Undergraduate Research Opportunities Programme in Science

Perspective in Mathematics and Art

By Kevin Heng Ser Guan
Department of Physics
National University of Singapore

Supervisor:
A/P Helmer Aslaksen
Department of Mathematics
National University of Singapore

Semester I 2001/2002

Table of Contents

I. The History and Theory of Perspective

II. Further Discussions on Perspective

III. The Moon in Art

IV. The Moon Illusion

V. Optical Illusions

VI. References

I. The History and Theory of Perspective

· Introduction

The philosophers of the Renaissance considered mathematical investigation integral to some of their theories. Such a combination of mathematics with natural philosophy was known as the “mixed sciences”, and this included the study of optics. The term “optics” is Greek in origin, and was introduced in the sixteenth century as part of the Renaissance determination to return to the Greek origins of science. The medieval name is the Latin term *perspectiva.* Unlike its contemporary counterpart, it was a complete science of vision, encompassing not only the nature and behaviour of light, but the anatomy and functioning of the human eye as well.

Initially, it was believed that sight was due to the active emission of “eye beams”, but it was later discovered to be the reception of light by the eye. In any case, the geometrical methods used to describe the way we see remain unchanged. Ignoring problems of physics and focusing on geometry, the theory of perspective describes how to project a three-dimensional object onto a two-dimensional surface. Prior to a more detailed discussion, it is necessary to understand what is meant by “cone” or “pyramid of vision”. Consider Fig. 1 shown below.

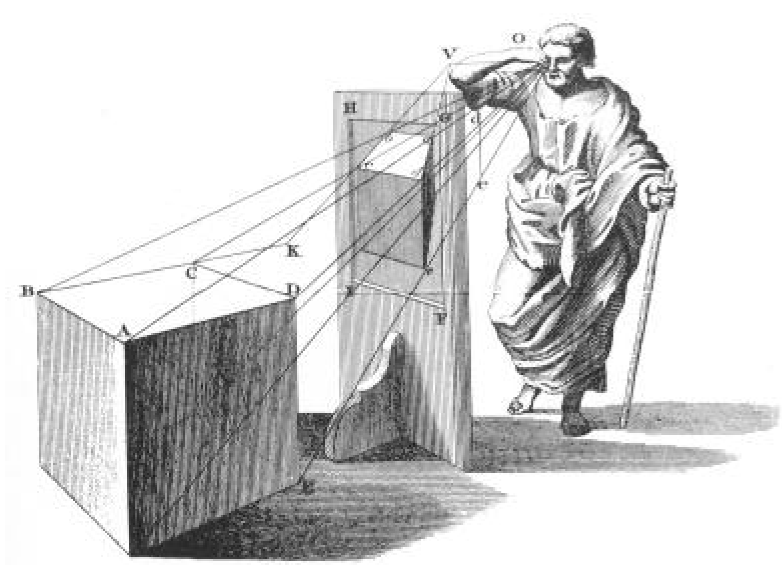


Fig. 1 – The pyramid or cone of sight is defined by the cube and the centre of rotation *O* of the eye of the spectator.

The centre of projection *O* is the centre of rotation of the eye, and the convergence of all light rays from the object. As a representation of these rays, straight lines are drawn from each point of the cube to point *O*. The intersection of each of these lines with the surface of projection *FGHI* forms the correct perspective of the cube in two dimensions. For instance, the point *A* is projected along the line *AaO* and *a* is its projection on the surface *FGHI*. Similarly, the corners *B*, *C* and *D* are projected on *b*, *c* and *d* respectively, such that the projection of the top face of the cube is *abcd*. Note that the visual angles subtended by the projections of these points on the surface *FGHI* are the same as those subtended by the cube. For example, the angle *AOB* is the same as the angle *aOb*.

Specifically, the perspective projection is the intersection of a plane with the pyramid or cone of sight. It should be noted that the view of the object in perspective is only valid for one particular viewing distance and location. When viewed in perspective, the object is said to be "degraded".

· Painting Before Perspective

Before the advent of perspective, it was generally accepted that the function of art was not naturalistic representation, but rather the expression of spiritual power. Artists tended to portray importance through size. An example is the mosaic shown in Fig. 2.

Fig. 2 – A mosaic with Christ made very large,
taken from scenes of the Last Judgement.
Notice the small figures near the feet of Christ.

A turning towards scientific naturalism began to appear with some force from the late thirteenth century onwards. The most striking examples of this newly naturalistic style are in the work of Giotto di Bondone (1266 – 1337). An example is his work "Joachim comes to the shepherds' hut", based on the story of Joachim and Anna (Fig. 3). The human and animal figures are all to scale, but the landscape has been shown rather small. It is clear that the picture is not shown in the correct perspective.

Fig. 3 – Giotto’s “Joachim comes to the shepherds’ hut”.
Fresco, Arena chapel, Padua.

Experts believe that Giotto is not constructing a landscape, but providing a landscape setting for the people in his story. Simple naturalism has taken second place to the demands of story telling.

· Filippo Brunelleschi

The quality of portraying real figures in real pictorial space is a distinct Renaissance characteristic. Ironically, the invention of mathematical rules for correct perspective came not from a painter, but from Filippo Brunelleschi (1377-1446), who was trained as a goldsmith. Brunelleschi made at least two paintings in correct perspective, but is best remembered for designing buildings and over-seeing the building works. Unfortunately, no writings on perspective by Brunelleschi have survived, and it is entirely possible that he in fact never wrote anything on the subject.

· Alberti’s Construction

The first written account of a method of constructing pictures in correct perspective is found in a treatise written by the learned humanist Leon Battista Alberti (1404 – 1472). The first version, written in Latin, was entitled *De pictura* (On painting), and was Alberti’s effort to relate the development of painting in Florence with his own theories on art. An Italian version (*Della pittura*) appeared the following year, and was dedicated to Filippo Brunelleschi. We shall look at how to properly construct a tiled floor or pavement. The details of construction are shown in Fig. 4, 5 and 6.

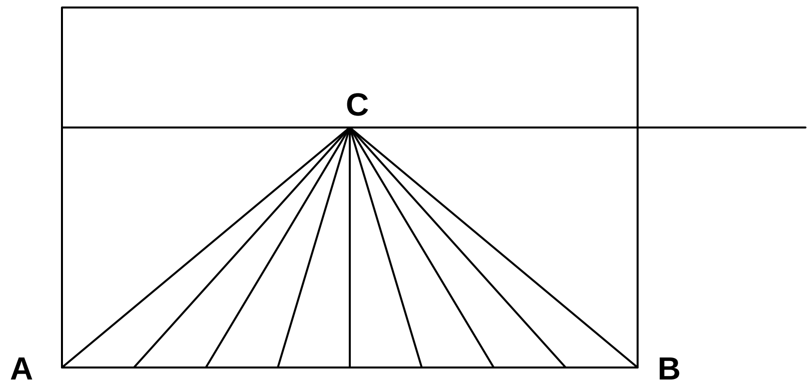


Fig. 4 – Choosing the centric point *C*.

First, the centric point *C* is chosen, and this is the point in the picture directly opposite the viewer's eye. It is also known as the "central vanishing point", the "point of convergence", or simply the "vanishing point". The ground plane *AB* in the picture is divided equally, and each division point is joined to *C* by a line. These are lines that run perpendicular to the plane of the picture, and are known as "orthogonals".

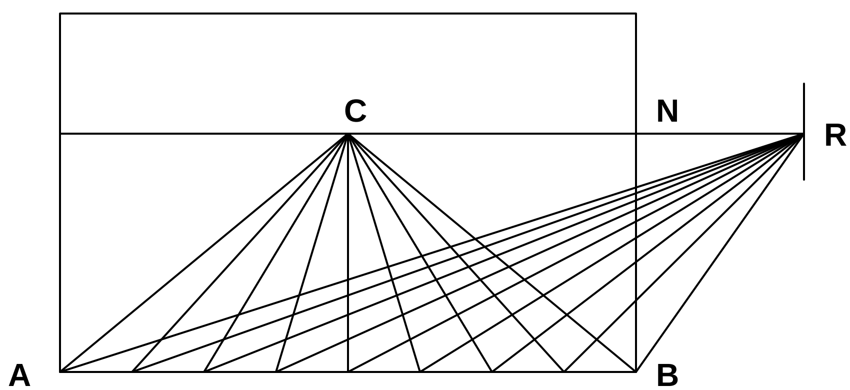


Fig. 5 – Choosing the right diagonal vanishing point *R*.

In Fig. 5, the point *R* is determined by setting *NR* as the "viewing distance". The "viewing distance" is how far the painter was from the picture. This is then how far a viewer should stand from the picture. *R* is known as the "right diagonal vanishing point".

