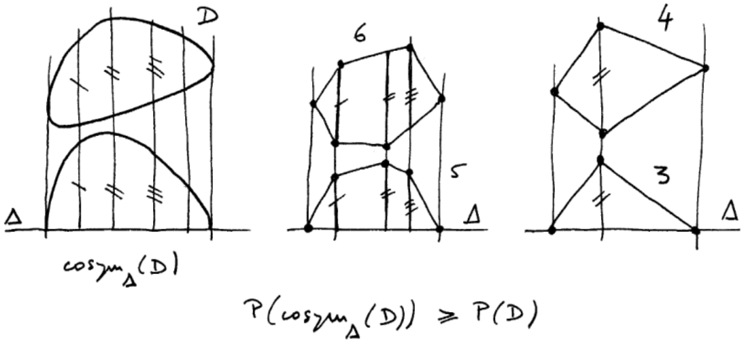


Fig. I.2.4.

joining two nonconsecutive sides, the cosymmetrization always has at least one vertex fewer than the initial polygon and we end up with a triangle, Q.E.D.

♦

Problem (I.2.1) seems well in hand, but in fact we have cheated a bit in requiring that the domain D be convex and bounded in order for the notion of probability to make sense. In truth it suffices for D to have *finite area*, which doesn't preclude "passage to infinity". Note that such a domain, extending to infinity, cannot – except for some very special cases – be convex and that Sylvester's second problem, cited frequently only for convex sets, continues to make sense for all sets of finite area. The four points may be in the domain, but the quadrilateral they determine may emerge from it, which actually needn't trouble us; it suffices to replace, in formula (I.2.2), $\text{Area}(x_1, x_2, x_3)$ by $\text{Area}(\text{Triangle}(x_1, x_2, x_3) \cap \text{D})$. This more general non compact study was undertaken only very recently and is not yet well understood. Here is what we know, a recent reference being (Scheinerman and Wilf, 1994): on the one hand, the shape that yields the lower bound $p \equiv 1 - \text{Sylv(D)}$ over all D isn't known or even conjectured precisely. On the other hand this work provides a result that is amazing at first glance: even though we don't know the exact value of the optimal probability, it is possible to show that it coincides with another number, also unknown and extensively studied in combinatorial geometry, see Pach and Agarwal (1995), for it is related to planar realizations of the complete graph K_n with n vertices (a graph is complete when every pair of vertices is joined by an edge, and a result of Fany states that we only need use segments for joining vertices). Let $\nu(\text{K}_n)$ be the minimum number of points of intersection of the edges in an arbitrary planar realization of a complete graph K_n. By putting all the points on a circle, we get C_n^4 intersections. Readers may find a smaller number with other examples, but a classical result states that n^4 is the right order of magnitude. More precisely, there exists a positive real number ν (finite) such that

$$\lim_{n\to\infty} \frac{\nu(\text{K}_n)}{\text{C}_n^4} = \nu.$$

The amazing result is that $\nu = p$.

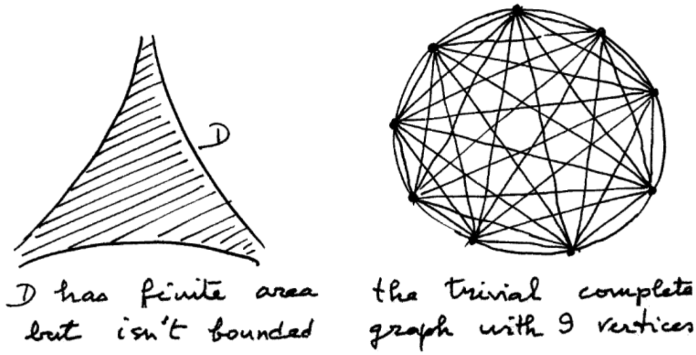


Fig. I.2.5.

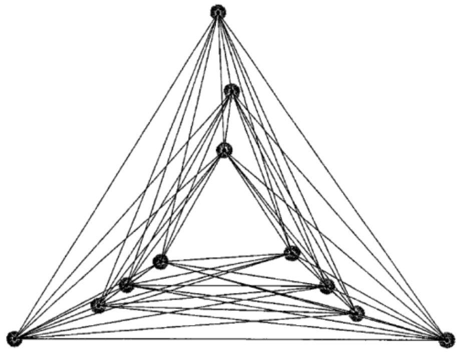


Fig. I.2.6. A very nice complete graph with 11 vertices, due to H. Jensen

For the proof, the connection between the two concepts is achieved thus: we choose n points in D at random and in a probabilistically independent manner. It is necessary to prove two opposite inequalities. In the one direction, we start with an optimal complete graph and surround each of its vertices with a small disk of radius ε. We can choose ε sufficiently small so that random points taken in this collection of disks yield another optimal graph. We then study Sylvester's probability, choosing for D the union of these disks; for ε sufficiently small we obtain the required inequality. Roughly speaking, $1 - \mathrm{Sylv(D)}$ is equal to the probability that the four points chosen at random form a convex set, i.e. the probability that among the three possible groupings of edges $ab - cd$, $ad - bc$, $ac - bd$, one of them gives rise to an intersection is $\frac{\nu(\mathrm{K}_n)}{\mathrm{C}_n^4}$. Thus we have roughly:

$$p = \min_{\mathrm{D}}(1 - \mathrm{Sylv(D)}) \leqslant \frac{\nu(\mathrm{K}_n)}{\mathrm{C}_n^4}.$$

More precisely, it is necessary to take into account the circumstance that the four points chosen may have the bad sense not to fall into four different small disks; but the asymptotic behavior of this bad case is in total of the order of $\mathrm{O}(\frac{1}{n})$ and thus goes to zero as n goes to infinity.

In the reverse direction, we start with any domain D in which we choose n points $\{p_i\}$ at random and probabilistically independently, and we assume that these n points are the vertices of a complete linear graph K_n. Then the number c of crossings of this graph is a random variable whose value is always at least $\nu(K_n)$. Moreover, consider the random variable

$$X = \sum_{a,b,c,d} \mathbf{1}_{\{p_a, p_b, p_c, p_d\}},$$

where the sum is taken over all quadruples of $\{1, \dots, n\}$ and where $\mathbf{1}_{\{\cdot,\cdot,\cdot,\cdot\}}$ is a random indicator that equals 1 if the convex envelope of $\{p_a, p_b, p_c, p_d\}$ is a convex quadrilateral, and equals 0 otherwise. Since a random graph can't have more crossings than the mean, we have $\nu(K_n) \leqslant E(X)$ for the mathematical expectation of X. The desired result is obtained by letting n go to infinity.

The optimal shape of D isn't known, as already mentioned.

♦

We haven't yet finished with the problem (I.2.1), which violates the strict rules of the game: staying in the plane, in dimension two.

(I.2.2) *We randomly throw 5 points at a bounded region* D *of three dimensional space; what is the probability that they form a true polyhedron with five vertices? And the same problem with $n+2$ points in the space of n dimensions.*

As before we compute the complementary probability to find the probability that the fifth point is in the interior of the tetrahedron formed by the four others. The formula is the strict generalization of that given above for an arbitrary dimension, which in dimension three will be:

$$\mathrm{Sylv(D)} = \frac{5}{\mathrm{Area}^5(\mathrm{D})} \int_{\mathrm{D}}\int_{\mathrm{D}}\int_{\mathrm{D}}\int_{\mathrm{D}} \mathrm{Volume}(x_1, x_2, x_3, x_4)\, dx_1 dx_2 dx_3 dx_4,$$

where $\mathrm{Volume}(x_1, x_2, x_3, x_4)$ denotes the volume of the tetrahedron with vertices x_1, x_2, x_3, x_4. Here half the problems remain open at present; we only know what happens on one side of the conceivable inequalities. First, the value is known for *ellipsoids* (here again, there is but one ellipsoid in affine geometry, in which we continue to be situated, given the nature of the problem (I.2.2)); this is due to Klingman in 1969. In each dimension d he finds for the ellipsoid $\mathcal{E}^d$ (the binomial coefficients have their usual sense when d is odd and we use the gamma function to define the necessary factorials when d is even; the gamma function provides a factor $\sqrt{\pi}$):

$$\mathrm{Sylv}(\mathcal{E}^d) = 2^{-d}\left(\mathrm{C}_{d+1}^{(d+1)/2}\right)^{d+1} \Big/ \left(\mathrm{C}_{(d+1)^2}^{(d+1)^2/2}\right).$$

In 1973, Grömer showed conversely that this value is attained only for ellipsoids. The value in question is thus a rational number when d is odd, and a rational multiple of π^{-d} when d is even. The method of proof is again the Steiner symmetrization, which is viable in all dimensions and which we will continue to encounter in

Sects. V.11 and VII.8. A recent reference is Sect. 5.2 of Gruber and Wills (1993) where there is a nice conceptual treatment.

On the other hand, for the maximum value, three problems remain open. Is it attained for tetrahedrons (in dimensions greater than three we say *simplex*)? Does it characterize the tetrahedrons? But above all, how can we calculate the above integral for tetrahedrons? Readers may find such ignorance surprising for so simple and ordinary a geometric object as the tetrahedron. In Sect. III.6 we will encounter two other unresolved problems on the volumes of tetrahedrons in the three dimensional sphere S^3. Readers may also try to see why Blaschke's cosymmetrization method doesn't work in dimension 3 or greater. We will encounter the P(D) in a remarkable way in Sect. VII.10. Many important results in this field have appeared quite recently; see a synthesis in Bárány (2008).

♦

In dimension 3 or more we will not be satisfied with only an estimate of P(D) in the case of ellipsoids. The problem is to estimate
(I.2.3),

$$\mathrm{P(D)} = \frac{n+2}{\mathrm{Area}^{n+2}(\mathrm{D})} \int_{\mathrm{D}}\int_{\mathrm{D}} \cdots \int_{\mathrm{D}} \mathrm{Volume}(x_1, x_2, \ldots, x_{n+1})\, dx_1 dx_2 \ldots dx_{n+1},$$

as a function of invariants attached to the convex set D, where we are dealing with the volume of the simplex generated by the $n+1$ points $x_1, x_2, \ldots, x_{n+1}$. We will find a partial answer in Sect. VII.10.F.

A final comment: we have just seen for the first time an interaction between geometry and probability. Historically the original problem is that of Buffon's needle; see the elementary exposition in Santalo (1976) and, for a contemporary treatment, Sect. 5.2 of Gruber and Wills (1993), already mentioned above. Recent directions in geometric research, in particular the Gromov's approach with *mm-spaces* (see Sect. I.XYZ), seem to indicate that the notion of *measure* – to which the notion of probability is equivalent – is every bit as important in geometry as that of distance, of metric. We will encounter other uses of geometric probability in Chaps. VII, XI and XII.

I.3. The essence of affine geometry and the fundamental theorem

We will attempt – as always without too much formalism – to enter further into a vision of the real affine plane. If we want to characterize affine geometry according to the philosophy of Klein at the turn of the twentieth century, it is necessary to study its automorphisms, by which we mean the bijections that map the affine plane onto itself and preserve its structure: lines, collinearity of points, intersections of lines, etc. In the modern definition given in Sect. I.XYZ these are the linear transformations combined with translations and thus the transformations that can be written, in arbitrary coordinates: $(x, y) \mapsto (ax + by + c, a'x + b'y + c')$, with the sole condition $ab' - a'b \neq 0$ for the six real numbers a, a', b, b', c, c'. Before returning to a purely geometric characterization of these automorphisms we will identify the affine invariants, that is to say the numbers, the situations, that are "respectable" and respected by all affine transformations. In any case, we must remember that affine transformations preserve lines (i.e. collinearity of points) and send parallel lines to parallel lines.

We begin with points. Two points do not give rise to any invariant since there always exists an affine transformation taking an arbitrary pair of points to another arbitrary pair; and it's the same for three points, which explains the fact noted above: all triangles in the affine plane are the same, are indistinguishable. This is furthermore plausible – although this is not a proof – because a set of three points depends on exactly $3 \times 2 = 6$ parameters and the affine transformations also depend on the 6 parameters written above: a, b, c, a', b', c'. But if the three points considered are collinear we come upon the first affine invariant: for three collinear points a, b, c the real number denoted by $\frac{\overline{ac}}{\overline{ab}}$ is a characteristic invariant, i.e. it is preserved by every affine transformation, and two collinear triples a, b, c and a', b', c' are transformable into each other if and only if the corresponding invariants are equal. This invariant may be defined thus: $\frac{\overline{ac}}{\overline{ab}}$ is the value of the (unique) coordinate of the point c on the line defined by a, b, c in a coordinate system where a is the origin and b is the point with coordinate equal to 1.

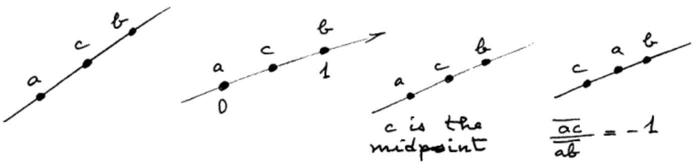


Fig. I.3.1.

As $\frac{\overline{ac}}{\overline{ab}}$ traverses the interval $[0, 1]$, the point c traverses the segment $[a, b]$ defined by a and b. The notion of *segment* is thus affine, as is that of *midpoint*: the midpoint of the segment $[a, b]$ is the point c such that $\frac{\overline{ac}}{\overline{ab}} = \frac{1}{2}$. Observe that this invariant is not Euclidean, but that if there is an additional Euclidean structure on our affine plane, then we may always compute it with the distances ab, ac (with assignment of the usual signs). Exercise: find conditions under which two sets of four (arbitrary) points can be transformed into one another.

♦

Passing now to lines, all lines are first of all the same; then, two pairs of lines are indistinguishable, under the obvious condition that they are simultaneously incident or simultaneously parallel. For three concurrent lines, there isn't any invariant: two triples of concurrent lines can always be taken into each other by an appropriate affine transformation. But it isn't the same for four concurrent lines D_i ($i = 1, 2, 3, 4$): we can attach an invariant to them in a canonical manner, i.e. their *cross ratio* $[\mathrm{D}_1, \mathrm{D}_2, \mathrm{D}_3, \mathrm{D}_4]$. This is a characteristic invariant: we can define it in an affine manner. But we won't do this, for it is in fact a *projective* invariant as will be shown in an entirely natural and simple manner in Sect. I.6.

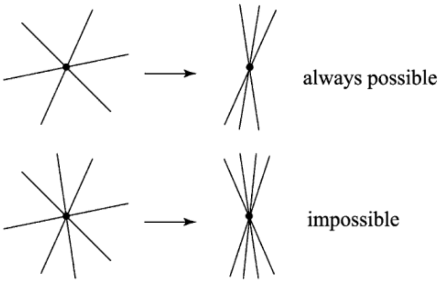


Fig. I.3.2.

Again, we can ask numerous questions on the subject of lines. Here is one of them: given two lines, three lines, or more, what is the number of possible *configurations*? For two or three, it's easy. For two: either they are concurrent or parallel. Difficulties begin with three and we encourage readers to sketch, to scribble: the lines may be concurrent or form a true triangle. But we must not forget the possibility of parallels, whence two other configurations: three parallel lines or two parallels and a third that intersects them. We see that for four and more, things become difficult; in particular we begin to get frustrated by the parallels. Here we find an additional incentive for projective geometry: parallelism doesn't exist!

♦

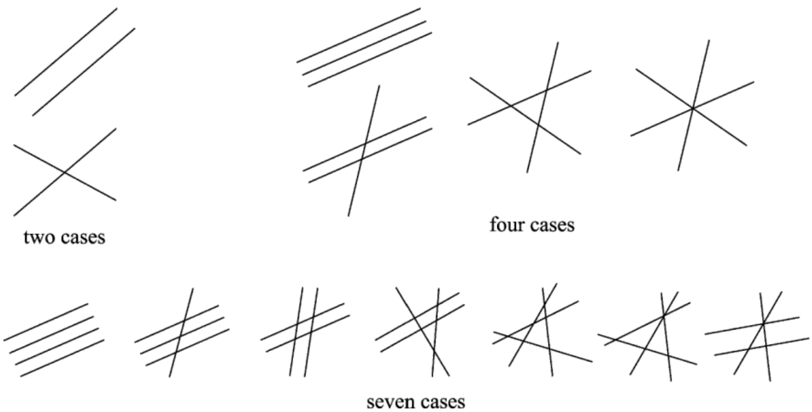


Fig. I.3.3.

The affine transformations map each line into a line, but if we want to completely capture the essence of affine geometry, e.g. by a purely axiomatic definition (without a vector space, etc.), we will want to be sure that there don't exist other transformations – beyond the affine ones defined above – of the affine plane to itself that transform each line into a line, i.e. that preserve the collinearity of points. We have a completely satisfactory answer to this question:

(I.3.1) *(Fundamental theorem of affine geometry) Each bijection of the affine plane to itself that takes lines into lines is an affine transformation.*

It is impossible to pass over the idea of the proof in silence, as much for its beauty and conceptual importance as for its allowing us to imagine what will happen in affine geometries over fields other than the reals – complex numbers, quaternions, etc. – that will be encountered in Sect. I.8. A detailed proof is found in 2.6 of [B]. We mention only this much: according to what has been said above we may suppose that our bijection f leaves three noncollinear points fixed, that we will use to define an origin and coordinates x, y; we then only need show that f is in fact the identity transformation. The fundamental remark is that parallel lines are transformed into parallels, since parallelism can be defined in a set-theoretic manner and f is bijective. Thus, in particular, parallelograms are transformed into parallelograms and it suffices to show that f acts identically on the first coordinate axis. To do this it will certainly be necessary to depart from this line, for any bijection of a line preserves that line, whether it acts identically or not. To define affine geometry we identify our line, the x axis, with the field $\mathbb{R}$ of real numbers. The figures below, based solely on parallelism, show that the restriction of f to $\mathbb{R}$ is an *automorphism*: $f(\lambda + \mu) = f(\lambda) + f(\mu)$ and $f(\lambda\mu) = f(\lambda)f(\mu)$.

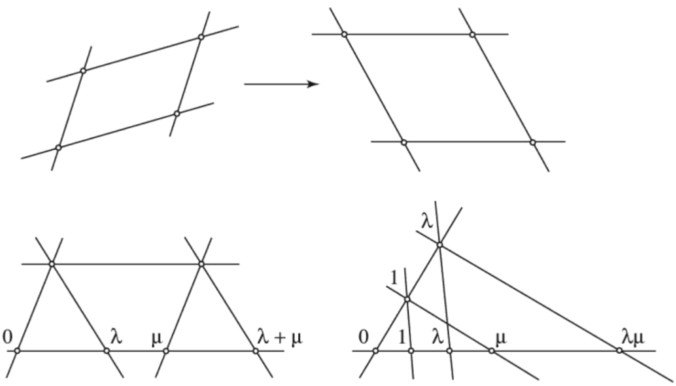


Fig. I.3.4. *Above*: a bijective mapping that preserves lines preserves parallelism. *Below*: construction of the abscissa points $\lambda + \mu$ and $\lambda\mu$ on the x axis

It is a classical exercise to show that only the identity is an automorphism of $\mathbb{R}$, but be careful not to use continuity, which has no reason to exist here; we have never required that f be continuous or even suggested that such a notion can make sense in the absence of distance!

We now pose a whole series of natural questions. First, the extension of (I.3.1) to all dimensions (> 2) is trivial; in contrast, (I.3.1) is false over the complex numbers, even adding continuity; see Sect. I.XYZ. A still more subversive question (a bit off the ladder) is to ponder the *local* and the *global*. But do we need the entire affine plane for our result? Certainly we do for the proof above, where parallelism is the key. But could we do without it? The answer is no. We will see definitively in Sect. I.5 that a bounded set in the affine plane admits plenty of other bijective transformations that preserve collinearity; these are the *projective transformations*. Thus for a deep knowledge of local affine geometry we need to climb at least one rung. In Sect. I.XYZ we will see that a good understanding of (I.3.1) in a general context and in good rapport with the axiomatics of the nineteenth century wasn't really achieved until 1950.

But back to the elaboration: what happens if we no longer require bijectivity or globality, or again if we study mappings between spaces of different dimensions? In the local but bijective case, readers will see, with the aid of passage to the infinite in the spirit of Sect. I.6, that the question is easily answered by reverting to the local affine case, but with full preservation of parallelism.

To finish our discussion of the essence of affine geometry, we pose two more questions. The first is that of incomplete duality: two distinct points determine a unique line, but in contrast two lines determine a point only if they are not parallel. Projective geometry will be the appropriate context (see Sects. I.5 and I.7) for having a *duality* without exception. A second question concerns topology: what is the topology of the set $\mathcal{D}$ of all the lines of the affine plane? What is its "shape"? The answer is that the topology of $\mathcal{D}$ is that of an open (no boundary) Möbius strip. We can convince ourselves with the sketches below. We puncture the plane at a fixed origin. With the exception of the lines that pass through the origin, the lines of the plane are associated in a one-to-one manner with the points of the punctured plane (take an auxiliary Euclidean structure and project the origin onto the line in question) and it only remains to "glue" (or sew) the punctured plane to the circle of lines that pass through the origin (caution! this is not the unit circle but is obtained by identifying antipodal points). The segments of the Möbius strip correspond to parallel lines. This operation, which consists of replacing a point by the set of lines that pass through it, is called the *blowing up* at the point; it is used in an essential way in algebraic geometry. More precisely, looking at the figure, we trace a disk about the blowing up point and replace it by a Möbius strip, while gluing the circle which bounds the disk to the circle bounding the Möbius strip. In this operation, the point is replaced by the median circle of the strip.

Analytically, the fact that the topology of the set of lines of the plane is not that of $\mathbb{R}^2$ is easily seen: it is not possible to obtain all lines with a single type of equation. For example, the two-parameter expression $y = ax + b$ allows the vertical lines with equation $x = c$ to escape. If we opt for the equation $ax + by + 1 = 0$, we lose the lines passing through the origin. We are thus forced to consider all the equations $ax + by + c = 0$; but then the triple (a, b, c) and the triple (ka, kb, kc), for $k \neq 0$, represent the same line. We are forced to pass to the quotient and to equivalence classes: this is precisely what we do in constructing projective geometry in Sect. I.5.

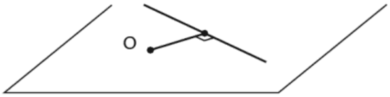


Fig. I.3.5. Correspondence between lines not passing through O and points of the punctured plane

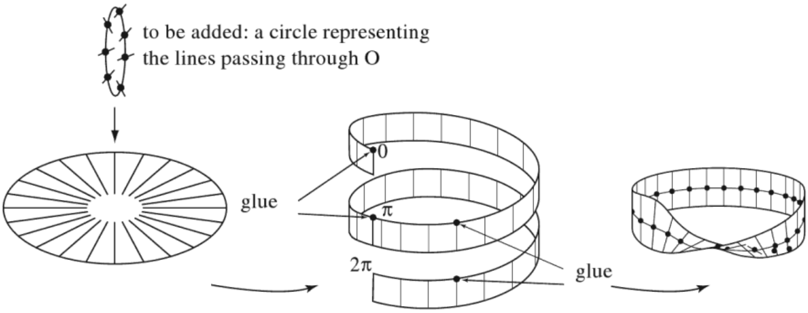


Fig. I.3.6. *Blowing up at the origin.* The half lines emanating from the origin O (and not containing the origin) are glued onto a *circle* of length π, two opposite half lines being glued to the same point of the circle. The punctured disk (or plane) thus becomes a Möbius strip

I.4. Three configurations of the affine plane and what has happened to them: Pappus, Desargues and Perles

We consider the three figures below, the first two are very old, the third dates from 1965. They seem innocent enough, but they are going to give rise, each in its turn, to very different phenomena. There are surely plenty of other plane affine configurations, but our choice has been dictated by the extensions for which the first two have given rise and the surprising consequences of the third.

Readers will be able to guess the significance of $(9_3, 9_3)$ and $(10_3, 10_3)$ and $(10_3, 10_3)$ or otherwise refer to Sect. I.XYZ or to Sect. I.9. The first configuration is that of *Pappus* (fourth century): given six points situated three apiece on each of two lines, then the three other points that can be derived from them, as indicated on the figure, are again collinear. In the second, *Desargues' theorem* (circa 1630), we have two triangles called *homological* (here abc and $a'b'c'$), which means that the lines joining corresponding vertices are concurrent. The conclusion is that the three points x, y, z indicated on the figure (the points of intersection of the homological sides) are again collinear. Finally, in the third, the conclusion is that the quotient of the affine invariants (see above) $\frac{\overline{12}}{\overline{13}}/\frac{\overline{42}}{\overline{43}}$ is forced, by the allignments drawn, to be equal to $\frac{1}{2}(3-\sqrt{5})$.

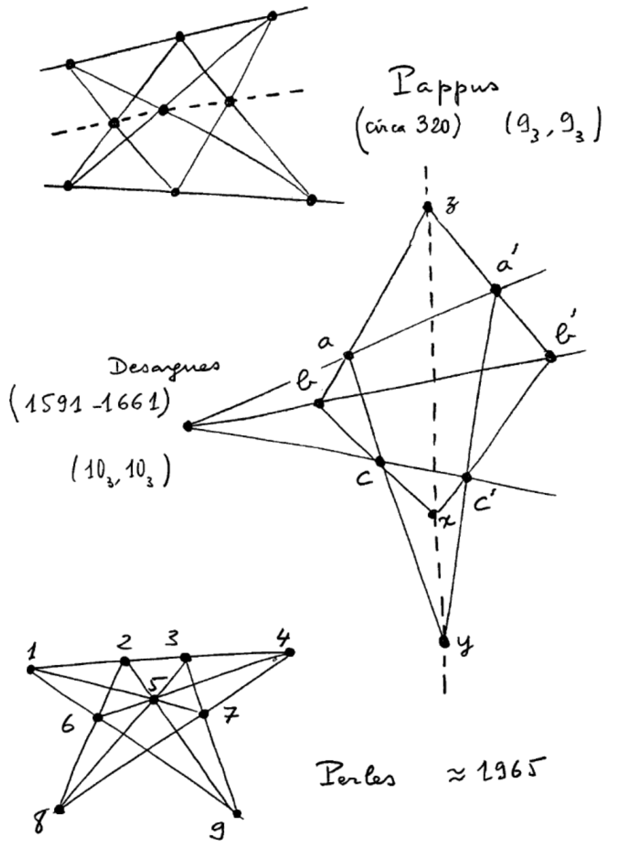


Fig. I.4.1.

♦

There are at least three things to mention regarding Pappus's theorem. The first, very briefly: when we have six points on two lines, we have a particular case of six points on a single *conic* because the pair of lines may be considered as a *degenerate* conic; see Chap. IV. In this more general case, the indicated collinearity still holds: this is the famous theorem of Pascal; see Sect. IV.2.

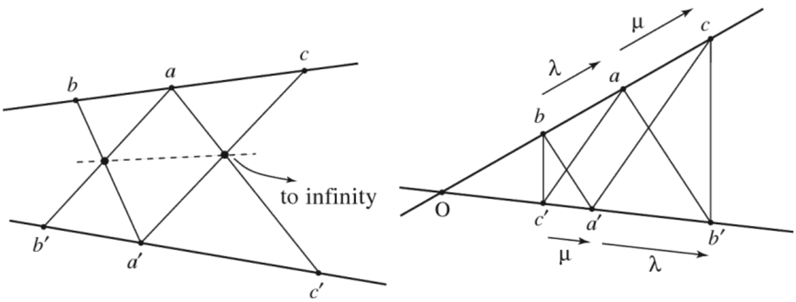


Fig. I.4.2.

We now speak about Pappus's proofs. The good proof, illuminating for the sequel, is one that uses projective geometry, considered amply in the next section. Suppose that two of the three points of intersection constructed are "*at infinity*"; see the figure. Then, as we will see, two pairs of lines that otherwise would intersect are parallel. It is required to show that the third pair is also made up of parallels. We pass from ac' to ca' by a *homothety* of ratio μ, and we pass from ba' to ab' by a homothety of ratio λ (all these homotheties have center O). Thus we pass from b to c by a homothety of ratio $\mu\lambda$ and from c' to b' by a homothety of ratio $\lambda\mu$. But since $\lambda\mu = \mu\lambda$, the proof is complete.

Alerted by what has been said in regard to the fundamental theorem of affine geometry, readers may ask what happens with the theorem for affine geometry over the other fields and thus deduce that Pappus's theorem is true for *complex* affine geometry, but not for *quaternion* affine geometry, since the quaternion field isn't commutative. It is a consequence of an axiomatic study of affine geometry that the commutativity of the underlying field can be characterized by the validity of the configurations of Pappus. All this dates from the time indicated above in Sect. I.3; see for example Artin (1957) or Baer (1952).

Recently Schwartz (1993) has given Pappus a second look. Here, very briefly, is what it's about, see the original text for more details. The starting point is this naive remark: to every pair of triples of collinear points, Pappus associates a third such triple; we then have an operation on such triples. Whence two questions: what is the algebraic nature of this operation? What happens if we iterate it a few or many times, or even indefinitely? In Schwartz (1993) these two questions are resolved and each is placed on an appropriate rung of the ladder; see also Berger (2005).

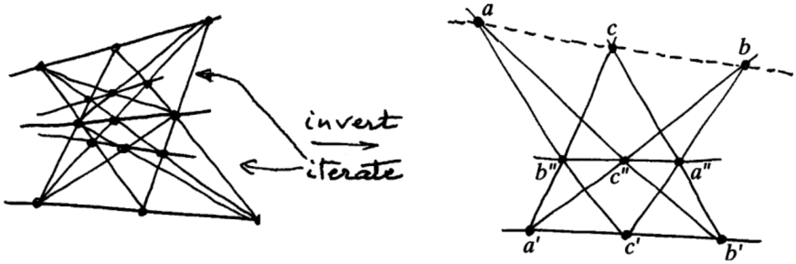


Fig. I.4.3.

The fundamental remark is that the operation "two triples produce a third" can be inverted: we can go backwards. The reversal is illustrated by the figure on the right. To study the iteration of this operation (after having composed two triples T and T′ to obtain T″, we may compose T and T″, or T′ and T″, and so forth), Schwartz introduced what he called "labeled boxes", consisting of two triples – in the box labeled $((a, b, c), (a', b', c'))$, we have that $abb'a'$ is the box and c, c' are the points labeled on the sides "above" and "below" – along with the transformations

$$\sigma : \big((a, b, c), (a', b', c')\big) \mapsto \big((a, b, c), (a'', b'', c'')\big)$$

and

$$\tau : \big((a,b,c),(a',b',c')\big) \mapsto \big((a',b',c'),(a'',b'',c'')\big).$$

It is easy to see that these two operations are related by only two conditions: $\sigma^2 =$ identity, $\tau^3 =$ identity. The group they generate is none other than the famous *modular group*, i.e. the group *denoted* $\mathrm{SL}(2,\mathbb{Z})$, defined as the group of matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer entries and determinant $ad - bc = 1$. It is interesting to encounter in connection with Pappus this group that governs a good part of mathematics and is the most important after $\mathbb{R}$ and $\mathbb{C}$. We find it in number theory, complex analysis, Riemann surfaces and algebraic geometry, i.e. for elliptic curves; see Sect. V.14. We will encounter it again in connection with polygonal billiards in Chap. XI.

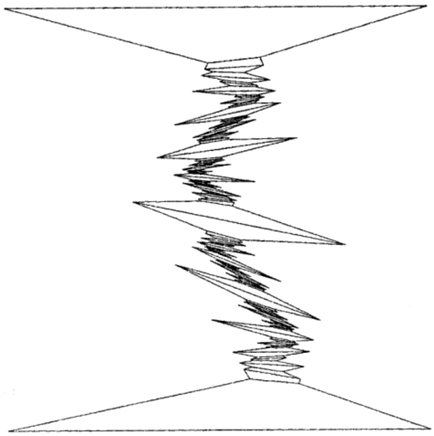


Fig. I.4.4. Schwartz (1993) © IHES

Now Schwartz has studied the figure obtained by applying the operations of this group to an initial box. It is drawn in Fig. I.4.4, to which we in fact need to add a whole complement (in order to go backwards), that turns out to be a Möbius strip (not drawn: this would be difficult). Schwartz shows that the discrete set of points marked by all the triples thus obtained can be extended by continuity to a closed continuous curve. If we start with one box that is an harmonic quadrilateral – and only in this case – all the marked points lie on a single line. In every other case the curve is *fractal*, but with an exceptional additional property: at each of its points, the line of support of the triple passing through this point intersects it in exactly one point, the one considered. This isn't the case for most fractals, where either there are plenty of lines that don't intersect the curve, e.g. the snowflake, or at the other extreme every line passing through this point intersects it amidst other points, e.g. a fractal curve that spirals. It seems that the only other known comparable example is that of the graph of Brownian motion in one dimension: at each of its points it behaves like the graph of the function $x \mapsto x^{1/2}$.

It is the moment to suggest that readers develop one or more purely affine proofs of Pappus's theorem, if only to appreciate projective geometry and in spite of the fact that they will need to climb a bit up the ladder.