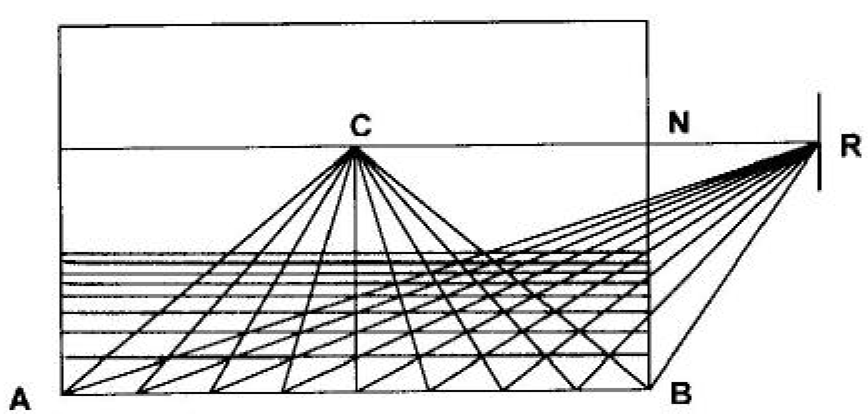


Fig. 6 – Alberti's construction.

The line *NB* is intersected by the lines converging at *R*. The final step is to draw lines perpendicular to the line *NB* from these intersection points. These are called the "transversals", and they run parallel to the ground line *AB* of the picture. They are also known as "horizontals".

Notice that the diagonal points of the squares in the grid can be joined by a straight line. This is an indication that Alberti's construction shows the ground plane of a picture in the correct perspective. Such mathematically correct "floors" are known as *pavimenti*, since most of the first treatises on their constructions were written in Italian. Notice that the spacing of the square grids gets smaller as the lines of the grid get farther away from the viewer. This is known as foreshortening.

Intuitively, it feels as if the distance *CR* is the viewing distance, and not *NR*. We can obtain a geometrical proof of why the correct viewing distance is *NR* by considering Fig. 7.

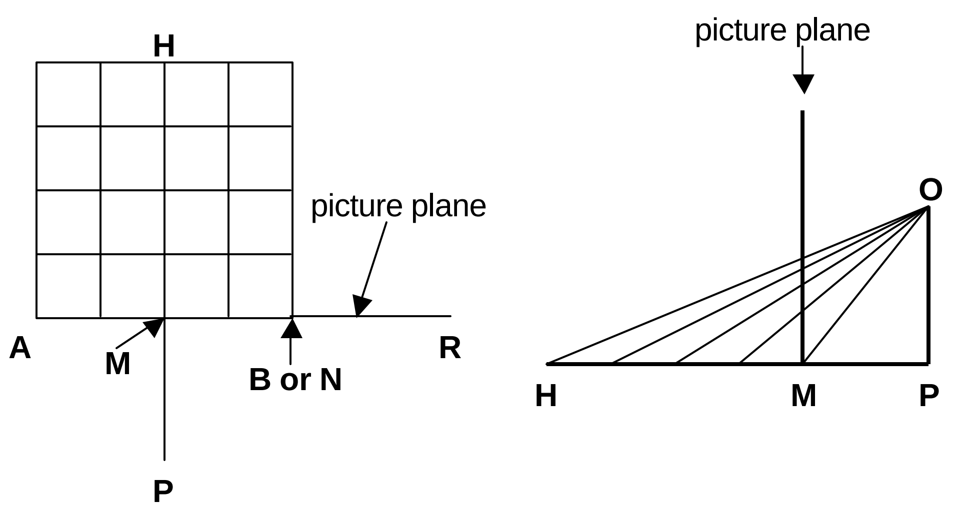


Fig. 7 – The plan and vertical section corresponding to Alberti's construction.

O and *P* are the positions of the eye and foot respectively, and are derived from the Italian words for eye and foot, namely *occhio* and *piede*. *MP* is the viewing distance. This is closely related to the *HP*, which is the distance from the viewer to the last transversal. If one was to stand at *R* instead of *P*, one can easily see that the distances *AR* and *HP* are the same. Hence, the viewing distances *NR* and *MP* are the same.

· Distance Point Construction

Historical records show that besides Alberti's construction, there were other methods for constructing *pavimenti*. One of them was known as the distance point construction, and was found in the treatise of Jean Pelerin (1445 – 1522), also known as the Viator. Entitled *De Artificiali Perspectiva*, it was first published in Toul in 1505 and later pirated at Nuremberg in 1509. It produces the same results as Alberti's construction, but constructs the *pavimenti* differently.

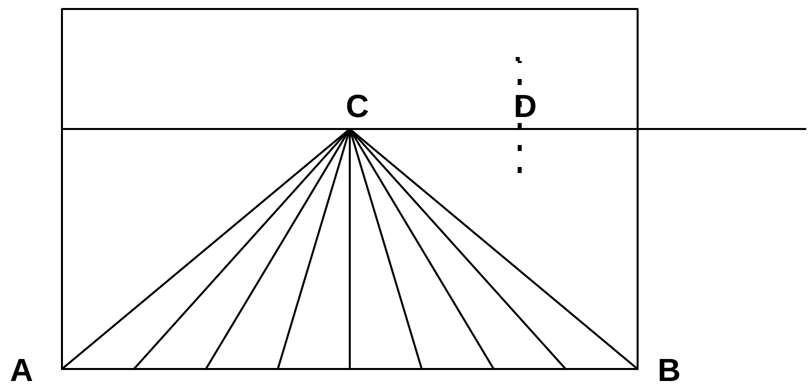


Fig. 8 – Choosing the distance point *D.*

As before, the ground line *AB* is divided equally, and each of these division points are joined to the centric point *C*. Next, the distance point *D* is chosen. The distance *CD* is the viewing distance.

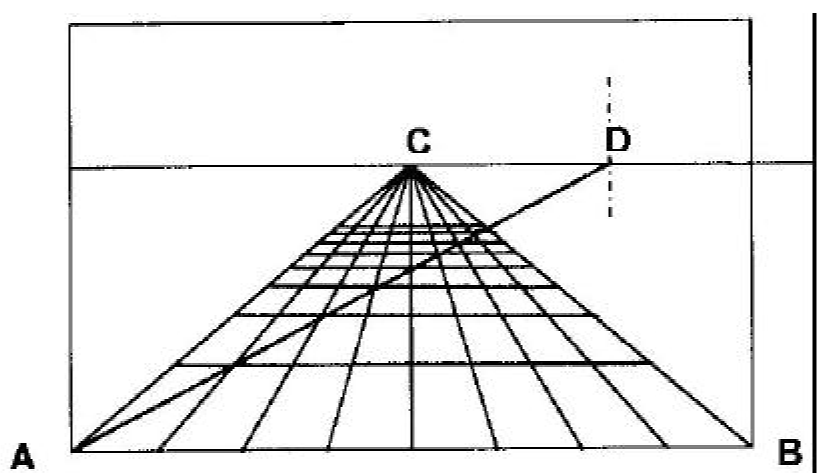


Fig. 9 – The distance point construction.

The line *AD* will intersect all the orthogonals. These intersection points are used to draw the transversals.

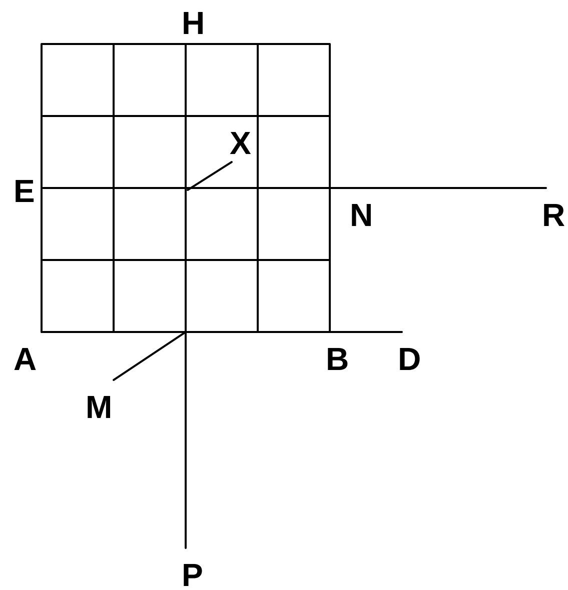


Fig. 10 – The plan and vertical section corresponding to the distance point construction.

We will next formulate a geometrical proof for the distance point construction. Fig. 10 shows the floor plan for the distance point construction. This is similar to Fig. 7. For Alberti's construction, it is the distance from the viewer to the last transversal which is of great importance. We can picture the line *ER* as the line *HP* rotated 90° anticlockwise about the point *X*. Hence, we show that *MP* is equal to *NR*.

For the distance point construction, we can picture the line *MD* as the line *MP* rotated 90° anticlockwise about the point *M*. *D* is the distance point. If one was to stand at *D* instead of *P*, one can easily see that the distances *MP* and *MD* are equal. It follows that *MD* is also the viewing distance.

We have shown that *MD* and *NR* are the correct viewing distances. Hence, *MD* and *NR* must be the same length. It follows that Alberti's and the distance point construction are equivalent.

To obtain a three-dimensional "proof" of the distance point construction, consider a square tile shown in the diagram below.