♦

We can also use projective geometry for a proof of the Desargues configuration by letting two of the collinear points go to infinity. We then only need use a homothety with center O. Thus the commutativity of $\mathbb{R}$ isn't needed, but the complete calculation will show readers that we use the *associativity* of $\mathbb{R}$: $\lambda(\mu\nu) = (\lambda\mu)\nu$ for all λ, μ, ν. This is important in the axiomatic theory of affine and projective spaces: we can replace the associativity of the object which must play the role of the underlying field by the requirement that Desargues' theorem hold. Interested readers will verify by calculation that to ascertain that two nonintersecting lines are parallel in an affine plane over an arbitrary field we need to use its associativity, it being understood that a line is a set defined by an equation $ax + by + c = 0$ and that two lines are parallel if and only if they are obtainable from each other by translation.

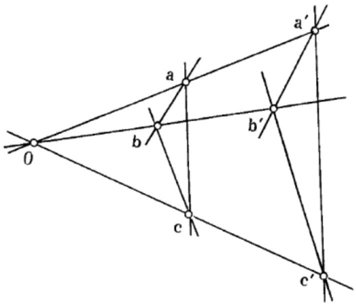


Fig. I.4.5. *Proof of Desargues theorem.* We can assume that x and y, the respective points of intersection of bc and $b'c'$ and of ac and $a'c'$ are at infinity, i.e. that bc and $b'c'$, and ac and $a'c'$, are parallel. It is then just a matter of showing that ab and $a'b'$ are parallel. But these hypotheses bring with them the existence of a homothety with center O that sends a, b, c to a', b', c', respectively. Hence the result

But there exists another proof that will subsequently appear less artificial. We embed the affine plane in the affine space of three dimensions and consider the figure obtained as the projection into dimension two of the figure below, where the three lines defining the projection between the two aren't coplanar. The result is then trivial: the three points x, y, z are collinear since they belong to the intersection of two planes, which is always a line.

The preceding explains why, in the axiomatic theory of affine or projective geometry, the situation in dimension two is completely different from the general case: affine or projective planes are hardly *categorical.* A typical example: there exists a

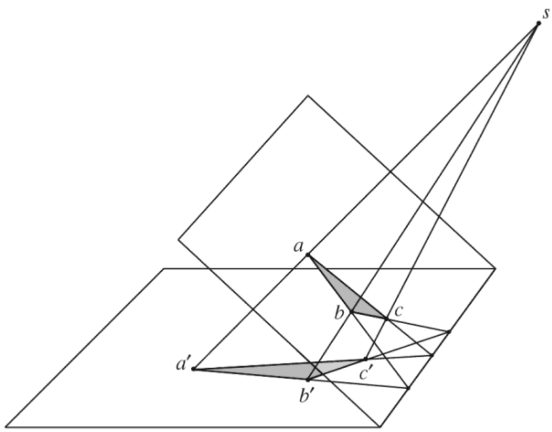


Fig. I.4.6. A figure necessarily drawn in the plane, but where we nonetheless see the perspective representation of a figure in space

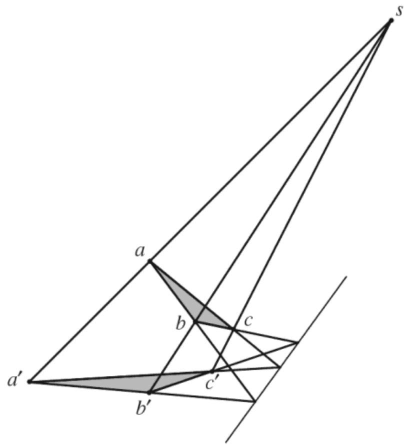


Fig. I.4.7. The same figure deprived of what allows us to see it "in space"

quasi-field, the *Cayley octonions*, denoted by $\mathbb{C}\mathbb{a}$, where there is no longer associativity; see Sect. I.XYZ. Although a projective plane, denoted by $\mathbb{C}\mathbb{a}P^2$, can be well defined over $\mathbb{C}\mathbb{a}$, we can never define $\mathbb{C}\mathbb{a}P^n$ for any $n \geqslant 3$; see Besse (1978). The reason for this is precisely that Desargues' theorem would be valid there according to the above figure; but we know that this would imply the associativity of the algebraic object, the Cayley octonions. We mention in passing that $\mathbb{C}\mathbb{a}P^2$ is for us one of the most beautiful of all geometric objects and that we could call it the *panda* of geometry. But in spite of this exceptional beauty, it is difficult to construct and extremely few authors construct it in detail; an exception can be found in 3.G of Besse (1978).

Finally, for a **dynamic** study of Desargues' configuration like that for Pappus, and by the same author, see Schwartz (1998). For another approach to iterations of geometric theorems, see Smith (2000), also cited at the end of Sect. II.1.

♦

The philosophy of Perles's example is as follows: the configuration can never be realized in the *rational* affine plane, i.e. the subset of the affine plane made up of all points whose two coordinates are rational numbers in a given coordinate system (modulo which we always have isomorphic objects); the reason is simply that $\sqrt{5}$ is irrational.

The existence of irrational affine configurations was known before Perles, see for example the notion of *accessible point* on p.126 of Coxeter (1964). For computer enthusiasts this means that such configurations are not, *in an exact sense*, visible on the screen. On the other hand we can inject the irrationals *in a formal way*, especially a number such as $\sqrt{5}$, which can be defined for example by the equation $x^2 - 5 = 0$. But the precise Perles configuration has a much deeper interest: it allowed him to show the existence of *polytopes* in dimension 8, that can never be realized with the same combinatoric and with vertices having rational (or, equivalently, integer) coordinates. We will return to this question amply in Sect. VIII.12.

I.5. The irresistible necessity of projective geometry and the construction of the projective plane

We have had reason to be unhappy on several occasions above: first, while Pappus's theorem – like Desargues' in the purely affine context – presents several variants because of possibilities of parallelism. We have an even simpler question, encountered at the end of Sect. I.3: into how many regions do two, three, four, etc. lines divide the plane?

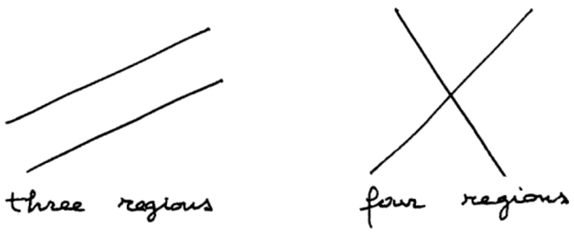


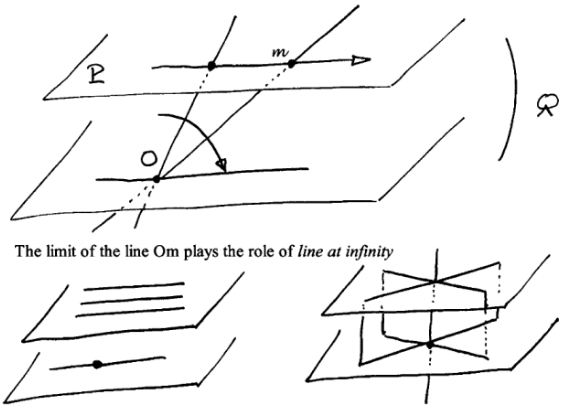
Fig. I.5.1.

Even though its formal definition in algebraic language may seem unproblematic, it requires a bit of time to begin to feel at ease with projective geometry and we thus beg readers to be patient and not to become discouraged. As further evidence of this difficulty it should suffice to remark that, even though introduced by Desargues at the beginning of the seventeenth century, projective geometry wasn't firmly established until the second half of the nineteenth century. Desargues' naive definition is as follows: the projective plane P^* associated with the affine plane P extending P is nothing other than P itself to which a line P_∞ of points at infinity is adjoined, the elements of P_∞ (the line "at infinity" of P) being the set of directions of lines of P: $P^* = P \cup P_\infty$. We then say that two distinct parallel lines intersect precisely at the point at infinity that corresponds to their common direction. As for a line D of P and the line at infinity, they intersect precisely at the point of P_∞ corresponding to the direction of D. Finally, for lines joining two distinct points of P^*: if one is in P and the other in P_∞, the line joining them is the one that passes through the first point with the direction given by the second; the line joining two points at infinity is the line at infinity. Thus for two lines in P^* – just as for two points – we can make existence statements without exception, without fear of parallelism.

Fig. I.5.2.

But this construction is abstract. It demands an act of faith and furthermore doesn't give us a basis for calculation, for which coordinates are needed. For finding a concrete geometric construction of P^*, we are inspired by the proof of Desargues's theorem obtained by embedding P in a space Q of three dimensions and taking an arbitrary point O of Q not in P. We have climbed a rung! With each point of P is associated a unique line of Q which passes through O. We will call lines through O "O-lines" for short and O the "origin" of Q; an O-plane of Q is likewise a plane through the origin. Among the lines passing through O precisely those are missing that are parallel to the plane P of Q; but we see that these are associated in a biunique fashion with the directions of the lines of P. We only need add that through a point at infinity corresponding to a direction of a set of parallel lines of P there is an O-line that has that direction. We thus define P^* concretely as the set of **all** the lines passing through the origin of Q. The lines of P^* will be the **planes** (always passing through the origin) of Q. The intersection axioms are now evident: two distinct O-lines uniquely determine an O-plane of Q; thus a line of P^* and two distinct O-planes Q intersect in a well determined O-line of Q. So now we no longer have any exception or particular case, just as we have wanted.

Even though it isn't really necessary (and has no significance for projective geometry over an arbitrary field), we can make this construction of P^* still more



The limit of the line Om plays the role of *line at infinity*

Parallel lines have the same point at infinity

The intersection of lines of P correspond to the intersection of planes of Q

Fig. I.5.3.

plausible as follows: if a point m regresses to infinity along a line D of P, the line OD tends toward the directed line parallel to D.

Historically this construction of projective space simply reflects the need that painters have for representing a portion of space in a picture. The point O above is nothing other than the eye of the painter (the observation point) and the plane P the picture (the picture plane). The "empirical" rules for geometric constructions employed in the arts are consequences of projective geometry.

We can now **calculate** in P* since we have the vectorial calculus in Q at our disposal: the points of P* are none other than those of Q **within multiplication by a scalar.** Let us quickly see how things work. For an arbitrary coordinate system (x, y, z) in Q, the points of P* will thus be triples of reals, not all zero, modulo an equivalence relation: the triple (x, y, z) is equivalent to the triple (kx, ky, kz) for all nonzero real k, a triple of *homogeneous coordinates* for the same point. Most important is the case where the coordinate system is such that the plane P is defined in Q by the equation $z = 1$. Then the points of P have for homogeneous coordinates the triples (x, y, z) with $z \neq 0$: the point (x, y) of P will have homogeneous coordinates $(x, y, 1)$ and all associated triples. Conversely, the triple (x, y, z) associated with $(x/z, y/z, 1)$ will be a triple of homogeneous coordinates of the point $(x/z, y/z)$ of P. Thus the points of the line of the equation $ax + by + c = 0$ satisfy, in homogeneous coordinates, the equation $ax + by + cz = 0$. The passage from the first equation to the second is called *homogenization.*

In contrast, the points at infinity are those of type $(x, y, 0)$, and the point at infinity of a line satisfies the homogeneous equation of this line.

It is convenient to use the notation $(x : y : z)$ to represent the set of all triples of homogeneous coordinates that can be obtained from the triple (x, y, z) by scalar multiplication. We then have $(x : y : z) = (x' : y' : z')$ if and only if there is a nonzero scalar k such that $x' = kx$, $y' = ky$, $z' = kz$, i.e. $(x : y : z)$ and $(x' : y' : z')$ represent the same point of P*.

As an example of significance for us in the spirit of Sect. I.1, see the expression of Hesse's configuration in homogeneous coordinates in Sect. I.8. On the other hand, what we **see** globally is the projective plane, a quite different story that we will touch on later. The human mind doesn't like objects obtained through an equivalence relation that can't be embedded in any ordinary space.

♦

This introduction of projective spaces may seem a bit artificial, but is in fact an essential tool for many problems where we have to consider things "within a scalar". We will see examples of this in II.6 and IV.7 for the space of all circles, or that of all spheres or of all conics.

An additional property of projective spaces is that they are **compact**, which is essential for certain problems; they are truly "round" (there are no longer points at infinity, they have been tamed): everything is "at a finite distance".

♦

To respond to a whole array of natural questions we now need to study projective geometry (planar here, but see Sect. I.XYZ) from the points of view of geometry, algebra (group of transformations) and topology (topology of the projective plane). This study must be done for the structure itself, initially independent of its being an **extension** of affine geometry. But of course we will want to know subsequently how to *return* to the affine plane. A (**the**) projective plane $\mathscr{P}$ is defined a priori as the set of vectorial lines (one-dimensional subspaces) of a (the) real vector space **P** of dimension 3, the lines of this projective plane being the vectorial planes (two-dimensional subspaces). For the algebraist this will be the quotient of $\mathbf{P}\setminus 0$ modulo the equivalence: $v \equiv v'$ if there exists a real k such that $v' = kv$.

What are the good transformations of $\mathscr{P}$? In the spirit of (I.3.1) it is now easy for us to find biunique transformations of an affine plane that preserve lines, but only locally: simply consider the figure below and the projection starting at the origin of the space of three dimensions Q, where we have embedded two copies P and P′ of the affine plane.

We shouldn't fail to mention that we have a injective transformation from all of P onto all of P′, with just one line of P and one line of P′ removed. Note that, in the projective coordinates of P and P′ obtained starting with systems of ordinary coordinates (x, y, z) and (x', y', z') in Q in which P and P′ are respectively the planes $z = 1$ and $z' = 1$, these transformations are expressed in a linear fashion. These are the transformations we need to apply if we want to assemble different aerial photos in order to compose a single map. Whence the following definition: the projective transformations of $\mathscr{P}$ are the (invertible) linear mappings of Q applied to (vectorial) lines. So much for geometry, but for the algebraist we consider linear transformations within a scalar. For example, in a coordinate system, we deal with all 3×3 invertible matrices modulo the multiplication of all their terms by a single nonzero scalar. More conceptually: the group of projective transformations of Q is the quotient of the linear group of Q by nonzero multiples of the identity.

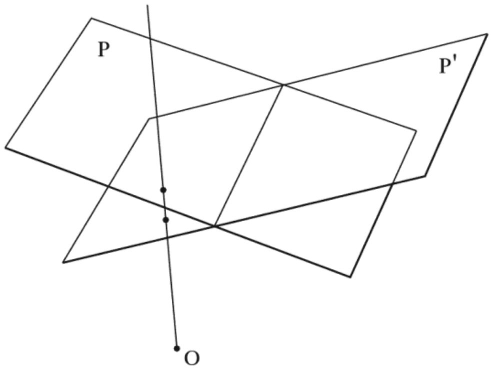


Fig. I.5.4.

In homogeneous coordinates a projective transformation will always have the form:

$$x' = ax + by + cz, \quad y' = dx + ey + fz, \quad z' = gx + hy + iz$$

or else, in affine coordinates:

$$x' = \frac{ax + by + c}{gx + hy + i}, \quad y' = \frac{dx + ey + f}{gx + hy + i}.$$

From this we deduce many things, in particular the important possibility of finding, for each quadruple of noncollinear points, coordinates written as

$$(1, 0, 0), \quad (0, 1, 0), \quad (0, 0, 1), \quad (1, 1, 1),$$

which is called a *projective frame*.

For the transformations of a projective line, see the following section. The preceding shows that a perspective (i.e. a central projection) of one line onto another is a homography (defined on the next page) and (see below) preserves the cross ratio: see Fig. 1.6.2.

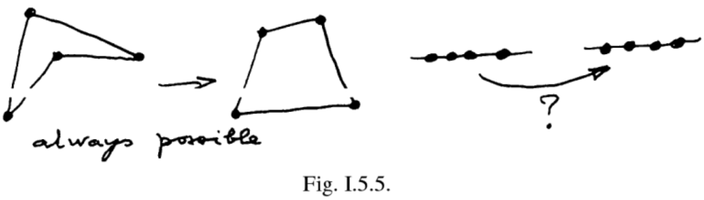


Fig. I.5.5.

♦

Now in the spirit of Sect. I.3 it's a rather easy exercise to show that, given two quadruples of non collinear points of $\mathscr{P}$, there exists a unique projective transformation taking one into the other. In the axiomatic theories this result is difficult, but essential; it is thus called the *second fundamental theorem of projective* *geometry*; see, in addition to Artin (1957) and Baer (1952), the classic (Veblen and Young, 1910–1918).

For us, in the vectorial context the result proceeds from the following fact (left to readers): for each quadruple a, b, c, d of noncollinear points, we can find a system of homogeneous coordinates such that $a = (1, 0, 0)$, $b = (0, 1, 0)$, $c = (0, 0, 1)$, $d = (1, 1, 1)$. The theorem then follows at once.

But now we have to answer the question for four collinear points. For collinear triples, we have of course transitivity. But for four, by definition of the projective plane, we need to know what happens for four vectorial lines of a vectorial plane, a question that we left open in Sect. I.3. In fact this opens an abyss under our very feet: we have completely forgotten to speak of the **projective line**! Otherwise expressed: what are the lines of $\mathscr{P}$? What is their geometry, assuming they have one?

I.6. Intermezzo: the projective line and the cross ratio

A (**the**) projective line is thus the set of lines of a vectorial plane, a set that we will denote by $\mathbb{R}\mathrm{P}^1$, in agreement with Sect. I.XYZ. The topologist is quickly satisfied here; the figure below shows that this set is in bijection with a (**the**) circle. This isn't astonishing, the construction of the projective line consists of completing the affine line by appending a single point ∞ at infinity; everything then closes up in a circle. It is well to emphasize that for the line there aren't two points at infinity, but one only. As in the affine plane it matters little in which sense we pass to infinity; we end up with the same point. This compactification of the line into a circle by a point at infinity is a particular case of a more general construction. Readers should be aware that there exist other types of compactification; we encounter some of these in Sect. II.3.

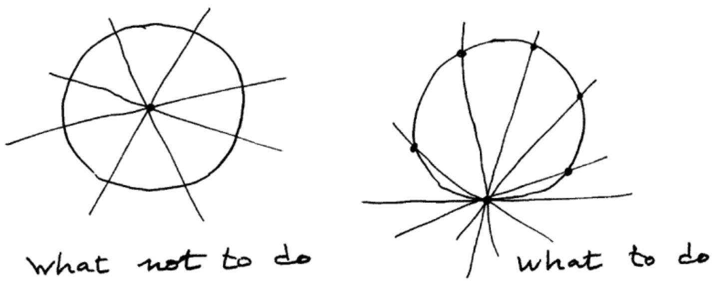


Fig. I.6.1.

♦

Algebraically, a projective line D, in particular $\mathbb{R}\mathrm{P}^1$, can always be written as the set of pairs (x, y) of real numbers, not all zero, within equivalence: (kx, ky) is equivalent to (x, y) for each nonzero real k. As in the case of triples, we can use the notation $(x : y)$ to designate the pairs taken within multiplication by a scalar.

We recover the affine line as the set of pairs where y is nonzero: $(x, y) \equiv (x/y, 1)$, which provides an embedding $t \mapsto (t, 1)$ of the affine line into the projective line. The projective transformations, called *homographies* of the projective line, are the mappings $(x, y) \mapsto (ax + by, cx + dy)$. Interpreted for the affine line, these are the mappings $t \mapsto \frac{at+b}{ct+d}$ which thus extend onto the projective line by $-\frac{d}{c} \mapsto \infty$ and $\infty \mapsto \frac{a}{c}$, consistent with the notion of limit, to comfort us once more if need be. The projective group of the line – the group of projective transformations, homographies – has three parameters, permitting us to uniquely map each triple of points into a given triple. It clearly does not preserve, when restricted to the affine line, the invariant $\frac{\overline{ac}}{\overline{ab}}$ encountered in Sect. I.3. On the other hand, there does exist an invariant for four (distinct) points $\{m_i\}$ $i = 1, 2, 3, 4$, called the *cross ratio* of these four points and *denoted* $[m_i] = [m_1, m_2, m_3, m_4]$. Two quadruples of points are in projective correspondence if and only if their cross ratios are equal. In an arbitrary coordinate system, for points $m_i = (t_i, 1)$, this cross ratio equals: $\frac{t_3-t_1}{t_3-t_2} / \frac{t_4-t_1}{t_4-t_2}$. After a moment's reflection its value is no longer surprising, it being the quotient of two affine invariants associated in a natural way with the quadruple considered. On a projective line, likewise for four distinct points, it is necessary that the cross ratio accept the value ∞, e.g. for all x we have $x = [0, \infty, x, 1]$. We note that the fact that the mapping $m \mapsto [a, b, m, c,]$ establishes a bijection between a projective line and $\mathbb{R}\mathrm{P}^1$ is equivalent to the fact that we are able to take (a, b, c) as the "projective frame", and is important for our being able to speak of "harmonic conjugation". Note that the cross ratio can be defined on an affine line; its invariance carries over by the fact that it is preserved by the point projection of the figure below, which furthermore allows it to be calculated for a quadruple of concurrent lines in the affine plane:

$$[[ap], [aq], [ar], [as]] = [p, q, r, s]_{\mathrm{D}};$$

see Fig. I.6.2. We have the interpretation: the lines passing through the point a form a projective line, which answers finally the question posed in Sect. I.3.

The above formula plays an essential role in certain geometric constructions. The case where the cross ratio equals -1 is particularly important; we say then that

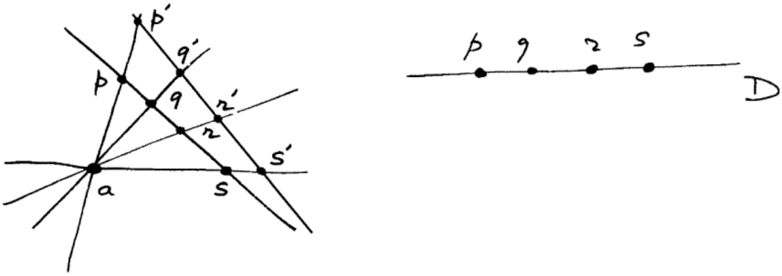


Fig. I.6.2.

the four points are in *harmonic division*. A purely geometric construction is given in Sect. I.7; note its systematic usage in Sect. IV.4.

The cross ratio is not invariant when we permute the points considered, but its behavior is simple and most interesting; see 6.3 of [B] for a detailed study. Direct calculation shows that $[b, a, c, d] = [a, b, c, d]^{-1}$ and $[a, b, c, d] + [a, c, b, d] = 1$, which allows us to calculate what happens for all the other permutations. But keep in mind for later (see Sect. V.14) that the simplest cross ratio λ which is invariant for all permutations of the four points is $\frac{(\lambda^2-\lambda+1)^3}{\lambda^2(\lambda-1)^2}$. We find in pp. 43–51 (Darboux, 1917) the calculation providing this invariance of λ for the four roots of an equation of fourth degree, as a function of the coefficients of this equation.

♦

The real projective line and its group of transformations is not an object that has been artificially concocted by geometers for their exclusive enjoyment. First of all, the "homographic" functions $t \mapsto \frac{at+b}{ct+d}$ are encountered everywhere; they are quotients of affine functions and are very important in the complex case. An important physical application of the notion of projective line is found in the theory of centered systems in optics. Lenses, mirrors, etc., are arranged in some way on a line that is their common axis; zoom lenses of the most sophisticated variety are of this type. Then the correspondence between a point of the axis and its image is always a homography. To convince ourselves of this it suffices to study the case of a mirror or of a single lens; we succeed since the homographies form a group. Readers will surely remember the following formula from school:

$$\frac{1}{x} + \frac{1}{x'} = \frac{1}{f},$$

where x is the abscissa of the object, x' that of its image, and f the focal length, positive or negative. Note that in optics infinity is essential; here it provides the *focal point* (image of the point at infinity) and the *focal objective* (reciprocal image of the point at infinity).

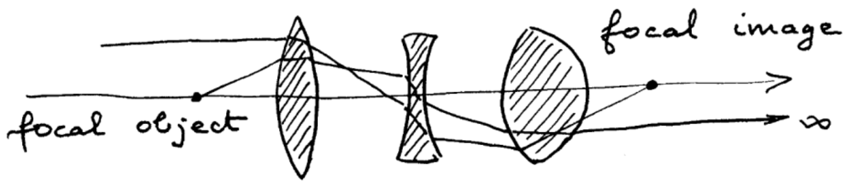


Fig. I.6.3.

The homographies, real (as here) or complex, are of primary importance in geometry; we will see this very soon in Sects. II.3 and II.4. Their classification – by their fixed points among other ways, especially as *involutions*, i.e. the homographies whose square equals the identity – is fundamental, but we will have but little to do with it (at the end of Sect. IV.6); see Chap. 6 of [B]. Finally, note that the complex projective line, in its role as a topological object, is nothing other than the sphere S^2, for it is obtained by appending a point to the complex line $\mathbb{C}$ (which is the real plane $\mathbb{R}^2$): $\mathbb{C}^* = C \cup \infty = S^2$; we will see this again in II.4.

I.7. Return to the projective plane: continuation and conclusion

We haven't yet finished with the projective plane. We first note that the cross ratio allows us to recognize when two quintuples of points are projectively equivalent; compare with Sect. I.3 for the affine case and its invariant. We now study in depth the relation between affine geometry and projective geometry, if only to make rigorous the proofs of the theorems of Pappus and Desargues that were outlined in Sect. I.3.

Starting with P, we constructed P^*, which contains the line at infinity. The essential things is that in $P^* = P \cup P_\infty$ and above all in any projective plane $\mathscr{P}$ whatever we can choose a line D and *make it* the line at infinity of the complement $\mathscr{P} \backslash D$ of D in $\mathscr{P}$. That is to say, in the construction of Fig. I.5.3, we replace the plane $z = 0$ by the plane defined by the origin and the desired line taken in the plane $z = 1$. The affine space so defined is "the" plane parallel to this new plane. For example the affine invariant $\frac{\overline{ac}}{\overline{ab}}$ of three collinear points on a line F equals the cross ratio $[c, b, a, \infty_F]$, where ∞_F denotes the point at infinity of the line F, i.e. $\infty_F = D \cap F$. To say for instance that c is the midpoint of ab is equivalent to saying that c, b, a, ∞_F form a harmonic division. In summary, we can accomplish this: in an arbitrary affine plane, completed to form a projective plane, we can alter the line at infinity, i.e. stay in the projective plane with all its advantages, all its properties, but decide on "*new parallels*". We can also speak of the "*transfer to infinity of one or more collinear points*": we let the new line at infinity pass through these points (or we can use a projective transformation sending the given line to the line at infinity). All this is certainly a rung up Jacob's ladder, where we can manage things a bit better. Fundamental is the fact that the cross ratio is conserved under these "transfers".

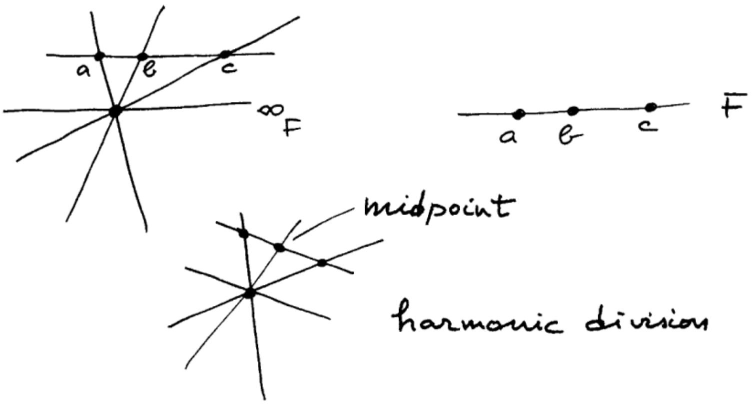


Fig. I.7.1.

The preceding technique was used to prove Pappus and Desargues in Sect. I.4. We now use it to demonstrate the classical property of the configuration of the complete quadrilateral:

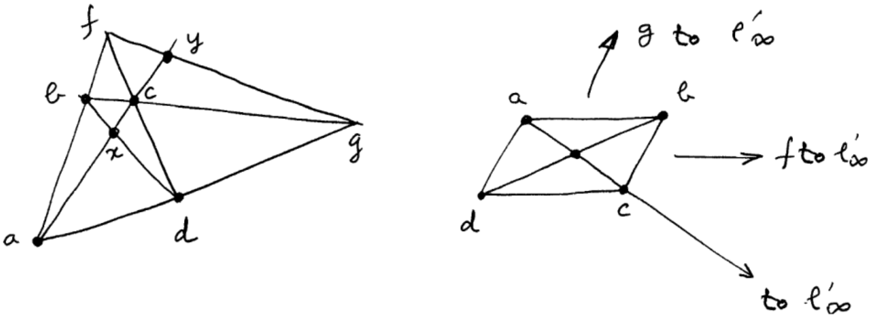


Fig. I.7.2.

In this figure the four points a, c, x, y are in harmonic division. To see this, it suffices to transfer the two points f and g to infinity. Then a, b, c, d becomes a parallelogram and our result simply translates the fact that the diagonals of a parallelogram intersect at their midpoints. Despite its simplicity, the configuration $(6_2, 4_3)$ of the complete quadrilateral may be seen as a geometric rendering of the fact that the solution of an equation of fourth degree may be reduced to that of a third degree equation. Indeed, this configuration associates in a canonical way a triple of points with a quadruple (find this in the figure). See Sect. I.XYZ for an entirely projective proof.

♦

We now attack the question of *duality*, used in Sect. I.1 and imperfect in the affine context: there points and lines played similar, but not identical, roles. Furthermore, the space of all lines had a topology different from that of the points (the affine plane), among other reasons because we could not find a good one-to-one correspondence between these two sets (see Sect. I.3).

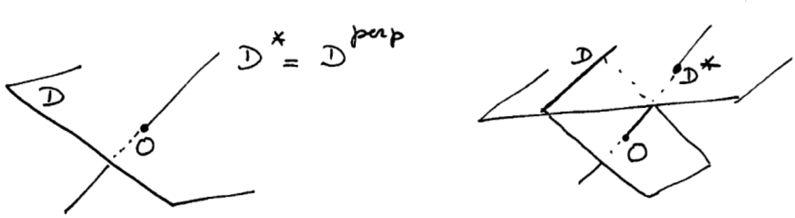


Fig. I.7.3.

In $\mathcal{P}$ the duality is perfect with regard to the line joining two points and to the intersection of two lines. However, we would like to obtain a one-to-one correspondence between $\mathcal{P}$ and the set $\mathcal{D}$ of all its lines: but this is utterly simple, since $\mathcal{P}$ is the set of lines through the origin of a vector space $\mathbf{Q}$ of dimension three and $\mathcal{D}$

the set of vectorial planes: we only need put a Euclidean structure on Q. With the directed line D we associate the perpendicular plane denoted by D^{perp}. There is a single defect: this bijection depends on the Euclidean structure chosen. The algebraist might prefer an alternative, but equivalent, presentation: let us choose some representation in projective coordinates, i.e. a projective frame of Q. Then the desired bijection consists of associating with the point (a, b, c) the line with equation $ax + by + cz = 0$. As for "modern" algebraists, they will observe that if $\mathcal{P}$ is the projective space associated with Q, then $\mathcal{D}$ is identified naturally with the vector space Q^* of Q. But, just as there doesn't exist a natural isomorphism between Q and Q^*, there doesn't exist a natural canonical isomorphism between $\mathcal{P}$ and $\mathcal{D}$.

For more on geometric dualities, the *correlations*, see 14.8.12 of [B] or p.260 of Frenkel (1973) for the general case, and Sect. I.8 below and Sect. IV.4 for the very particular case of Möbius tetrahedra. Duality will be unavoidable in a large part of Chap. VII. Furthermore this duality is completely geometric: given two points, the point of intersection of their two image lines has for an image precisely the line that joins the initial two points. This allows us to systematically obtain twice as many theorems, or to relate a desired theorem to another, perhaps simpler, theorem. In the sequel we will encounter examples in various contexts; see Sects. IV.4 and VIII.8 (conics, Pascal and Brianchon, inscribability of polyhedra). Right away readers can look for the duals of the theorems of Pappus and Desargues (see Sect. IV.4 as needed). Note that the mapping $D \to D^*$ of the left part of Fig. I.7.3 is imperfect: the origin doesn't have a dual; it is in fact the line at infinity.

♦

Attentive readers will not have missed noticing that, even though $\mathcal{P}$ and $\mathcal{D}$ are now in good bijection and have the same topology, this doesn't at all divulge the nature of the topology of $\mathcal{P}$. The first thing to observe is that $\mathcal{P}$ is **compact**, and this is also true for all the more general projective geometries of Sect. I.XYZ. In fact, if $\mathcal{P}$ is the set of points of $Q \setminus 0$ considered within multiplication by a nonzero scalar, this is also the set of points of the unit sphere of Q (a Euclidean structure is chosen for Q) modulo multiplication by ± 1, in other words the set obtained by identifying antipodal points of this sphere.