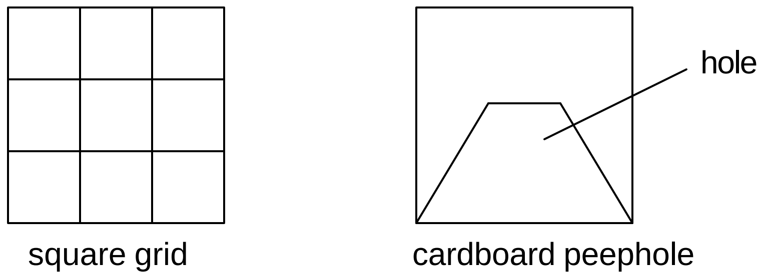


Fig. 11 – Three-dimensional setup of the distance point construction.

Suppose we have a “cardboard peephole”, which is a piece of cardboard with the trapezium cut out. Now we wish to position the square grid behind the cardboard peephole, in such a way that the entire grid can be seen through the trapezium hole.

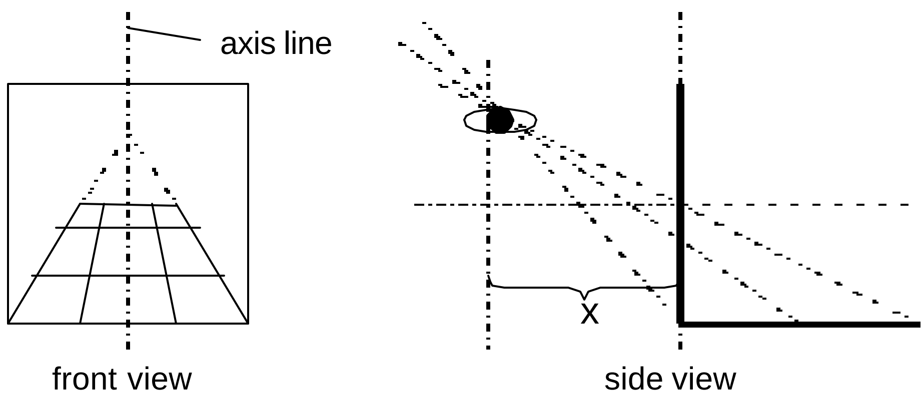


Fig.12 – Viewing the grid through the cardboard peephole.

The viewing distance is x. Now imagine that the cardboard peephole is rotated about its axis line, as shown in Fig. 13.

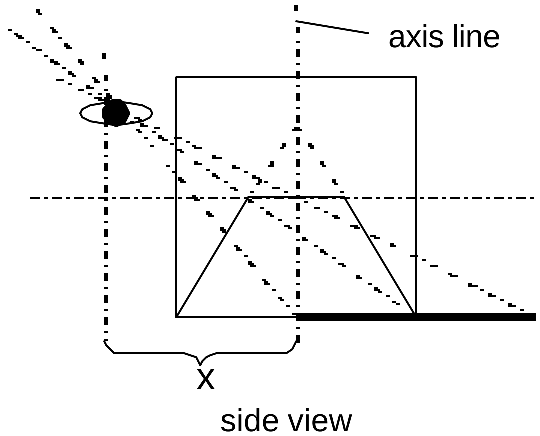


Fig. 13 – Notice that the viewing distance *x*
is due to the distance point construction.

Now imagine that the cardboard peephole is displaced to the right, such that it sits exactly on top of the grid (Fig. 14).

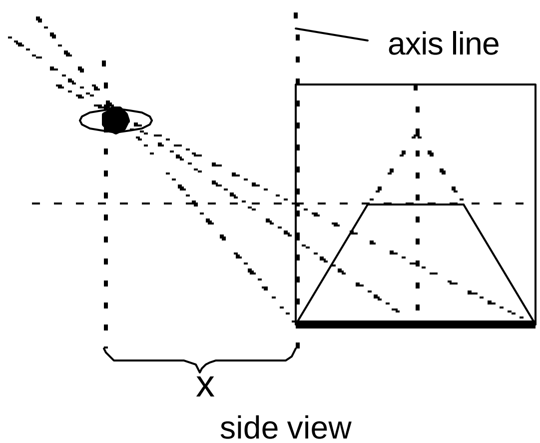


Fig. 14 – Notice that the viewing distance *x*
is due to Alberti's construction.

After the displacement, notice that *x* is now the viewing distance obtained from Alberti's construction. Though hardly a rigorous proof, it suggests a way of obtaining viewing distances from both methods using physical models, and provides an intuitive "feel" for the distance point construction. Most importantly, the model shows that Alberti's and the distance point construction are equivalent, as they yield the same viewing distances.

· A Physical Model of Alberti's vs. the Distance Point Construction

A physical model was built, according to Figs. 11, 12, 13 and 14. The square grid is 12cm by 12cm, with each of the squares measuring 1cm by 1cm. The cardboard peephole measured 12cm by 15cm, with the central vanishing point 10cm above the base. We pick the viewing distance to be 26.3cm. The top of the trapezium is then 3.5 cm from the bottom.

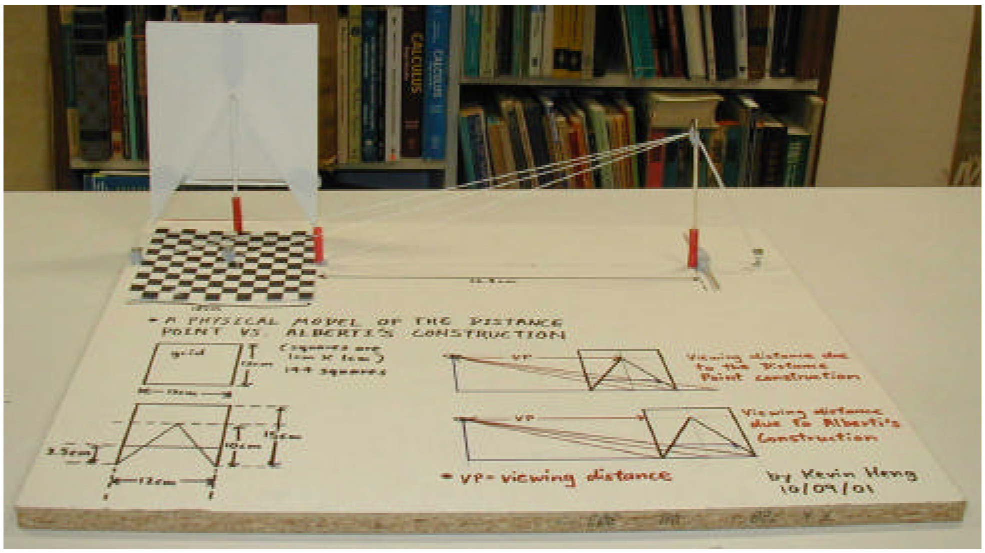


Fig. 15 – A physical model of Alberti's vs. the distance point construction.

The model is designed such that two cardboard peepholes can be mounted simultaneously to show the relation between Alberti's and the distance point construction.

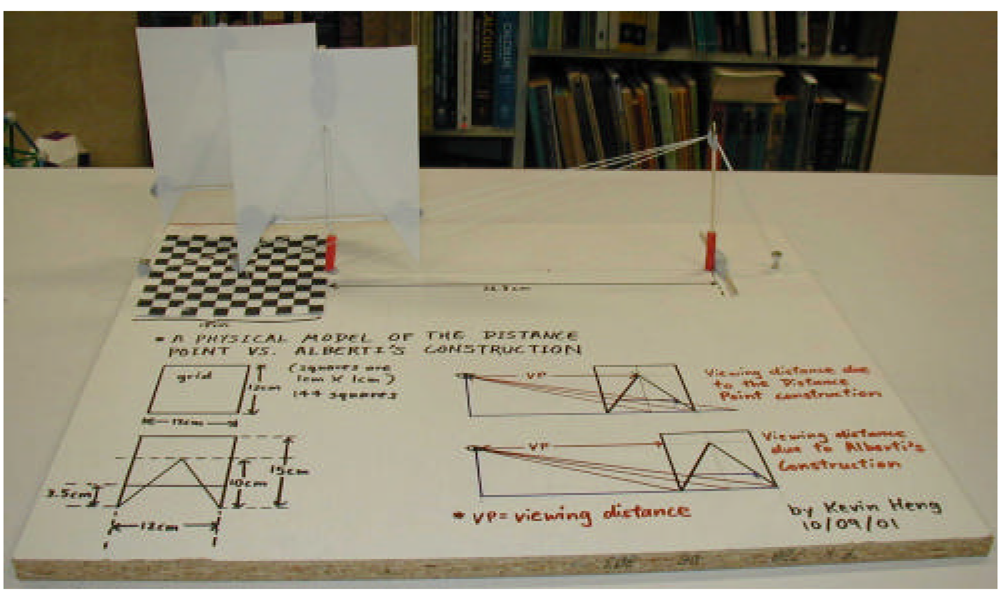


Fig. 16 – Model with two cardboard peepholes.

Placing the eye at the correct viewing point, 10 cm above the base, will show the square grid in the correct perspective. This is shown in Fig. 17.