Nonetheless, the "shape" of $\mathcal{P}$ is not simple, and for an essential reason: if $\mathcal{P}$ is clearly a surface, moreover compact, it can nonetheless never be realized as a surface that is embedded in three dimensional space, this since $\mathcal{P}$ isn't *orientable*. The trick that was used for the projective line in Sect. I.6 doesn't work anymore. For

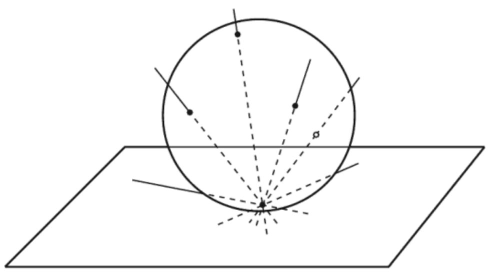


Fig. I.7.4.

on the oriented sphere, which must replace the oriented circle (compare Fig. I.6.1 with Fig. I.7.4), it is necessary to append all the points at infinity, and not just a single point; and in order to do that, cause the intervention of a "blowing up" (see Sect. I.3 and Fig. I.3.6).

A better way of understanding the topology of $\mathcal{P}$ is to see that not only can we obtain $\mathcal{P}$ by identifying antipodal points of the sphere, but that we can also be content to let this identification operate just on a hemisphere (boundary included): we need then only identify antipodal points of the equator. We can still choose to keep a band about the equator, it still being required that we identify antipodal points in this band. We obtain in this way a Möbius strip and $\mathcal{P}$ then appears as the union of a Möbius strip and a spherical cap, i.e. a disk sewn together without ambiguity.

Fig. I.7.5. *Ways of seeing the projective plane. At left*: identify antipodal points. *Middle,* identify antipodal points of the equator. *At right,* identify antipodal points of a band (which comes down to preserving only the middle line of the band, while identifying ab and $b'a'$; pay attention to the direction of travel)

It is because $\mathcal{P}$ contains a Möbius strip that it is not an orientable surface; and it is for this reason that $\mathcal{P}$ is not embeddable in $\mathbb{R}^3$: if fact, to embed $\mathcal{P}$ in $\mathbb{R}^3$ would be to define a transformation of $\mathcal{P}$ into $\mathbb{R}^3$ that is continuous and injective; the geometer says *without double point* or *without self-intersection*. In view of compactness such a transformation would automatically realize a homeomorphism of $\mathcal{P}$ onto its image. However, a result from topology states that each compact surface in $\mathbb{R}^3$ without boundary possesses an interior and an exterior and is, for this reason, orientable.

A more direct proof of the fact that $\mathcal{P}$ is not embeddable in three dimensional space amounts to observing that a Möbius strip cannot be glued, following a common boundary, to a topological disk without the band and the disk intersecting. The boundary of the Möbius strip is indeed a circle (moreover unknotted), but a circle intertwined with the strip:

Fig. I.7.6.

the fact that Steiner discovered it when he was in Rome in 1844. We will encounter the Veronese surface in Sects. II.0 and V.9.

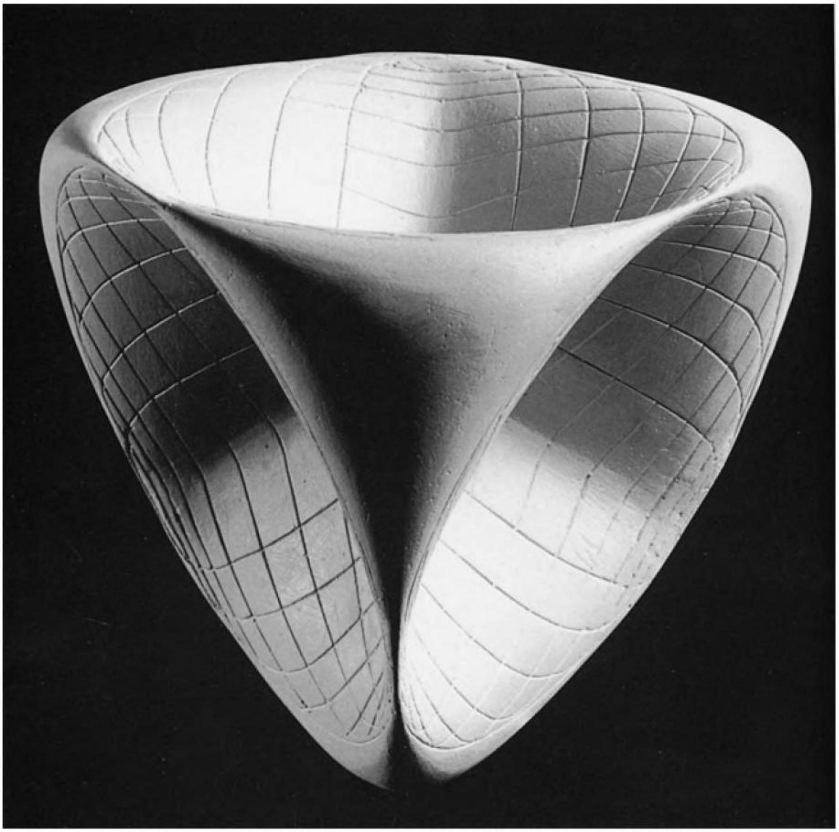


Fig. I.7.9. Steiner's Roman surface. Fischer (1986a) © G. Fischer

Note that the complement in $\mathcal{P}$ of a projective line is *connected*, in contrast to the affine case, infinity serving as the connection bond. The same thing is true for the median line of the Möbius strip. How do we see that we have the same phenomenon? By considering, in the projective plane, a band containing a given line D. The band situated between two lines parallel to D won't do, since it contains only a single point at infinity, but the region contained between the two branches of a hyperbola (situated on both sides of D) contains a whole segment of points at infinity, and it clearly has the topology of a Möbius strip, since it is obtained by identifying, in a rectangle, two opposite sides traversed in opposite senses.

This fact explains why a curve located on one side of its asymptote, when it tends toward infinity, always returns from infinity in the opposite direction. The curve is in fact tangent to its asymptote at its point at infinity, and if the contact is ordinary the curve does not cross its tangent. In particular this is the case for the point at

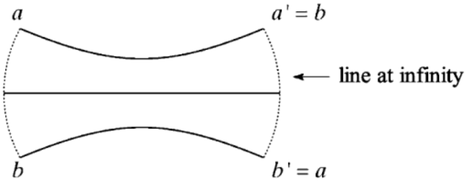


Fig. I.7.10. A neighborhood of a *line* in the projective plane has the topology of a Möbius strip

infinity of a hyperbola: the homography or homogeneous coordinate transformation, $(x, y, z) \mapsto (z, y, x)$, or $(x, y) \mapsto (1/x, y/x)$ in affine coordinates, transforms the hyperbola with equation $xy = z^2$ ($xy = 1$ in affine coordinates) into the parabola with equation $zy = x^2$ ($y = x^2$ in affine coordinates), tangent to the x axis at the origin. The curve does not cross its asymptote; it's the plane that makes a half turn like a Möbius strip.

As for the connectivity property indicated above, it explains the well-known trick of cutting a paper Möbius strip along its center curve and then continuing to cut along new median curves. It is left to readers to carry out the necessary experiments.

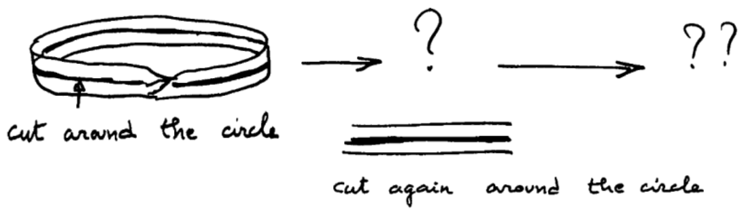


Fig. I.7.11.

♦

We have seen that there exists a canonical metric structure on $\mathcal{P} = \mathbb{R}\mathrm{P}^2$, derivative from that for the sphere: the distance between the two points $p, q \in \mathbb{R}\mathrm{P}^2$ is the angle, between 0 and $\pi/2$, between the two lines of $\mathbb{R}^3$ which give rise to p and q. This geometry is called *elliptic*; it must be seen as a generalization of Euclidean geometry, for any two projective lines intersect in a single point (which is not the case for the great circles of spherical geometry). Here there are never any parallels, whereas in hyperbolic geometry in contrast there is an infinity of parallels for any given line. For more details on elliptic geometry, see Chap. 19 of [B]. But here is an example to which we should pay attention. It has to do with studying the cases of the equality of triangles: are two triangles for which two sides are the same equal, in particular are their three angles the same? An initial remark: two points in $\mathcal{P}$ are joined by a unique shortest path (projection of an arc of a circle onto the sphere) if their distance apart is less than $\pi/2$, otherwise there are exactly two shortest paths; but we will only discuss triangles with distances between vertices all less than $\pi/2$.

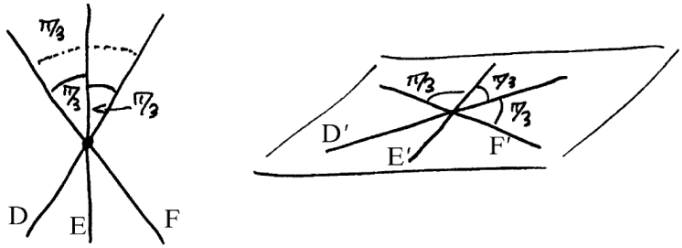


Fig. I.7.12. The *three lines* (D, E, F) and the *three lines* (D′, E′, F′) form *equilateral triangles* of side$\pi/3$ in $\mathcal{P}$. But these *triangles* are completely different: in the second the three angles equal π, in the first the three angles equal A, where $\cos A = \frac{1}{3}$ (apply the fundamental formula of spherical trigonometry)

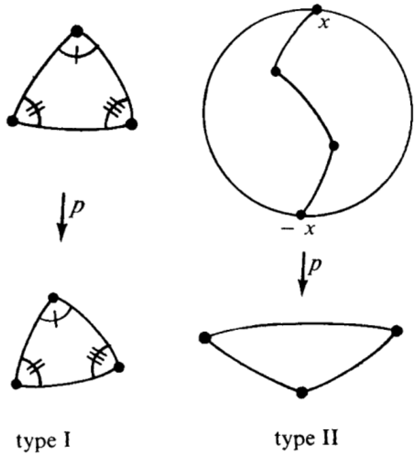


Fig. I.7.13. The two types of triangles in projective space, distinguished by their rise

We return to the equality of triangles, the exemplary case being, viewed as sketched in $\mathbb{R}^3$, that of a trihedron with three equal angles of $\pi/3$ and of a degenerate trihedron formed by three lines of a plane which makes among them the equal angles $\pi/3$. What can happen here, seeing that everything goes well in spherical geometry? To understand this, it is natural to go back to the sphere; but just one point of $\mathcal{P}$ provides two different points (antipodes) of the sphere. A curve of $\mathcal{P}$, here a side of a triangle, once a vertex has been lifted onto S^2, is lifted without ambiguity into S^2 because the projection of the sphere onto the projective plane is bijective and bicontinuous when restricted to a sufficiently small open set of S^2, typically the open hemisphere (spherical cap of aperture π) centered at a given point. Continuing in like manner for the two remaining sides we obtain a curve formed by three arcs of circles in S^2, but for which the terminal point is either the chosen point of departure or its antipode. These two cases are exactly those of the two triangles in $\pi/3$ considered. There are thus two types, I and II, of triangles in $\mathcal{P}$, but note that the type II will only be encountered if the sum of the sides is greater or equal to π. Readers will easily show, by deftly applying the case of equal spherical angles, that the equal angle case holds if, besides the equality of the respective sides, the two triangles considered are of the same type.

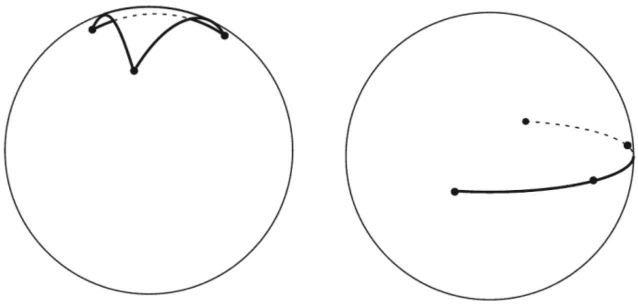


Fig. I.7.14. Lifting into S^2 of the two "exemplary" triangles of Fig. I.7.13

With the canonical metric structure of $\mathcal{P}$, the associated duality of Sect. I.7 is expressed thus: the projective line that is dual to a point p is made up of points of $\mathcal{P}$ located at a distance $\pi/2$ from p. For this associated geometry, perpendicular bisectors, etc., see [B]. It is interesting to note that elliptic geometry, which furnishes a trivial counterexample to Euclid's parallel postulate, was not known until well after hyperbolic geometry was discovered. This is due, among other reasons, to the difficulty of "seeing" the real projective plane.

Finally, we indicate why the only transformations that preserve lines are the projective transformations: the proof is achieved by fixing any line at infinity and applying the fundamental theorem of affine geometry to the complement. This result is often called the *second fundamental theorem of projective geometry*; see Sect. I.XYZ for the "first fundamental theorem".

I.8. The complex case and, better still, Sylvester in the complex case: Serre's conjecture

In Sect. I.1 we briefly alluded to affine geometry over the field of complex numbers, and even over the quaternions. The definition of the affine plane (over the reals), in which we have worked until now (see Sect. I.XYZ), extends trivially to the case of an arbitrary base field, not only to number fields, in particular the complex numbers $\mathbb{C}$ (commutative) and the quaternions $\mathbb{H}$ (non commutative), but also to all other fields, in particular the finite fields, the most simple among them being the field of two elements $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$. But there is no reason at all to restrict ourselves to dimension 2, since everything is constructed using only the algebraic theory of vector spaces (if necessary, refer to Sect. I.XYZ). Beyond what will be said in this section, see p. 9 of the introduction of Orlik and Terao (1992). For example, there is no strict notion of cross ratio in the non commutative case, typically over the quaternions, but only of conjugate classes, which take its place.

Let us repeat that Sylvester's theorem of Sect. I.1 is false over the complex numbers, the simplest numerical example is written in projective coordinates for the complex projective plane $\mathbb{C}P^2$; specifically, the nine points with projective coordinates

$$\begin{array}{lll} (0,1,-1), & (1,0,-1), & (1,-1,0), \\ (0,1,-\omega), & (1,0,-\omega), & (1,-\omega,0), \\ (0,1,-\omega^2), & (1,0,-\omega^2), & (1,-\omega^2,0). \end{array}$$

Here ω denotes a cubic root of unity other than 1, e.g. $(-1+i\sqrt{3})/2$. These nine points are the inflection points of the cubic (projective) equation $x^3+y^3+z^3-3axyz=0$ $(a \neq 0)$. We verify by hand, without having need of the theory of planar cubics, that on each line joining two of these nine points there is always a third. Algebraically the condition for the collinearity of three points is translated by the fact that the determinant of their nine coordinates is zero. This configuration $(9_4, 12_3)$ is called *Hesse's configuration*. Readers who like simple calculations but hate projective coordinates will be able, with the aid of an appropriate projective transformation, to write the coordinates for a system of nine points of this type in the complex affine plane. There are too many zeros in the three possible places in the above array to allow us to proceed solely by division of the same coordinate. We will have observed that collineation here is complex, but we can regard the condition by taking a complex vector space as a real vector of twice the dimension; then Sylvester's condition will be that we always have a third point on the complex line generated by two given points; but of course this complex line is in effect a real **plane**.