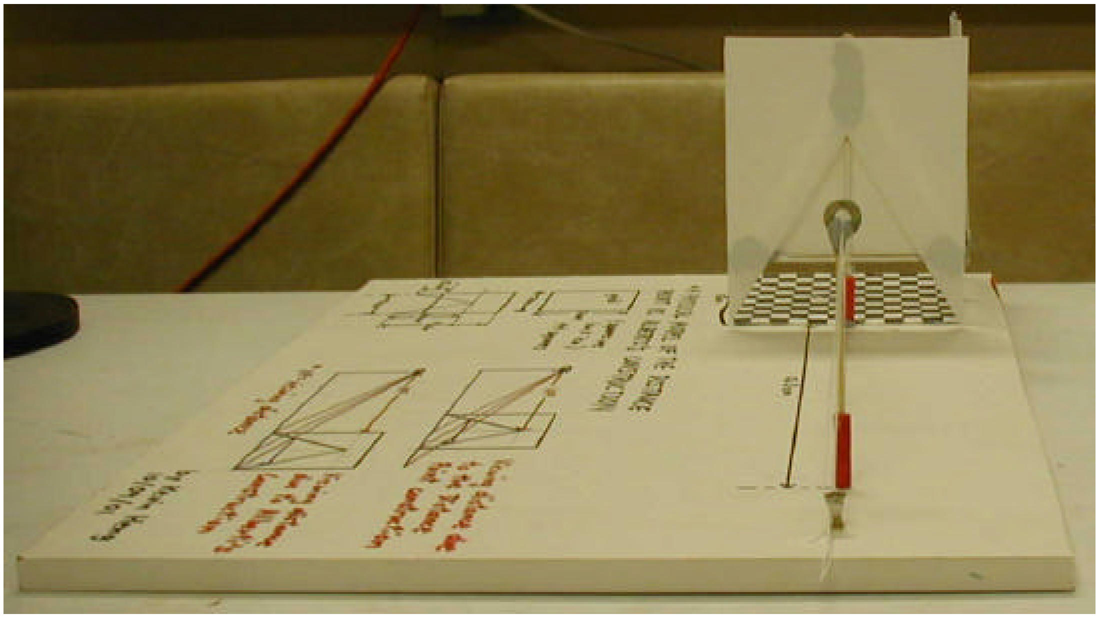


Fig. 17 – The square grid viewed in perspective.

This idea of investigating perspective is at least a few hundred years old. In fact, Albrecht Durer (1471 – 1528), too, substituted sight lines with string when he was constructing the image of a lute in perspective. This is shown in Fig. 18.

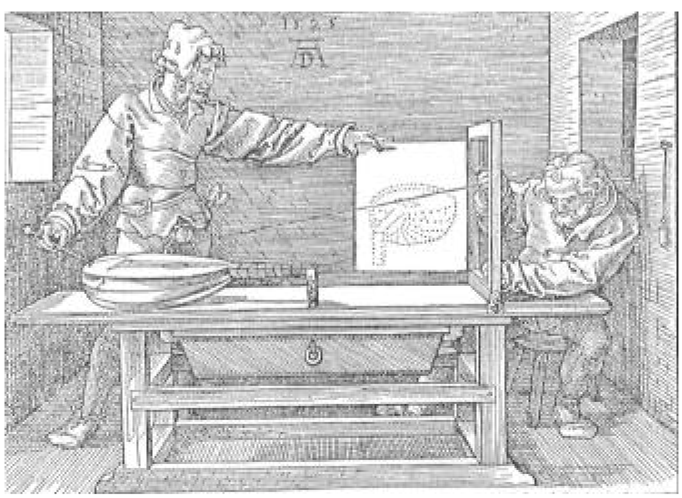


Fig. 18 – The use of strings to construct a perspective image of a lute. Taken from Albrecht Durer's "Treatise on measurement with compasses and straightedge (*Underweysung der Messung mit dem Zirkel und Richtscheyt*, Nuremberg, 1525).

In Fig. 18, the pointer held by the man on the left marks where the light ray ends. The man on the right marks the position where the ray intersects the picture plane by adjusting strings stretched across the rectangular wooden frame fixed vertically to the table. This is seen more clearly in Fig. 19. Once the crossing strings are in the appropriate position, the string *LNGH* is moved away and the panel with the drawing paper is swung into the picture plane so that the new point can be marked on it. In this way, all the necessary coordinate points for drawing the lute in perspective can be obtained.

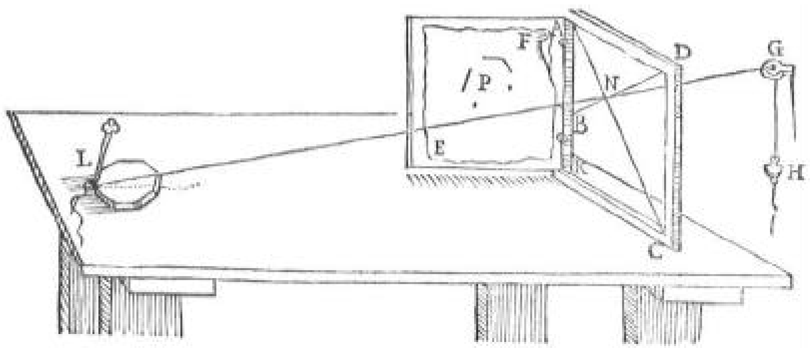


Fig. 19 – The instrument Durer used for drawing a lute.

· Piero della Francesca

Piero della Francesa (1412 – 1492) is acknowledged as one of the most important painters of the fifteenth century. Besides being a talented painter, he was also a highly competent mathematician. Piero wrote three treatises: "Abacus treatise" (*Trattato d'abaco*), "Short book on the five regular solids" (*Libellus de quinque corporibus regularibus*), and "On perspective for painting" (*De prospectiva pingendi*).

Piero's perspective treatise is believed to be the first of its kind. It was concerned not with ordinary natural optics, but with what was known as "common perspective" (*perspectiva communis*), the special kind used by painters. Piero felt that this new part of perspective should be seen as a legitimate extension of the older established science. He divided his treatise into three books, and his first original theorem appeared in one of them (Fig. 20).

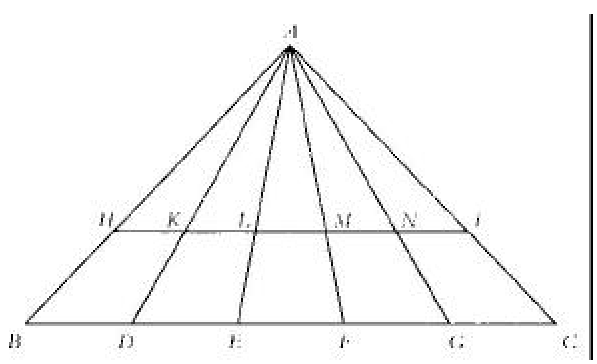


Fig. 20 – Proposition 8, Book 1 of
Piero della Francesca's *De prospectiva pingendi*.

A given straight line *BC* is divided into several parts. Another line *HI* is drawn parallel to the first. The theorem states that if lines from the points dividing *BC* converge at point A, *HI* will be divided in the same proportion as *BC*.

The proof uses similar triangles. Since the two lines are parallel, $AK/AD = HK/BD$. Also, $AK/AD = KL/DE$. It follows that $HK/BD = KL/DE$. Since $BD = DE$, this means $HK = KL$. The same reasoning can be extended to *LM*, *MN* and *NI*.

Indeed, so influential was Piero's treatise that many following it were merely simplifications of it. It was found that examples used in many treatises on perspective were derived from Piero's work. Furthermore, it is possible Piero had shown that Alberti's and the distance point constructions are equivalent. Nevertheless, we will attempt to prove this in the next section.

· Equivalence of Alberti's and the Distance Point Construction

The following is a direct, geometrical proof of the equivalence of Alberti's and the distance point construction. Specifically, this means that the viewing distances *CD* and *NR* are the same.

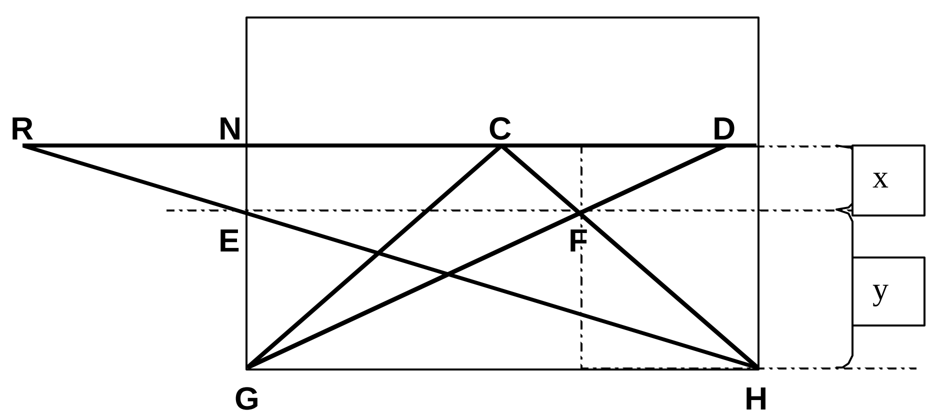


Fig. 21 – Alberti's construction and the distance point construction superimposed.