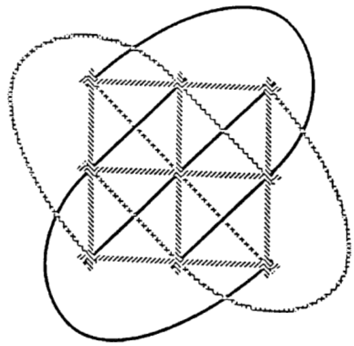


Fig. I.8.1. Hesse's configuration

In 1966 Jean-Pierre Serre announced the following conjecture: "*Let there be given, in a complex affine space of arbitrary dimension, a finite system of points satisfying Sylvester's condition: show then that this set of points is necessarily contained in a (complex) plane V*". We emphasize that the proof, if there is one, cannot be purely combinatorial; and that if Sylvester was wrong in the complex case, it's because a line D does not separate the complex plane $\mathbb{C}^2$ into two regions: $\mathbb{C}^2 \setminus \mathrm{D}$ is

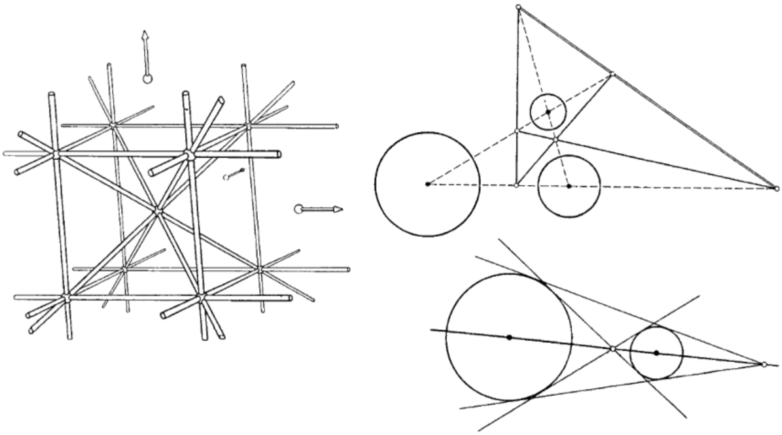


Fig. I.9.2. *At left*, Reye's configuration. *At right*, the configuration formed by the centers of homothety of *three circles* (below, recall the properties of the centers of homothety of *two circles*). Hilbert, Cohn-Vossen (1996) © Springer

infinity. But why can't we do the same thing for triangles? It is easy to convince ourselves of this impossibility "by hand", but that would not really get to the heart of the matter. We must climb the ladder, which we do in recalling the end of Sect. I.6, which turns out to be profitable in all dimensions. There we constructed a bijection between the points and lines of the plane, which with each triangle then associates a new triangle; but we can't require that this bijection have the property that the line associated with a point pass through that point. However, this is possible in a projective space (projective, so as to avoid exceptions) $\mathcal{Q}$ of three dimensions: there exists a bijective transformation f of $\mathcal{Q}$ onto the set $\mathcal{Q}^*$ of all its planes such that $p \in f(p)$ for each point p of $\mathcal{Q}$. Of course we require further that the properties of collinearity and intersection be preserved. Thus any tetrahedron defines a pair of Möbius type by adding to it the tetrahedral image under f. And here is the "parachuting" of such an f: we set up projective coordinates in some way (any choice will do) and define $f((a,b,c,d))$ to be the plane with equation $-bx + ay - dz + ct = 0$. The membership condition is obviously satisfied and, things being linear, all properties of intersection, collinearity, etc. are preserved as we would like.

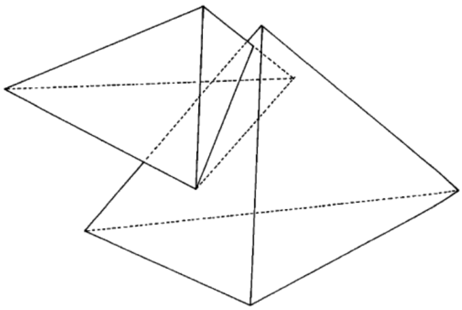


Fig. I.9.3. Möbius tetrahedra. I [B] Géométrie. Nathan (1977, 1990) réimp. Cassini. (2009) © Nathan Édition

In fact, we have indeed climbed the ladder somewhat. First, we see that we will be able do the same thing for every space of uneven dimension, the number of projective coordinates being even in this case. But above all the theory of all this was accomplished geometrically and very laboriously at the end of the nineteenth century. Now, algebraically, linear and multilinear algebra permit us to resolve completely all the questions that can be posed, and this in arbitrary dimension and over an arbitrary field. The essential problem is knowing what are the bijections between a space and its dual (the space of its hyperplanes) that preserve intersections and collinearity. The answer is that there are two possible types, ones that we have encountered: the type given by a Euclidean structure, i.e. a quadratic form (corresponding to a *symmetric* bilinear form) of maximum rank, and the type called *symplectic*, i.e. given by an antisymmetric bilinear form of maximum rank, it being understood that a symmetric or antisymmetric bilinear form of maximum rank on a vector space E defines a bijection of E onto its dual E^* and, by passage to the quotient, a bijection of the projective space $\mathcal{P}(E)$ onto the space of its hyperplanes. In the second type, we have the Möbius property that each point belongs to its dual; and it is trivial that these structures cannot exist in even dimensions (odd dimension for the vector space). Since the quadratic forms of maximum rank – at least in the real case – are categorical, as are the symplectic forms of maximum rank, we now know everything. The proof of this fundamental result is Exercise 14.8.12 of [B], a more complete reference is Frenkel (1973).

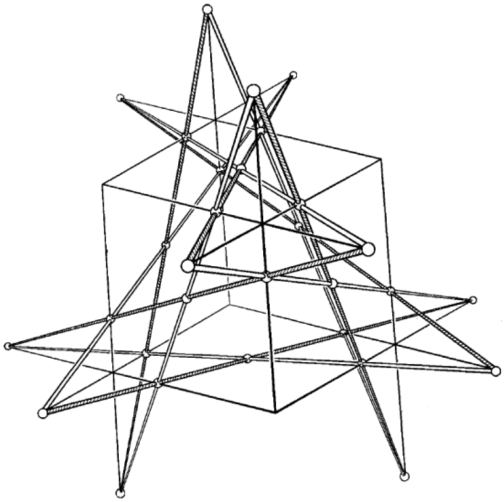


Fig. I.9.4. Schläfli's double six. Hilbert, Cohn-Vossen (1996) © Springer

♦

Schläfli's *double six* is given in the above figure. It's a configuration of type $(30_2, 12_5)$, but the notation here indicates 30 points and 12 lines (no planes). The proof of its existence isn't elementary; it makes essential use of algebraic geometry, specifically that a cubic surface (of degree 3) in three dimensional space and without singularities contains exactly 27 lines in the complex case, and in the real case 27, 15, 7 or 3 lines.

An "elementary" exposition is contained in §25 of Hilbert and Cohn-Vossen (1952). The point of departure is to take four lines in space in general position. There exist two lines that intersect them all, which is seen using what is found in Sect. IV.10: the set comprised by the lines intersecting three given lines in general position is a quadric surface, so that the desired lines, based on four lines, are obtained by looking for the points of intersection of the fourth line with the quadric surface defined by the first three. In general there are two such points, our four lines being "generic". The construction of the double six begins thus: we start with a line D_1, then construct at random five lines E_2, E_3, E_4, E_5, E_6 intersecting D_1. Let D_6 be the next line, other than D_1, intersecting the four lines E_2, E_3, E_4, E_5. Then

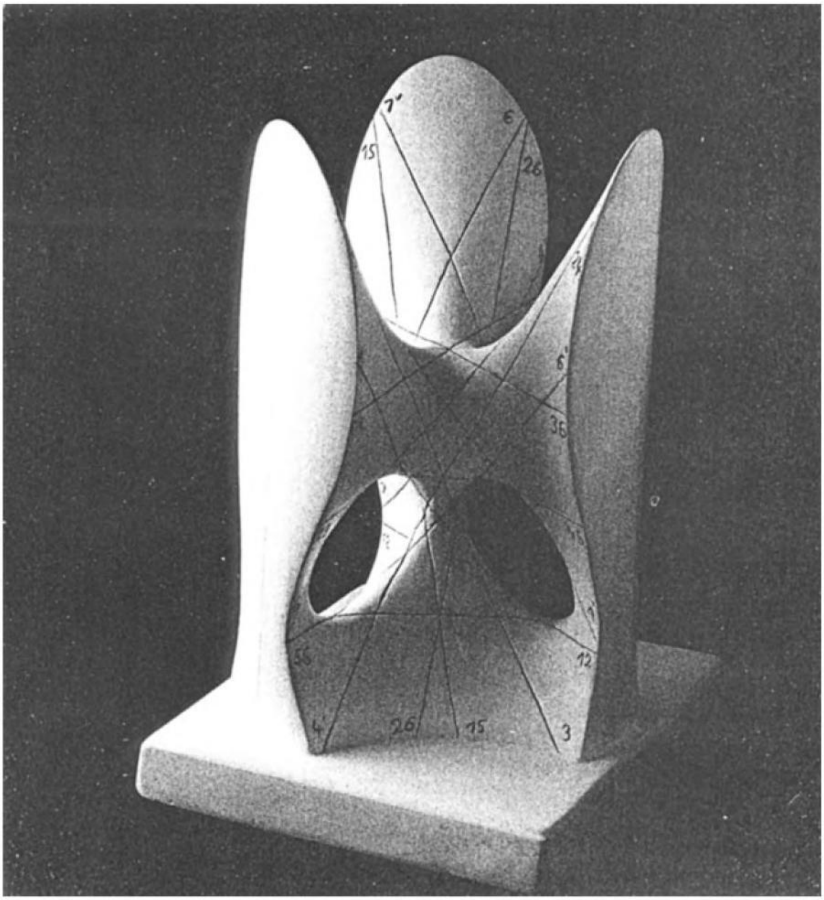


Fig. I.9.5. Clebsch's diagonal surface. Fischer (1986a) © G. Fischer

D_2, which in addition to intersecting E_3, E_4, E_5, E_6, likewise intersects D_3, D_4, D_5. It remains to define E_1 as the line likewise intersecting D_2, D_3, D_4, D_5. Now, we have numerous other intersections that form, by construction, a configuration of type $(30_2, 12_5)$. But the proof is difficult and, as already indicated, requires the use of algebraic geometry; see the book cited above for an heuristic proof.

We will encounter this configuration in Sect. V.16. There indeed exist real cubic surfaces containing 27 lines, all without singularities, e.g. Clebsch's diagonal surface. The theoretic considerations are found in Fischer (1986b), they depend on Fig. I.9.5.

I.10. Arrangements of hyperplanes

We have seen that there is a difficulty in studying the finite sets of lines in the real affine plane, because of the possibility of parallels. Nonetheless, the question of the **maximum** number of connected components of the complement of n lines was resolved before the twentieth century, and in fact by Schläfli for the complement of n hyperplanes in arbitrary dimension. In order to have the maximum number, it is necessary to be in the generic case; what happens when more than two lines intersect, and for the parallels, dates from 1889. This does not give a classification; such a study leads in fact to problems in topology, algebraic geometry, convexity, combinatorics, number theory, (arithmetic) analysis and geometry without any obvious relations. We will not speak further here of the real case; we refer readers to its introduction in Orlik and Terao (1992) or pp. 679–706 of volume II (Hirzebruch, 1987b), not to forget the commentaries of pp. 802–804.

The complex case has an even greater richness, and provides much pleasure when connections are found between seemingly different things. The book cited above (Orlik and Terao, 1992) is entirely devoted to it; it's a subject that ascends quite high on the ladder and has a very recent development. We can't speak about it in detail, or even partially. We have seen a first approach in Sect. I.7 above. In a second approach, we first remark that, if the complement of a hyperplane in a complex space is connected, contrary to the real case, in return it isn't *simply connected*: if we "make the tour" of a hyperplane, we obtain a loop (closed curve) that isn't contractible to a point. More generally, with a finite set of hyperplanes we associate the *fundamental group* of its complement, i.e. the group generated by the closed curves through a fixed point (in the complement, of course) that "turns" about the hyperplanes of the arrangement under consideration. We obtain this group by considering two closed curves to be equivalent if they are deformable each into the other, in a continuous fashion, without of course intersecting the figure about which they are turning. The nature of this group and more generally the topology of this complementary set are much studied nowadays and yield surprising relationships.

The simplest case is the topology of the complement of two intersecting lines of $\mathbb{C}^2$: it is easily seen that this set can be continuously deformed onto the product of two circles, i.e. a torus. In particular, the fundamental group of this complement is that of the torus, thus $\mathbb{Z}^2$.

I. XYZ

For the mathematical objects described above and for proofs, a general and complete reference is [B], to which we add Hilbert and Cohn-Vossen (1952) for the configurations.

Before mathematicians, curious and practical minds would ask how to make lines straight, planes "flat" – all this both in drawing instruments and in industrial practice of a more or less high degree of precision; and finally for the highest degree of precision, that of metrology. We will give some brief indications about this problem in Sect. II.XYZ.

The real affine plane is defined as follows: we consider a real vector space P of dimension 2, or, what amounts to the same thing – modulo isomorphism – the set $\mathbb{R}^2$ of pairs (x, y) of real numbers. The *points* are thus the elements of $\mathbb{R}^2$; the *lines* are all the subsets that are translations of a vectorial line of $\mathbb{R}^2$. We might say that our affine plane is a real vectorial plane **for which the origin has been forgotten**; there is no longer an origin, no special point. The properties enjoyed by lines and planes, and their relationships, are direct consequences of the axioms for vector spaces. This definition will not satisfy very abstract minds; there remains a mental trace of an origin. In [B] there is a more axiomatic presentation. Here we will retain our point of departure: with each pair (p, q) of points of the affine space there is associated the vector ("free" in the old language) that may be *denoted* $q - p$, or often also $\overrightarrow{pq}$, which belongs to the vector space, giving rise to the affine space under consideration.

An important element, which serves as the foundation for the more conceptual construction mentioned above (i.e. not favoring an origin, or even the ghost of one), is the **barycentric calculus**. Two points p, q cannot be added in an intrinsic manner (we can only subtract them, but the result is a vector and not a point of the space considered. On the other hand, the midpoint of the two points can be written $\frac{p+q}{2}$, and more generally we can divide a segment by a given ratio; see the affine invariant introduced in Sect. I.3. What matters in $\frac{p+q}{2}$ is that it can be written $\frac{1}{2}p + \frac{1}{2}q$. More generally we can define in an intrinsic and practically trivial fashion any finite sum $\sum \lambda_i p_i$ such that the sum of the coefficients (called barycentric) satisfies $\sum \lambda_i = 1$. Two facts are essential: the associativity of this operation and the uniqueness of the coefficients for the sums for three points forming a true triangle. This furnishes "purely affine" coordinates for the plane, coordinates called *barycentric*; for all the details see Chap. 3 of [B]. Briefly and in any dimension, since no changes are needed: if $\{p_i\}$ $(i = 1, \ldots, d + 1)$ denotes a set of $d + 1$ points in real affine space of dimension d forming a *simplex* – which means that none of them belongs to the hyperplane defined by the remaining d – then for each point x of the space there exists a unique expression (a "barycentric sum of the $\{p_i\}$") $x = \sum \lambda_i(x) p_i$

consisting of the identity element 1 and seven other generators $\{e_i\}$ $(i = 1, ..., 7)$ for which the multiplication table is furnished by the triangles inscribed in a heptagon as in Fig. I.XYZ.2.

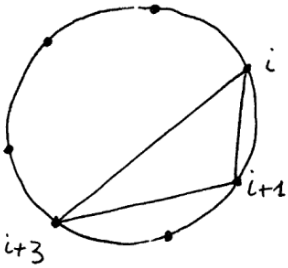


Fig. I.XYZ.2.

More precisely, we consider all triples of the form $\{e_i, e_{i+1}, e_{i+3}\}$ for i going from 1 to 7 and where the additions in the subscripts are computed modulo 7. Each of these triples obeys the same laws as the triple $\{i, j, k\}$ from the definition of the quaternions: $e_i e_{i+1} = e_{i+3}$, $e_{i+1} e_i = -e_{i+3}$. Even though this law is not ultimately associative, we can nonetheless define a reasonable geometry over $\mathbb{C}\mathbb{a}$, not just in dimension 2, but also in dimension 3 and beyond. For references on this topic, which is especially subtle in the case of the construction of the "panda" encountered in Sect. I.4, the projective plane over $\mathbb{C}\mathbb{a}$, see 4.8.3 of [B] and the references given there. See above all the entirety of Chap. XIV of (Porteous, 1969). A recent reference on octonions is Baez (2002).

♦

We must pay strict attention to the tricks intuition can play, as soon as we are working over the complex field instead of over the reals. A plane over the complex numbers has real dimension 4. A typical example is this: what is the topology of the complement of a point in $\mathbb{C}$? It is that of a cylinder, which may be contracted to a circle, whereas in $\mathbb{R}$ it amounts to two intervals, which contract to two points; in each instance the complement in $\mathbb{R}$ is not connected, which is no longer the case in $\mathbb{C}$. And for the complement of two intersecting lines in the plane $\mathbb{C}^2$? In the case of the real plane, we find four connected components. But here the complement is connected; if we contract it onto the intersection of $\mathbb{C}^2$ with the unit sphere S^3 in $\mathbb{R}^4$, it turns out, according to Sect. II.5, that the complement has the same topology as the complement in S^3 of two orthogonal circles, i.e. of a torus T^2. Another motivation for geometry over the complex numbers is the recent notion of *quantum computers*; these are objects that truly operate in $\mathbb{C}$.

♦

Another caveat is that, when we work in **higher dimensions**, our intuition loses practically all its rights, all its effectiveness. We will see numerous manifestations of this in Chap. VII. But high dimensions are absolutely necessary in numerous mathematical studies, both theoretical and applied, especially since about 1950. Loading an image on a computer requires working in spaces of, say, dimension 10 000. Think also about credit cards, of cryptography, of biology and the error correcting codes that are studied in Chap. X. All of linear programming involves a potentially large number of "consumers", and hence enormous dimensions.

Another need is that of functional analysis, which leads to the study of convex sets in very high dimensions; see Chap. VII. Our intuition will be put to the test there, starting with the volume of spheres of radius 1, for which the volume tends ultra-fast to zero as the dimension becomes large.