The left portion of Fig. 21 shows Alberti's construction, with the *NR* being the viewing distance. The right portion has the viewing distance as *CD*, a result of the distance point construction.

Since the lines *RD* and *GH* are parallel, triangles *CDF* and *GFH* are similar. Hence,

$$\frac{CD}{GH} = \frac{x}{y}$$

In addition, triangles *RNE* and *EHG* are also similar.

$$\frac{NR}{GH} = \frac{NE}{EG}$$

However, *NE* = *x and EG* = *y*. It follows that

$$\frac{CD}{GH} = \frac{NR}{GH}$$

Hence, *CD* is the same length as *NR*.

· Historical Footnotes on the two Constructions

As late as the 1620s, it was still not recognized that Alberti's and the distance point constructions were actually equivalent. In fact, Pietro Accolti misunderstood part of a serious mathematical discussion of perspective by Giovanni Battista Benedetti and believed that out of the two constructions, only Alberti's was correct. Accolti accordingly gave the Albertian method the name *costruzione legittima* or "legitimate construction", which unfortunately stuck till this day. It is perhaps an irony that the treatise in which Accolti coined this misnomer is entitled *Lo inganno de gl'occhi* or "Deceiving the eyes".

It is unclear whether Alberti was aware of the distance point construction, though he proposed using the straight diagonal line to check if his method was correct. He also believed that the centric point should be at the eye level of the figures in the painting, though many painters do not follow this rule. Nevertheless, the centric point is the focus of attention of the painting. It is usually the job of a curator to position the paintings such that viewers can obtain the right eye height and viewing distance.

Furthermore, there is a practical advantage of using the distance point construction. By definition, the right vanishing point *R* is always outside the edge of the painting, whereas it is possible for the distance point *D* to be within the painting. Hence in practice, it is easier to determine *D*. In fact, renowned mathematician Egnazio Danti (1536 – 1586) once wrote: "And first we shall consider the (rule that is) more well-known and easier to comprehend, but longer and more tedious in application; and in the second (part) we shall consider the (rule that is) more difficult to comprehend, but easier to follow (in practice)". As expected, the "first rule" refers to the Albertian construction, while the "second rule" is that for the distance point construction.

· An Incorrect Method of Constructing *Pavimenti*

Some authors refer to a way of *pavimenti* construction called the "2/3 method", which was in used by painters till about 1435. We will now see why it is incorrect. In Fig. 22, the distance between the ground line and the first transversal is taken to be *l*. The distance between the first and second transversal is then $(2/3)l$. It follows that the distance between the n-th and the $(n+1)$-th transversal is $(2/3)^n l$. However, one only has to join the diagonals in all the grids to observe that the method is incorrect, as the diagonals formed is curved.

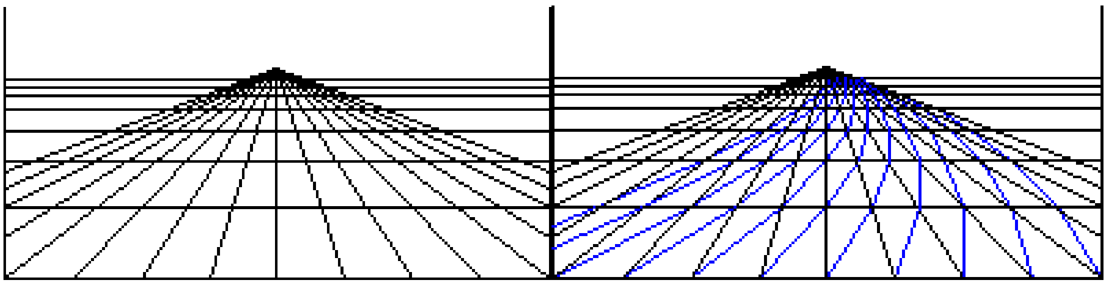


Fig. 22 – The "2/3 method"

• Case Study: "The Flagellation of Christ" by Piero della Francesca

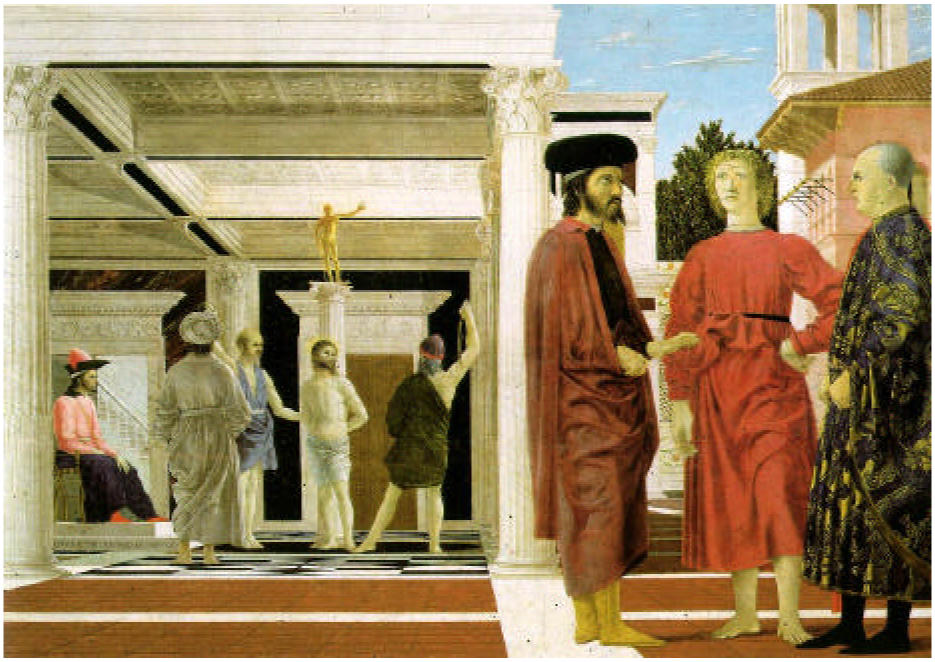


Fig. 23 – "The Flagellation of Christ" by Piero della Francesca.

"The Flagellation of Christ" by Piero della Francesca is one of the most striking examples of a painting done in the correct perspective. Using the distance point method, the viewing distance AO, as shown in Fig. 24, is measured to be about 1.1 times the height of the picture.

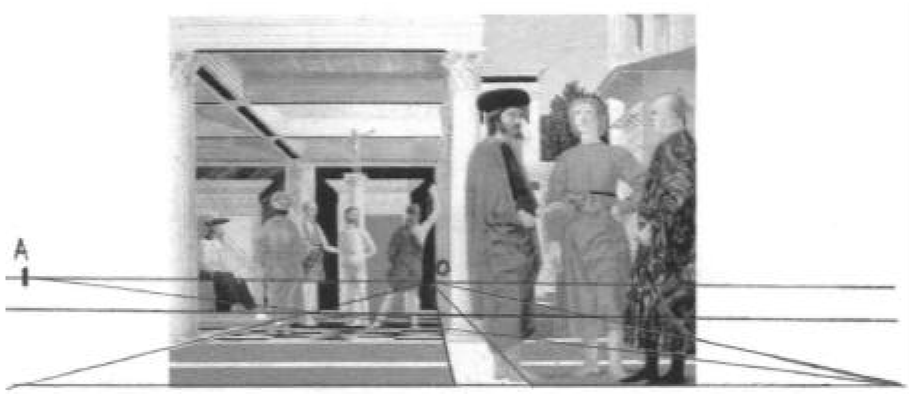


Fig. 24 – Determination of the viewing distance using the distance point method.

Incidentally, the "Flagellation" is also one of the few paintings where one can determine the viewing distance easily. In most other paintings, the architecture which extends over most of the painting is drawn in perspective as a whole, from one main centre of projection. However, the human figures and other objects are drawn with their own centres of projection.

For example, in his "School of Athens", Raphael (1483 – 1520) depicted two spheres at the right side of his painting (Fig. 25) and drew their outlines as circles, not ellipses. This is an instance of the modifications made to the strict theory of central projection. In addition, the individual human figures in Fig. 25 have their own centres of projection.

Fig. 25 – "The School of Athens" by Raphael.

· Federico Commandino

Federico Commandino (1509 – 1575) had a reputation as a learned humanist, a linguist of ancient languages, and reasonably competent mathematician. His rare combination of talents made him the foremost editor and translator of Greek mathematics, including the works of Archimedes, Apollonius of Perga and Pappus. At the time of publication (1558 – 1588), these texts were also of much intrinsic mathematical interest. By making them available to mathematicians, Commandino's editions exercised considerable influence over the development of mathematics in the later sixteenth and seventeenth centuries.

For instance, Johannes Kepler credited his use of approximation methods in his invention of the "calculus of indivisibles" to procedures found in the work of Archimedes. It was found that the edition used by Kepler was that of Commandino.

· The Theory of Conics

Conics are curves obtained when a cone is intersected by a plane in a certain way. They include ellipses, circles, parabolas and hyperbolas.

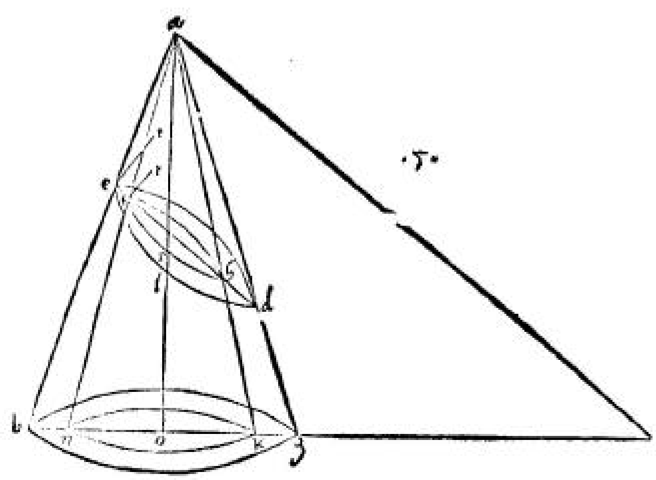


Fig. 26 – An ellipse is formed when the cone is intersected by a plane.

Conics belonged almost exclusively to the world of learned mathematics, and treatises before the sixteenth century rarely mentioned them as they were considered difficult subjects. A rare treatise on conics was written by Albrecht Durer, entitled "Treatise on measurement with compasses and straightedge (*Underweysung der Messung mit dem Zirkel und Richtscheyt*, Nuremberg, 1525).

Johannes Kepler (1571 – 1630) was one of the first to give a greater degree of unity to the theory of conics. He developed the idea that the properties of conics will change slowly from one type to the next (Fig. 27). For instance, as the foci move closer together an ellipse gradually becomes a circle. Then as the ellipse gets longer and its foci moves further apart, it gradually becomes a parabola.

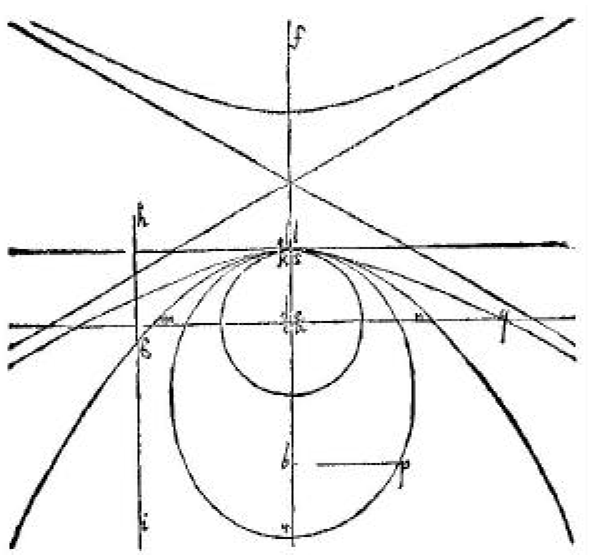


Fig. 27 – A plane system of conic sections by Johannes Kepler.

Giovanni Battista Benedetti (1530 – 1590) further proposed that if a cone is cut by two planes parallel to one another, then the two conic sections will be similar (Fig. 28).

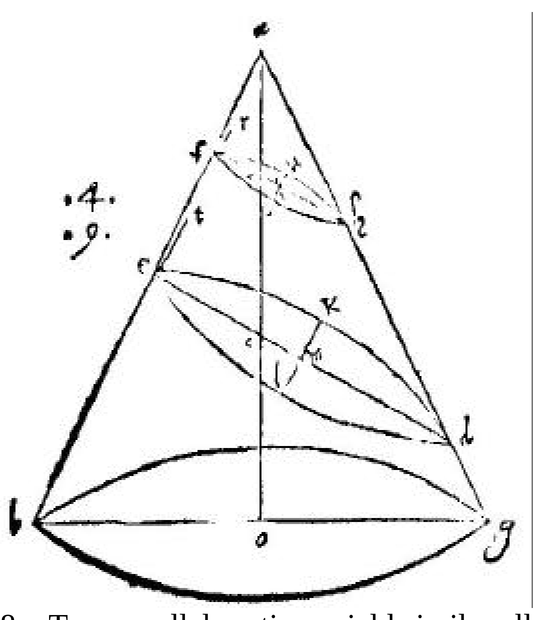


Fig. 28 – Two parallel sections yield similar ellipses.

Specifically, Fig. 28 shows ellipses with the same eccentricity. Benedetti also proved that a plane intersecting two cones with different vertical angles will yield two different conics.

· Girard Desargues and Projective Geometry

Born to one of the richest men in France, Girard Desargues (1591 – 1661) is credited with unifying the theory of conics. A highly competent and original mathematician, he conceived problems in three-dimensional terms, just like Benedetti did fifty years earlier. The difference was that unlike Benedetti, Desargues fully appreciated the power of this method, and used it to give a projective treatment of conics.

Desargues wrote his most important work, the treatise on projective geometry, when he was 48. It was entitled "Rough draft for an essay on the results of taking plane sections of a cone" (*Brouillon proiect d'une atteinte aux evenemens des rencontres du cone avec un plan*). In the course of his work, he became recognized as the first mathematician to get the idea of infinity properly under control, in a precise, mathematical way. For example, Desargues defined a line as one to be produced to infinity in both directions, while a plane was similarly taken to extend to infinity in all directions. A set of parallel lines were defined as a set of concurrent lines whose meeting point lied at an infinite distance.

One of the most important parts of his treatise can be viewed as a generalization of a theorem proved by Piero della Francesca (Fig. 20). Desargues' theorem allowed a way of defining the pattern of division even when the second line was not parallel to the first. In a sense, Desargues was generalizing perspective into becoming a technique of use to mathematicians.

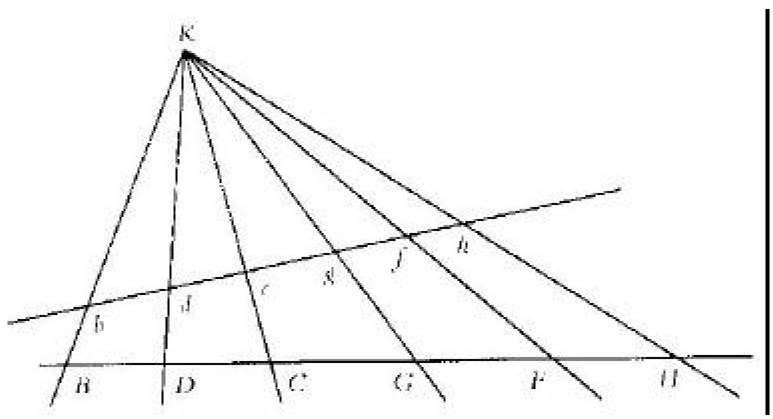


Fig. 29 – Diagrams for Desargues' proof of the theorem that if we have six points in involution, *BDCGFH,* then their images under projection from a point *K* onto another line, *bdcgfh,* will also be six points in involution.

Desargues is not looking at what is changed by perspective, as artists were, but is instead looking for what is unchanged. Perspective is not being seen as a procedure that "degrades" but merely as one that transforms, leaving certain mathematical relationships unchanged (invariance). These are now what we term the "projective properties" of the figure concerned. Specifically, Desargues' theorem states that six points in involution are projected onto another six points in involution.

II. Further Discussions on Perspective

· The Column Paradox

Fig. 30 – Photograph of the Basilica at Paestum. The Basilica is a Greek Doric temple dating from the sixth century B.C.

Fig. 30 shows a photograph of the Doric columns at Paestum, near Naples, Italy. The row of columns is parallel to the plane of projection. Curiously, the columns exhibit the counter-intuitive property that those further away from the centre of projection appear wider, and not thinner. This apparent paradox can be explained by Fig. 31.

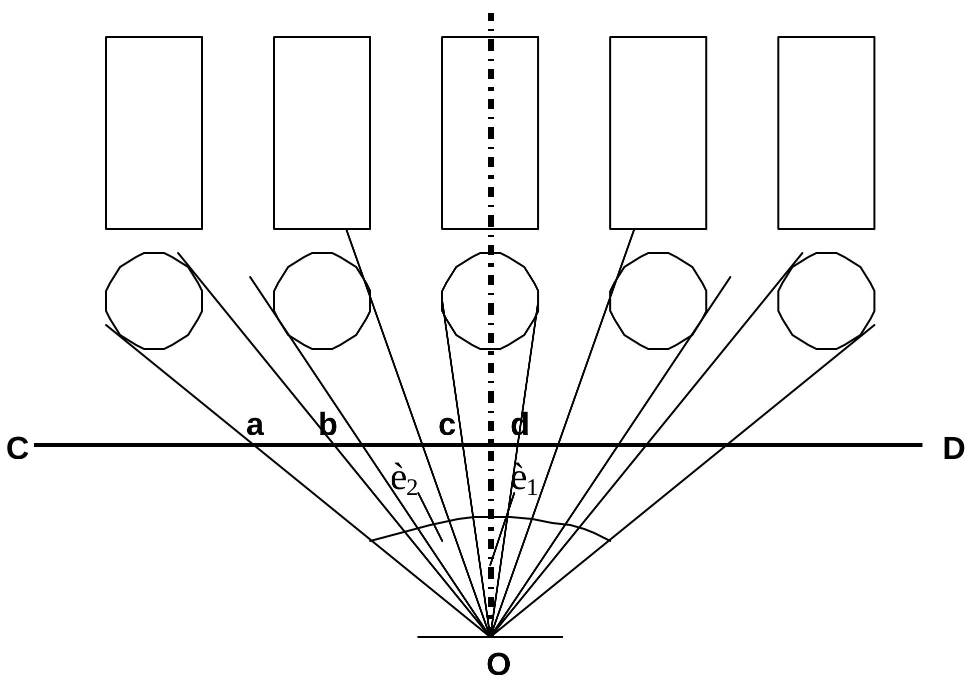


Fig. 31 – Schematic diagram of the Doric columns, with the centre of projection *O*. The surface of projection is the line *CD*.