Still another need is that of physics, where statistical mechanics treats sets of particles with numbers of order 10^{23}, which thus lives in spaces having such dimensions.

♦

Also interesting is the remark due to Pierre Cartier that, in fact, we scarcely comprehend more than dimension 1. The reason isn't merely that our brain works in a linear fashion, sequentially, for in fact (as far as is known today) it works rather like a parallel computer. It seems that this is simply for the crude reason "*that it's simpler*". For in fact dimension 2 isn't yet well understood; pattern recognition, for example, is in its infancy. Dimension 3, however indispensable for all the objects of space, harbors numerous open problems, as we shall see. Some people think that our difficulties in dimension 3 are due, at least in part, to the fact that the group of rotations in space is not commutative. But all these reflections on the profound nature of mathematics and the functioning of our brain are still in limbo. To paraphrase Pierre Cartier, in (Cartier, 1991): "The very subjective, if not to say blurred, character of art criticism is without doubt due to this characteristic. In our epoch of intensive use of computers and of computer-aided creativity, there is a regrettable gap." Making a leap from dimension 1 to dimensions 2 and 3 is one of the present obstacles to progress in mathematics. Similar obstacles have led to the creation of non-Euclidean geometry and to the discovery of Gödel's theorem.

♦

The fundamental theory of affine geometry, seen for the real plane in Sect. I.3, is proved by the same technique for all dimensions and all fields. But there are two very important modifications that we give here; for more, see 2.6 of [B]. We proceed first to the obvious exclusions (the miracle is that there aren't others). First the case of dimension 1, where the collinearity condition is vacuous: every bijection preserves lines, since there is but one line here! But things get more complicated according to the qualities of the field K considered. It is necessary at first to completely exclude the field $\mathbb{Z}_2$ of two elements, for then the affine line contains only two points and the collinearity condition is again vacuous. But above all the proof of I.3.1 goes through thanks to the fact that only the identity is an automorphism of the real number field.

It's not entirely the same for other fields, e.g. the field of complex numbers admits the automorphism that transforms z into its complex conjugate $\bar{z}$: conjugation preserves sums and products. A *semi-affine* transformation of an affine space into a field K is, by definition, modulo a translation, a *semi-linear* transformation (more generally we can speak of a semi-affine transformation of one affine space into another), i.e. a transformation f such that $f(\lambda x + \mu y) = \sigma(\lambda) f(x) + \sigma(\mu) f(y)$, where $\sigma : \mathrm{K} \to \mathrm{K}$ is any automorphism of the field K. We point out that $\mathbb{C}$ admits lots of other automorphisms besides the identity and the conjugation $z \mapsto \bar{z}$, but if we further require continuity, then there remain only those two. The fields of characteristic k different from 0 admit the celebrated *Frobenius automorphism* $x \mapsto x^k$, whose importance in mathematics should not be underestimated, e.g. it enters Frobenius, Georg into Deligne's proofs of Ramanujan's conjectures; see Sect. III.3 (and Sect. III.6). On the other hand, the quaternion automorphisms are easy to classify; see 8.12.11 of [B]. And so these "refined" mappings are never so frightening as they can be in the case of the complex numbers.

♦

We now define a projective space of dimension n over the field K as the set of "vectorial lines" (one-dimensional subspaces) of a vector space of dimension $n + 1$ over K. All these spaces are in fact the same, so we can speak of *the projective space of dimension n over* K; we *denote* it by KP^n. Algebraically, it's the quotient of $\mathrm{K}^{n+1} \setminus 0$ modulo the equivalence relation such that $w \equiv v$ if and only if there exists $k \in \mathrm{K}$ (necessarily nonzero) for which $w = kv$. The (projective) *lines, planes, hyperplanes* of KP^n correspond to vector subspaces of K^{n+1} of respective dimensions 2, 3 and n. The case of the projective line is the object of study of Chap. 6 of [B].

We introduced the projective spaces essentially for the purpose of completing affine geometry so that the intersection theorems could be presented without exceptional cases in their statements; but apart from the miracles mentioned below, the idea of considering the elements of a vector space "within a scalar factor" is very natural, if not to say indispensable. We will see two examples in this book, the first is the space of circles and spheres in Sect. II.6; the second is that of conics and quadrics (see Sect. IV.7). It isn't possible to really thoroughly understand circles, spheres, conics and quadrics without introducing the projective space formed **by their equations**. This is very much in the spirit of Jacob's ladder. Here now, very briefly, are the essential properties of projective geometries.

The complex projective spaces are among the most important objects of algebraic geometry.

The projective transformations are (bijective) linear mappings of $\mathrm{K}^{n+1} \setminus 0$ **retracted** onto KP^n; they are also called *homographies*. They obviously form a group, called the *projective group*, denoted by $\mathrm{GP}(n\,;\mathrm{K})$, whose structure is that of the quotient of the linear group $\mathrm{GL}(n\,;\mathrm{K})$ by the group of multiples of the identity. In terms of coordinates, it is the multiplicative group of the $(n+1) \times (n+1)$ matrices with elements in K, modulo the group of multiples of the identity matrix. The projective spaces over finite fields are encountered in combinatorial geometry. Here is a typical example in the spirit of Sylvester's problem of Sect. I.1: whatever its dimension, the projective space $\mathbb{Z}_2\mathrm{P}^n$ over the field $\mathbb{Z}_2$ of two elements is a finite set such that, for each pair of points, there is a third point on the line that joins them. In fact, all projective lines over $\mathbb{Z}_2$ have three elements: two affine and their point "at infinity".

The homographies of the projective line are classified and studied in detail in Chap. 6 of [B]. The involutions (homographies which when squared yield the identity) play a particularly important role. If a homography has two distinct fixed points a and b, then the cross ratio $[a, b, m, f(m)]$ is constant; conversely, the relation $[a, b, m, f(m)] = k$ defines a homography, an involution when $k = -1$.

♦

As indicated above in Sect. I.7, there is a *second* fundamental theorem of projective geometry valid in any dimension and over any field; but as in the affine case it is necessary at first to make the obvious exclusion of dimension 1. But subsequently there will be no need to exclude the field of two elements; for a projective line, in contrast to an affine line, always contains at least three points. The conclusion is that we find only semi-projective transformations, i.e. projective transformations modified if need be by an automorphism of the field; see as needed 5.5.8 of [B] and the references mentioned there. The *first fundamental theorem of projective geometry* states that the transformations in projective dimension n are sufficiently abundant in order to transform two arbitrary $(n+2)$-tuples of points into one another; see as necessary 4.5.10 of [B]. The result is trivial using linear algebra, although it bears a ponderous name, which comes from the fact that it is difficult to prove and is much more concealed if we pursue the *axiomatic theory of projective geometry*. Here we have parachuted the projective spaces with linear algebra, while axiomatic projective theory constructs them (more or less completely) with axioms bearing on points, lines, etc., and properties required for their various intersections, properties that are trivial in linear algebra. Readers interested in the axiomatic theory can consult the two basic books that exist Artin (1957) and Baer (1952). The theory remains difficult and, in our opinion, of mediocre elegance for the case of projective *planes*, since there is no Desargues' theorem – seen in Sect. I.4 – at our disposal; readers may be able to form an opinion with a cultural text such as (Lorimer, 1983).

In complex geometry it is very important to know that, while topologically the real projective line is the circle S^1, the complex projective line is the sphere S^2; see the end of Sect. I.6. Note that the real plane – as a real object – is projectified into $\mathbb{R}\mathrm{P}^2$ by appending a line at infinity, topologically a circle. But the same real plane, when seen as a complex line, is projectified with a single point at infinity, which yields the sphere S^2 topologically. Thus we have two quite different compactifications of $\mathbb{R}^2$. In Sect. II.3 we will see a third, also utterly essential, where compactification provides the topology of a disk; the interior is $\mathbb{R}^2$ but the compactification adds the boundary circle. For more on the topology of projective spaces, which are elements of constructions essential in algebraic topology, there are some references in Chap. 4 of [B]. We mention only that $\mathbb{R}\mathrm{P}^n$ is orientable for all odd

n, nonorientable otherwise. Let us add here, however, that the complex projective spaces are the generating elements of the cobordism ring in algebraic topology. The theory of cobordism, which dates from the 1950s, is a classification of compact differential manifolds; see Husemoller (1975). As for homographies $z \mapsto \frac{az+b}{cz+d}$, which dominate a large part of mathematics, we will encounter them in Sects. II.3 and II.4. For a modern proof of the fundamental theorem of geometry, affine or projective, see Faure (2002).

♦

A *configuration* of an affine or projective plane is simply a specification of a set of points and the lines joining those points. Such a configuration evidently does not have any interest unless we require particular properties. We say that a configuration τ is of type (p_q, r_s), where p, q, r, s are integers, if there are p points and r lines, such that for each point of τ there pass exactly q lines of τ and if every line of τ contains exactly s points of τ. An accounting implies the relation $pq = rs$. The proof of the existence of a configuration of a given type may be trivial, but also more or less difficult.

The complete quadrilateral is a configuration $(4_3, 6_2)$; existence is trivial but nevertheless interesting to interpret in Euclidean geometry as that formed by the six centers of homothety of three circles; this has served us above for the Reye configuration. In Sect. I.7 we saw its very useful property of harmonic conjugation, proved by "transfer to infinity". Here we give a purely projective proof "in situ".

The complete quadrilateral is thus the figure formed by four lines in general position (the sides) and their six points of intersection (the vertices). These points of intersection are joined pairwise by three diagonals. The property of harmonic conjugation mentioned above is the following: two vertices a and b not located on the same side (and thus located on one of the diagonals) having been chosen, and i being the point of intersection of the two other diagonals, the sides issuing from a divide the segment bi harmonically. This means that, if we let x and y denote the points of intersection of the sides passing through a with the line bi, we have $(b, i, x, y) = -1$.

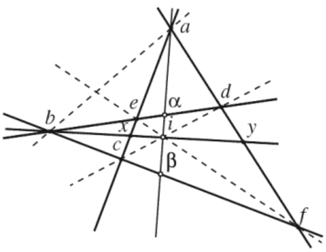


Fig. I.XYZ.3.

In Sect. I.7, this property was proved by transferring the points a and b to infinity, and thus in making a parallelogram out of the quadrilateral. The "classical"

Baez, J. (2002). The octonions. *Bulletin of the American Mathematical Society, 39*(2), 145–205

Bárány, I. (2008). Random points and lattice points in convex bodies. *Bulletin of the American Mathematical Society, 45*, 339–365

Berger, M. (2000a). Encounter with a geometer I, II. *Notices of the American Mathematical Society, 47*(2), *47*(3), 183–194, 326–340

Berger, M. (2005). Dynamiser la géométrie élémentaire: introduction à des travaux de Richard Schwartz. Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti Lincei. Serie IX. Matematica e Applicazioni. Accad. Naz. Lincei, Rome, Ser. 25, 127–153

Besse, A. (1978). *Manifolds all of whose geodesics are closed*. Berlin/Heidelberg/New York: Springer

Bokowski, J., & Richter-Gebert, J. (1992). A new Sylvester-Gallai configuration representing the 13-point projective plane in Υ^4. *Journal of Combinatorial Theory B, 54*, 161–165

Bott, R. (1988). Morse theory indomitable. *Publications mathématiques de líInstitut des hautes études scientifiques, 68*, 99–114

Bott, R., & Tu, L. (1986). *Differential forms in algebraic topology*. Berlin/Heidelberg/New York: Springer

Cartier, P. (1991, octobre). "Le calcul des structures à deux ou trois dimensions est un défi pour les mathématiciens". *Pour la Science, 168*, 8–10

Coxeter, H. (1964). *Projective geometry*. New York: Blaisdell

Coxeter, H. (1974). *Regular complex polytopes*. Cambridge: Cambridge University Press

Coxeter, H. S. M. (1989). *Introduction to geometry*. New York: Wiley

Croft, H., Falconer, K., & Guy, R. (1991). *Unsolved problems in geometry*. Berlin/Heidelberg/New York: Springer

Darboux, G. (1917). *Principes de géométrie analytique*. Paris: Gauthier-Villars

Du Val, P. (1964). *Homographies, quaternions and rotations*. Oxford: Oxford University Press

Faure, C.-A. (2002). An elementary proof of the fundamental theorem of projective geometry. *Geometriae Dedicata, 90*, 145–151

Fischer, G. (1986a). *Mathematische Modelle* [Mathematical models]. Braunschweig: Vieweg

Fischer, G. (1986b). *Mathematical models: Commentary*. Braunschweig: Vieweg

Frenkel, J. (1973). *Géométrie pour l'élève professeur*. Paris: Hermann

Gromov, M. (1999). *Metric structures for Riemannian and non-Riemannian manifolds*. In J. Lafontaine & P. Pansu (Eds.). Basel: Birkhäuser

Gruber, P., & Wills, J. (Ed.). (1993). *Handbook of convex geometry*. Amsterdam: North-Holland

Hilbert, D., & Cohn-Vossen, S. (1952). *Geometry and the imagination*. New York: Chelsea

Hilbert, D., & Cohn-Vossen, S. (1996). *Anschauliche Geometrie*. Berlin/Heidelberg/New York: Springer

Hirschfeld, J. (1979). *Projective geometry over finite fields*. Oxford: Clarendon Press

Hirzebruch, F. (1987a). *Selecta*. Berlin/Heidelberg/New York: Springer

Hirzebruch, F. (1987b). *Collected papers*. Berlin/Heidelberg/New York: Springer

Husemoller, D. (1975). *Fibre bundles*. Berlin/Heidelberg/New York: Springer

Kelly, L. (1948). The neglected synthetic approach. *The American Mathematical Monthly, 55*, 24–26. (Kelly's solution of Sylvester's problem can be found at the end of an article by H.S.M. Coxeter in the same issue.)

Kelly, L. (1986). A resolution of the Sylvester-Gallai problem of J.-P. Serre. *Discrete & Computational Geometry, 1*, 101–104

Klee, V. (1969). What is the expected volume of a simplex whose vertices are chosen at random from a given convex body? *The American Mathematical Monthly, 76*, 286–288

Kuiper, N. (1984). Geometry in total absolute curvature theory. In W. Jäger, J. Moser, & R. Remmert (Eds.), *Perspectives in mathematics: Anniversary of oberwolfach* (pp. 377–393). Basel: Birkhäuser

Lidl, R., & Niederreiter, H. (1983). *Finite fields*. Cambridge: Cambridge University Press

Lorimer, P. (1983). Some of the finite projective planes. *The Mathematical Intelligencer, 5*, 41–50

Orlik, P., & Terao, H. (1992). *Arrangements of hyperplanes*. Berlin/Heidelberg/New York: Springer

Pach, J., & Agarwal, P. (1995). *Combinatorial geometry*. New York: Wiley

Porteous, I. (1969). *Topological geometry*. London: Van Nostrand-Reinhold

Santalo, L. (1976). *Integral geometry and geometric probability*. New York: Addison-Wesley

Scheinerman, E., & Wilf, H. (1994). The rectilinear crossing number of a complete graph and Sylvester's "four point problem" of geometric probability. *The American Mathematical Monthly, 101*, 939–943

Schwartz, R. (1993). Pappus's theorem and the modular group. *Publications mathématiques de líInstitut des hautes études scientifiques, 78*, 187–206

Schwartz, R. (2001). Desargues theorem, dynamics and hyperplane arrangements. *Geometriae Dedicata, 87*, 261–283

Smith, A. (2000). Infinite regular sequences of hexagons. *Experimental Mathematics, 9*, 397–406

Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publications mathématiques de líInstitut des hautes études scientifiques, 81*, 73–205

Veblen, O., & Young, J. (1910–1918). *Projective geometry*. Boston, MA: Ginn and Co.

Chapter II
Circles and spheres

II.1. Introduction and Borsuk's conjecture

If the first chapter was essentially about affine and projective geometry, we now want to enter the Euclidean realm, i.e. we will now have a metric at our disposal, a notion of distance between points, with subsidiary notions such as *circles* and *spheres*. The basic reference for circles and spheres, completely authoritative at the time of its publication, is Coolidge (1916). We have made a critical selection from the enormity of classical results; see the very beginning of Sect. II.2. But of course above all we have chosen to talk about recent results, all the more if they require a climb up the ladder.

♦

Borsuk's conjecture. In the spirit of this book and before touching on problems leading us to configurations that are natural but rather sophisticated, we must speak about Borsuk's conjecture. Its statement is trivial, except that it deals with arbitrary dimension. It is one of the simplest assertions in all of Euclidean geometry, for it doesn't involve anything but distance. Here it is formulated as a question, where $\mathbb{E}^d$ denotes d-dimensional Euclidean space without reference to any particular coordinatization (in contrast, $\mathbb{R}^d$ is d-dimensional Euclidean space with canonical coordinates):

(II.1.1) *In the Euclidean space* $\mathbb{E}^d$, *can we decompose any bounded part* E *into* $d + 1$ *parts of diameter strictly less than the diameter of* E*?*

Fig. II.1.1.