The angle subtended at the centre of projection *O* by the width of the columns becomes smaller for columns further away from the centre. For example, θ_2 is smaller than θ_1. However, the lengths due to the intersection of the limiting light rays of the columns by the projection surface *CD* become wider for columns further away from *O*. This is due to the increasing obliquity of the light rays. For instance, *ab* is longer than *cd*. Ultimately, this means that columns further away from *O* will project a larger width onto the projection surface *CD*.

In practice, most artists omit this weird perspective effect from their paintings, instead giving equal width to all their columns. However, a row of square pillars does not give rise to this problem. As shown in Fig. 32, when projected onto the projection surface *EF*, the rectangular fronts of the pillars all yield projections of the same shape and size. This is because the projection of a plane figure onto a plane parallel to it always yields a figure that is geometrically similar to the original. The proof for Fig. 32 uses similar triangles, and was discussed under the section on Piero della Francesca (Fig. 20).

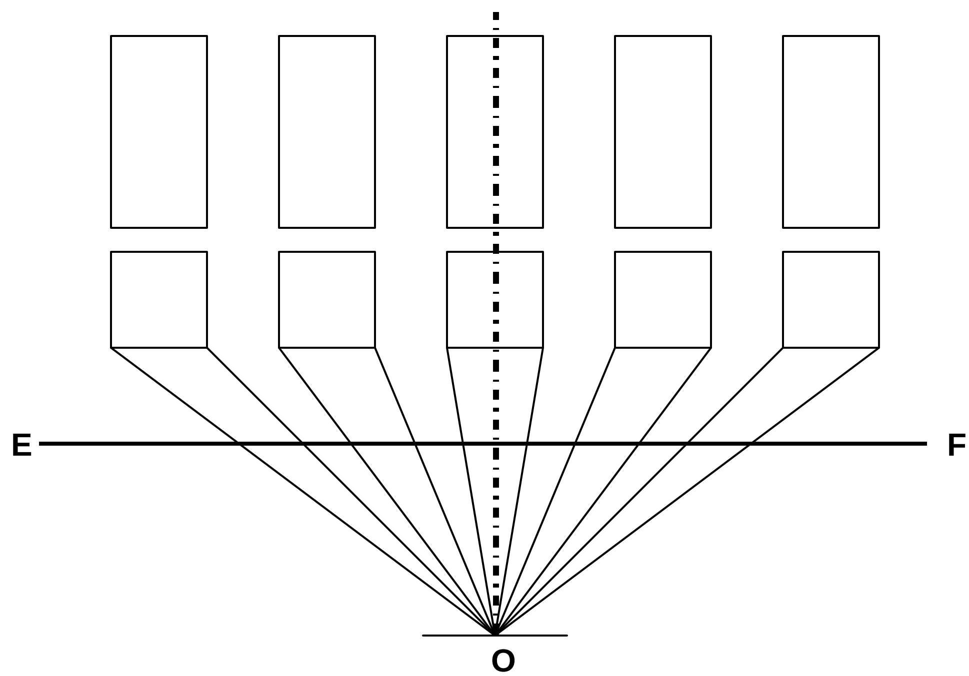


Fig. 32 – Schematic diagram of a row of square pillars. The centre of projection is at *O*, while the projection surface is *EF*.

The problem arises for columns because their surfaces are curved, and our eyes find it difficult to distinguish what is the “front” of each column. By contrast, even if we look at a square pillar from an oblique angle, it is easy to tell how much of the “front” we are observing.

· The Parthenon

Fig. 33 – The Parthenon in Greece.

Erected between 447 and 438 B.C., the Parthenon in Greece was designed by Iktinos, when Greek architecture was at the height of its sophistication. Due to the presence of optical illusions, the Parthenon has what are known as "optical refinements" built into its structure. It must be stressed that these illusions are physiological and psychological in nature. They are not geometrical effects.

To the unaided eye, columns tend to look narrower in the middle than at the top or bottom. Each of the columns in the Parthenon was built with a slight bulge in the middle, to make them appear "straight". Columns tend to "contract" near the top, and hence the base of each column was built a little thicker. Columns further away from the centre appear thicker. To counteract this effect, the columns in the centre were built a little thicker. Furthermore, the spacing between the columns appear smaller towards the centre. Therefore, they were spaced wider apart accordingly.

Horizontal lines appear to "dip" in the middle, and hence the centre portion of the floor was slightly raised. Furthermore, the columns were slanted inwards so that they would meet if they were extended one mile into the sky. This to counteract the effects of hatched-line illusions (Fig. 64). The triangular outline of the roof makes the top part of each column appear to slant outwards.

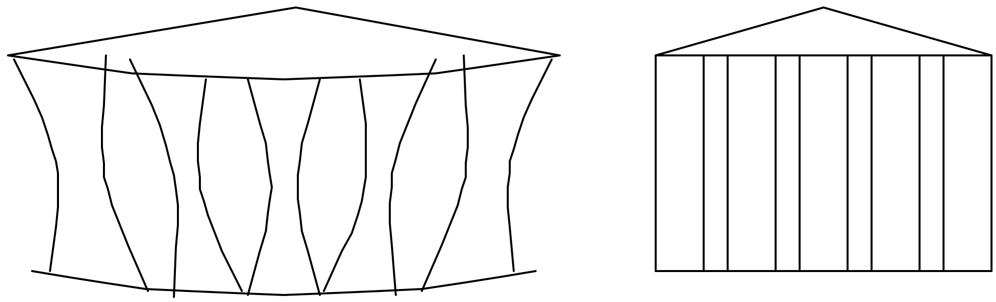


Fig. 34 – Schematic diagram of how the Parthenon would appear before (left portion) and after (right portion) optical refinements.

The left portion of Fig. 34 shows how the Parthenon would appear before the optical refinements. The optical illusions shown are grossly exaggerated for effect. After the corrections were made to the columns and floor, the Parthenon now appears “correct”, as shown in the right portion of Fig. 34. It is interesting to note that none of the “straight lines” seen in the Parthenon are geometrically straight. Such concepts of making architecture appear to look “correct” are known as “counter-perspective”.

· The Palazzo Spada

The Palazzo Spada in Rome, built in 1638 by Francesco Boromini, is an illuminating example of how perspective can be used to deceive the eyes.

Fig. 35 – A view of the Palazzo Spada in Rome.

From the view in Fig. 35, it appears as if the pathway is rather long, leading to a statue at the end. The statue is about three-quarters the height of the distant doorway. However, when a person who has the same height as the statue stands at the near doorway, we find that she is only about a quarter the height of the near doorway (Fig. 36).

Fig. 36 – Comparing the heights of the doorways.

This puzzling phenomenon can only be explained if one realizes that the two doorways are of vastly differing heights. However, as the tunnel seems to be constructed in perfect perspective, we are tricked into believing that both doorways are of the same height.

· A Mathematically Incorrect Regular Icosahedron

In 1985, Branko Grunbaum pointed out that the logo used by the Mathematical Association of America (MAA) was mathematically impossible. The image used was a regular icosahedron shown in the left portion of Fig. 37. The three lines shown are parallel in three dimensions. If rendered by orthogonal projection from 3D to 2D, they will remain parallel. If rendered in perspective, they will all intersect at a single vanishing point. Clearly, this logo is wrong. Mathematics was not used to draw it. The right portion of Fig. 37 shows a correctly drawn regular icosahedron.

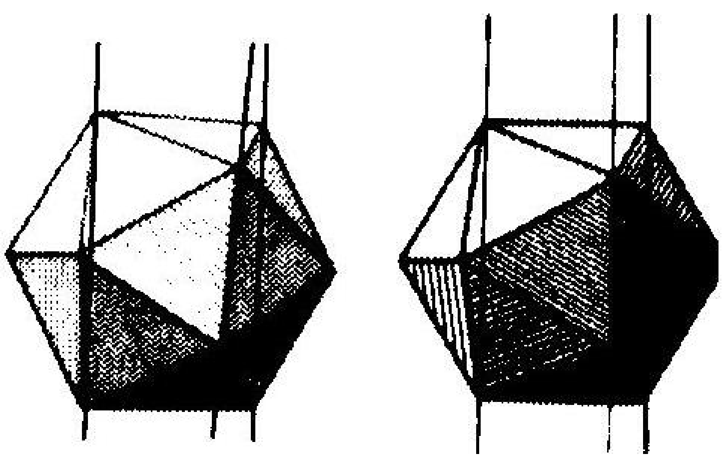


Fig. 37 – The left portion shows an incorrectly drawn regular icosahedron.
The right portion shows a correctly drawn regular icosahedron.