The figures above seem to show that the problem is indeed trivial. Recall that the *diameter* of a bounded set E of an arbitrary metric space is the supremum of the distance between points of E: $\operatorname{diam}(E) = \sup\{d(x, y) : x, y \in E\}$. The conjecture is certainly true for $d = 1$, and it is also evident that $d + 1$ is necessary; that d doesn't suffice is left to the readers. We are going to describe in a bit of detail some

M. Berger, *Geometry Revealed*, DOI 10.1007/978-3-540-70997-8_2,

of the history of this conjecture, whose various aspects seem enriching. For the algebraic topologist, we need first say that in Borsuk (1933) that – among other things – an important result was proved that had been conjectured by Ulam, specifically that each continuous mapping of the sphere S^d into $\mathbb{E}^d$ sends at least one pair of antipodal points onto the same point, which is intuitive enough for $d = 1$ or $d = 2$. At the end of the text, Borsuk conjectured (II.1.1), or rather he merely stated the problem, for he was much too good a mathematician to make a conjecture without knowing a lot more. In fact his theorem just quoted implicitly suggests that we can't ever decompose the sphere S^d into only d pieces of smaller diameter, whence the idea for the problem. Borsuk's proof of the quoted theorem was a bit complicated, but H. Hopf pointed out to him that it is really instantaneous thanks to degree theory; see Chap. 7 of [BG].

Fig. II.1.2. Borsuk's theorem for $d = 1$ and $d = 2$

Borsuk's conjecture didn't lack for dedicated investigators; but even for the plane, the matter isn't quite as trivial as it perhaps seems. For this and for the more recent history, the reference is 19.3 of Grünbaum (1993). We only mention that, for the plane, we can proceed thus: we inscribe E in a convex set E′ of constant width (see the end of Sect. VI.9) equal to the diameter of E. This is easily done by taking the intersection of all the circular disks containing E. Next the continuity principle (turn around!) shows that there exists a regular hexagon circumscribed about this new body whose opposite sides are a distance apart equal to our diameter. We cut this hexagon into three (non regular) hexagons of diameter equal to $\frac{\sqrt{3}}{2}$, supposing $\mathrm{diam}(E) = 1$; this cuts E a fortiori in the same way. We have thus succeeded, moreover with a gain of a factor always equal to at least $\frac{\sqrt{3}}{2}$.

For three dimensional space, it was necessary to await (Eggleston, 1957) for the first proof. No proof is really very simple. The best gain presently known is 0.9987...; whereas $0.888... = \sqrt{(3+\sqrt{3})/6}$ has been conjectured. The fact that it is much closer to 1 than for the plane allows us to predict a difficult prospect for dimension 4. Here again the proof consists of inscribing E in an "adapted" polytope, specifically and for the moment a regular octahedron whose parallel faces are a distance apart equal to the diameter in question and, after having trimmed it down at the vertices, subjecting it to some subtle dissections (see Fig. II.1.4). Later proofs of a combinatorial nature were found using finite point sets. The properties of their

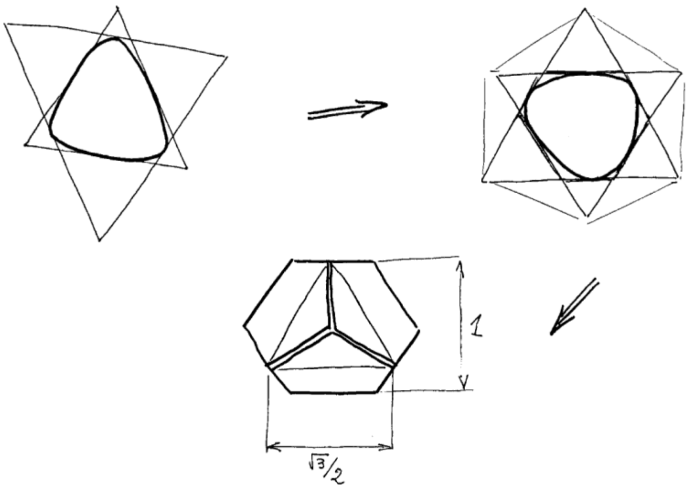


Fig. II.1.3. Circumscribe about our body *equilateral triangles* in all directions; by continuity at least two of these *triangles* are congruent. Moreover, as these two *triangles* have their parallel sides at a constant distance, their intersection must be a regular hexagon

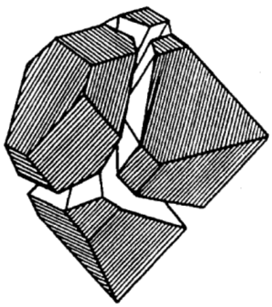


Fig. II.1.4. Boltyanski, Martini, Soltan (1997) © Springer and H. Martini

mutual distances are studied; see e.g. 13.15 of Pach and Agarwal (1995), Chap. 1 of Zong (1996) and Boltyanski, Martini, and Soltan (1997).

Meanwhile, for higher dimensions, it was necessary to make do with results for sets E having special properties. First, the solid balls in all dimensions are easily cut up – see the figure above – e.g. by inscription in a regular simplex. Next, every E having a center of symmetry is cut up using the following elementary observation: *when two points* p, q *of* E *attain the diameter of* E, *then* E *is contained between the two parallel hyperplanes that pass through* p *and* q *and are orthogonal to the line* pq. If there is a center of symmetry for the cut (and there must be, since E has a center of symmetry), this can't be anything other than the midpoint O of pq; and thus E is entirely contained in the ball whose center is O and whose boundary sphere passes through p, q. We now need only cut this ball as above or in some other way.

real projective space $\mathbb{RP}^N$, which is realized isometrically by the Veronese manifold above. Let A be a **connected** portion of $\mathbb{RP}^N$ with diameter less than $\pi/2$. We can then lift it within S^N into A', which will be contained in a hemisphere. Now in S^N we have an *isodiametric* inequality which generalizes that of the penultimate subsection of Sect. VII.7:

> *Among all the domains of* S^N *contained in a hemisphere and of given volume, the smallest diameter is attained exactly by the spherical caps.*

The proof proceeds by the Steiner symmetrization (see Sect. VII.5) – for this method extends without difficulty onto the sphere – while paying attention to the convexity. It's for this reason that we need to be in a hemisphere from the outset. The diameter here being less than $\pi/2$, the greatest volume will be obtained by a cap of radius strictly smaller than $\pi/4$. We then finish with a calculation of volumes; the disjoint parts partitioning A' cannot be greater in number than the quotient of the total area of the hemisphere with that of a cap of radius $\pi/4$.

A recent text on Borsuk's conjecture is Raigorodskii (2004). For more on Borsuk's conjecture, as well as for other "strange phenomena" in geometry, see the book that is entirely devoted to them: Zong (1996), especially Sects. VII.11, VII.12, VII.13.D and the slicing conjecture in Sect. X.6.

II.2. A choice of circle configurations and a critical view of them

We present some figures formed by circle configurations (see Sect. II.XYZ and in particular "a scandal to repair") in *the Euclidean plane* $\mathbb{E}^2$. We will comment on them later, restricting our attention to those that arise from the idea underlying this work. Specifically, we study geometric situations that can be stated quite simply, but which have led, and possibly still lead, to more intensely conceptual developments, to rungs up Jacob's ladder. We comment on what exists for finding proofs (entire or partial) or for understanding them deeply or placing them in a more general context, etc. We hope that readers will take the trouble to contemplate these figures at length, to compare them and decide which are interesting and which, to the contrary, seem unaesthetic or otherwise unappealing. Chapter I was set in the **affine** context; we are now entering into the **metric** context. Readers destined to want to embrace everything, i.e. to climb the ladder in order to unify their vision, will find in Sect. II.XYZ how to connect the metric to the complex projective setting by projectifying, and then complexifying, the Euclidean plane.

♦

For those who like constructions and practical matters, we should point out that circles are constructed with compasses, whereas **linear rulers** and also **graduated goniometers** (protractors) for **measuring** angles are always produced in a "physical" and thus approximate way; see Sect. II.3 below, Sect. V.7 and above all Sect. II.XYZ of this chapter for more on this.

And then there is the question of critical perspective: an historical example is that of “triangle geometry”, a discipline which has seen a disproportionate flowering, but which to our knowledge has contributed absolutely nothing in the way of Jacob’s ladders. The analysis (Davis, 1995) is very interesting, even if we don’t completely share the conclusions. References on the subject can be found there of course, among them the classic (Lalesco, 1952).

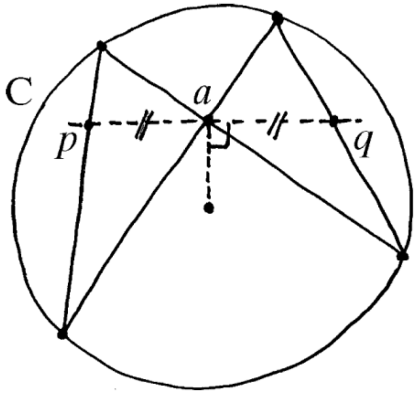


Fig. II.2.1. The butterfly theorem

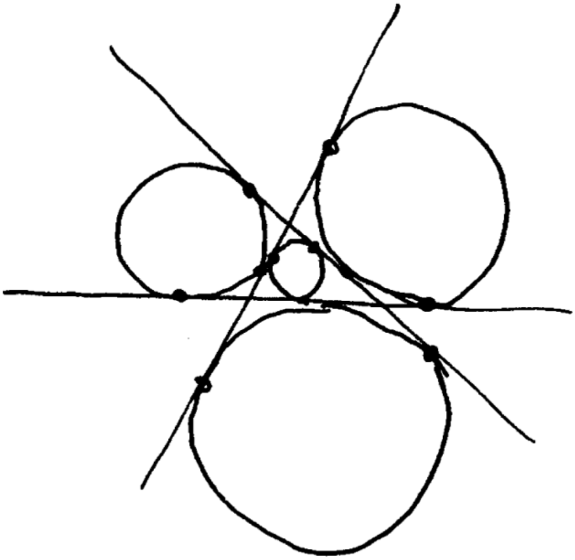


Fig. II.2.2. The four circles tangent to the three sides of a triangle

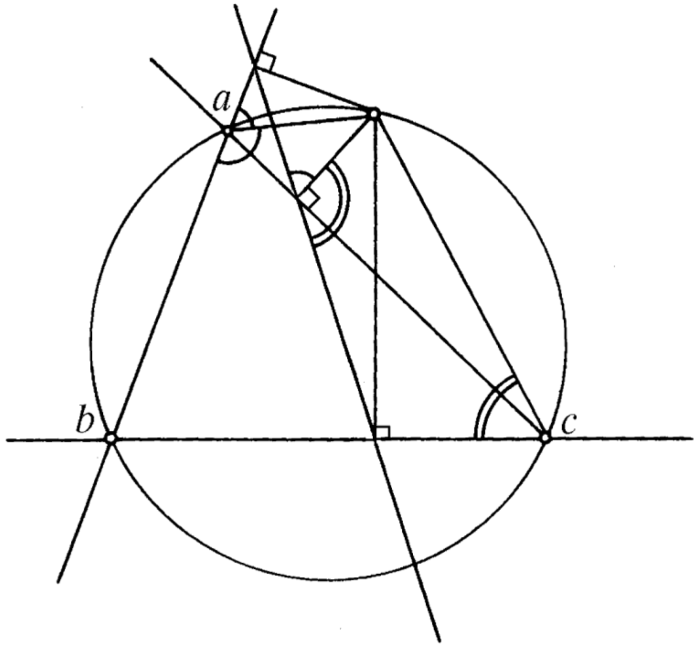


Fig. II.2.3. The Wallace-Simson line

♦

To our knowledge, the butterfly (Fig. II.3.1) has only limited interest: the correctness of the middle of this completely naive figure isn't easy to show and we leave conviction to readers. In fact it's a trick problem: the result is formulated in a metric (Euclidean) fashion, although it is really a theorem of projective geometry. In fact, let us introduce the polar D of the point a with respect to the circle C and consider the (harmonic) homology defined by the pair (a, D); it's an involutive projective transformation (see Sect. I.7). By construction, this homology preserves C and D and thus interchanges the two points p and q. But on the line D it is a homography that preserves the point at infinity, thus an affine transformation. Now, it preserves the midpoint a of pq: it's the metric symmetry of D with center a, and thus a is clearly also the midpoint of pq. Apart from this, to our knowledge, the butterfly theorem doesn't yield any movement on Jacob's ladder.

♦

The fact (Fig. II.3.2) that there exist four circles tangent to the three sides of any triangle in the Euclidean plane is not profound; we use the interior and exterior bisectors of the triangle (which gives a new way of obtaining a complete quadrilateral, cf. Sect. I.7). Nonetheless, we want to mention this figure, for its generalization to dimension 3 (and higher) sets a trap, pointed out to the author by his school mathematics teacher, the late Jean Itard, when he was 16; a trap into which practically all the mathematicians to whom the author has posed the problem have fallen, at least

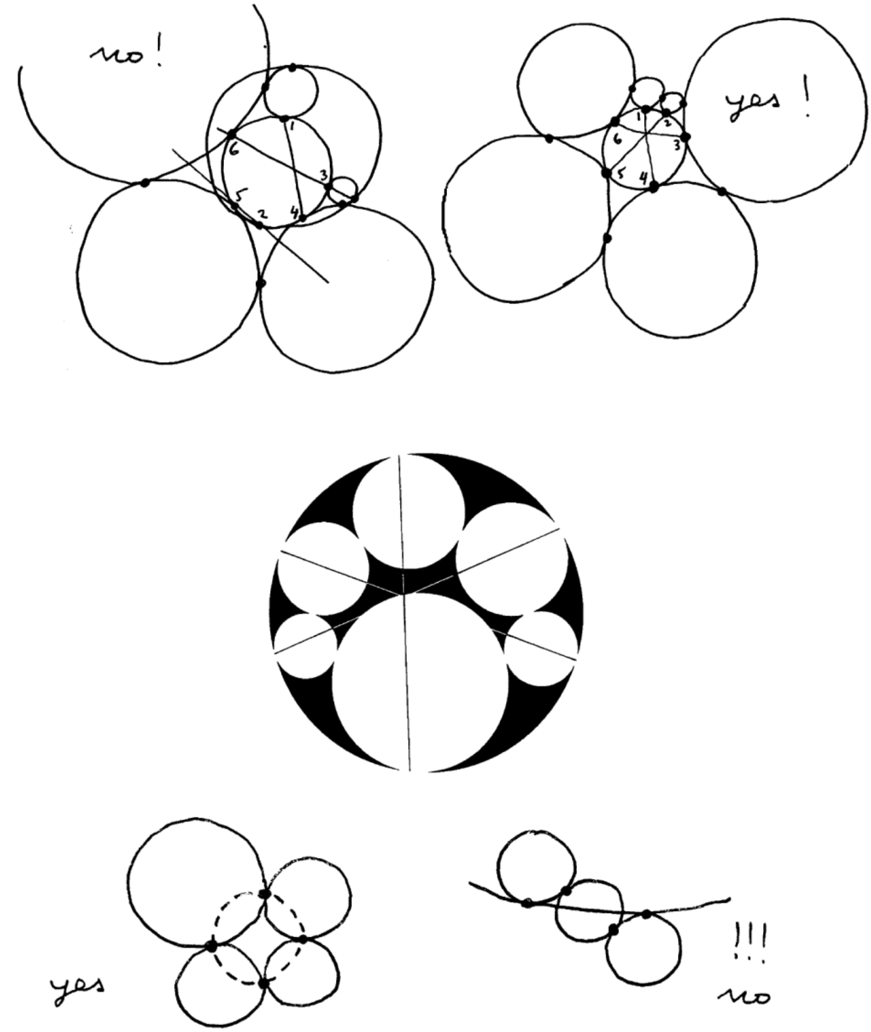


Fig. II.2.6. The seven circles theorem and that of five circles. Above and below, the conditions for the tangency of the circles are satisfied, but the lines are not coincident; or the four cocyclic points, and "one time in two"

can be found in detail in Sect. 10.6.8 of [B]; but we are going to give the essentials, for the result is fascinating and the proof trivial, once barycentric coordinates have been introduced. One of the appeals of this very elementary geometry problem is that, once the dimension exceeds 2, the number of spheres tangent to the faces of a simplex depends on the initial figure; for example, in dimension 3 this number depends on the areas of the faces of the tetrahedron considered; it is equal to 8 typically, but decreases to 5 for the regular tetrahedron. It is equal to 2^d in general. In dimension three, as far as pure geometry is concerned, it's a matter of looking at the

Name Index