· A Photograph of a Photograph

When a photograph of a photograph is taken, curious perspective effects occur. This is vividly illustrated in Fig. 38.

Fig. 38 – A photograph of a photograph.

The bald man appears deformed as we are viewing it well away from its center of projection and viewing location. As discussed in Part I, an image in perspective is meant to be viewed only from one point. However, if we look at the other man in the figure from a position away from the center of projection of Fig. 38, he will not appear deformed.

When the shape and position of the picture surface can be seen by the observer, an unconscious intuitive process of psychological compensation takes place. This restores the correct view even when the picture is viewed from the wrong position. In fact, this important characteristic has been noted by Albert Einstein.

· Pozzo's Ceiling

Usually, an observer is able to distinguish between a real geometrical entity and a pictorial representation of it. However, when the distance between the observer and the picture surface is very large, the human eye can be tricked into believing that aspects of the picture are real geometrical entities. A famous example is the ceiling in the church of St. Ignazio in Rome. Painted on the ceiling is a very convincing scene of St. Ignatius being received into Heaven, the masterpiece of Fra Andrea Pozzo (1642 – 1709).

The ceiling is hemicylindrical in shape and is about 30 metres off the floor. When one is at the correct viewing location (Fig. 39), the painted architecture appears in three dimensions as an extension of the real architecture. It is impossible to determine where the ceiling surface actually is, much less distinguish the painted architecture from the real ones.

Fig. 39 – Pozzo's painting on the ceiling in the church of St. Ignazio.
This is a photograph taken from the correct viewing location.

However, when one observes the ceiling well away from the correct viewing location, it becomes possible to detect the hemicylindrical nature of the ceiling (Fig. 40).

Fig. 40 – Pozzo's ceiling away from the correct viewing location.

Notice that the painted columns in Fig. 40 appear deformed. It is interesting to note that although deformations appear when an observer walks away from the correct viewing location, the illusion of depth remains. This is because the distance between the observer and the ceiling remains high, and the eye is unable to determine the picture surface. When the ceiling is observed from a point 20 metres above the floor, the illusion of depth disappears (Fig. 41). The painted ceiling now appears hemicylindrical, as it should be, and the painted columns look severely deformed.

Fig. 41 – Pozzo's ceiling as seen from a point 20 metres above the floor.
Notice that the three-dimensional illusions have vanished.

· Multiple Vanishing Points

It is possible for objects in perspective to have more than one vanishing point. When an objective has two vanishing points, it is said to appear in two-point perspective. Similarly, three vanishing points correspond to three-point perspective.

When we view an object, we can usually make out which is the "front" of it. We shall call the surface of this "front" the front plane. The surface which the object is projected onto is called the projection plane (or the picture plane). Both the observer and the object rest on the ground plane.

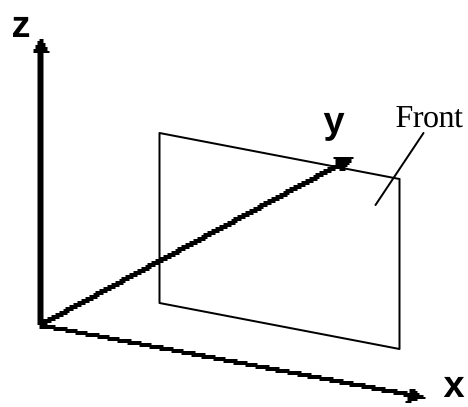


Fig. 42 – Diagram showing the front plane, projection plane and the surface plane.

In Fig. 42, the projection plane is taken to be the *xz*-plane. The ground plane is the *xy*-plane. The front plane is represented by the rectangle. Hence, the front plane is always parallel to the projection plane, and perpendicular to the ground plane. When this occurs, only one vanishing point is obtained when the object is viewed in perspective. This is called one-point perspective.

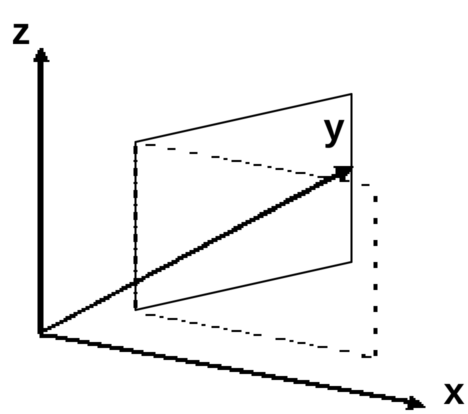


Fig. 43 – Rotating the front plane about the *z*-axis.

When the front plane is rotated about the *z*-axis, it is no longer parallel to the projection plane. However, it is still perpendicular to the ground plane. When the front plane is viewed in perspective, two vanishing points will be observed. This is shown in Fig. 43.

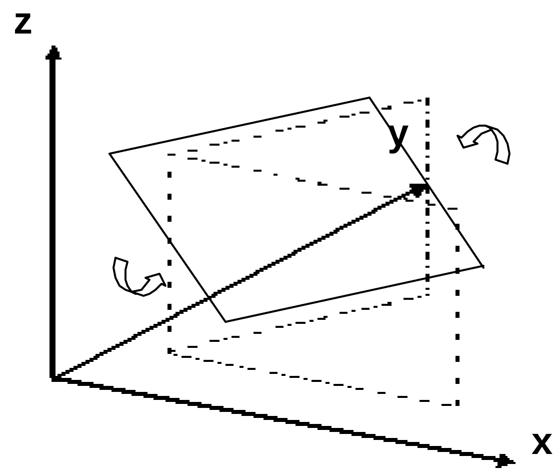


Fig. 44 – Further rotating the front plane.

If the front plane is further rotated as shown in Fig. 44, it will no longer be perpendicular to the ground plane. It is also not parallel to the projection plane. The combination of these two factors will yield three vanishing points, when the front plane is view in perspective.

Consider a cube placed in the *xyz* coordinates axes. First let the "front" of the cube be placed parallel to the projection plane. In perspective, the cube will appear as shown in Fig. 45.

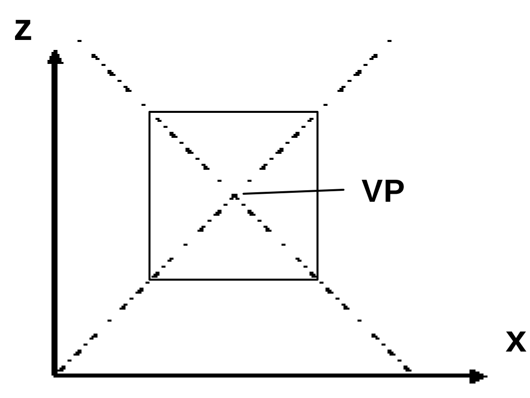


Fig. 45 – A cube in one-point perspective.
The vanishing point is denoted by *VP*.

Next consider the cube with its "front" rotated about the z-axis. The "front" will not be parallel to the projection plane, but will still be perpendicular to the ground plane. In perspective, the cube will appear as shown in Fig. 46.

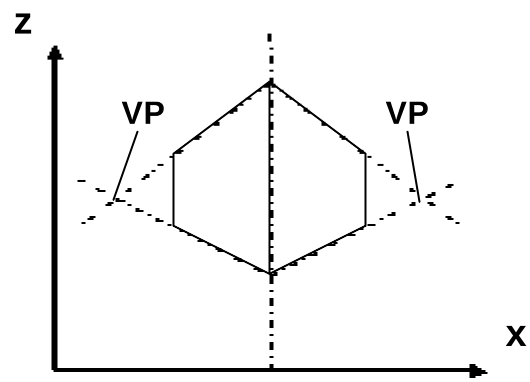


Fig. 46 – A cube in two-point perspective.

One can further orientate the cube such that its "front" is neither parallel to the projection plane, nor perpendicular to the ground plane. The cube will then be viewed in three-point perspective, as shown in Fig. 47.

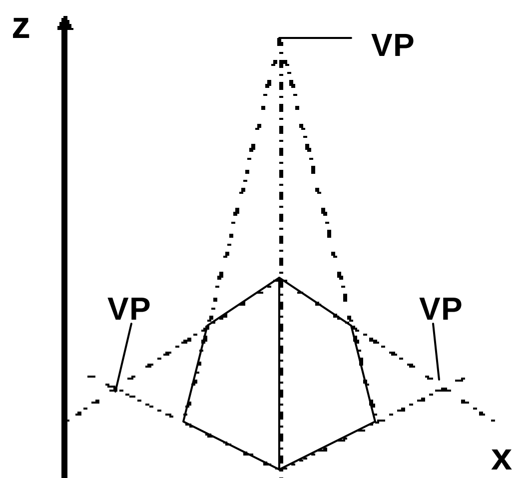


Fig. 47 – A cube in three-point perspective.

The properties governing two- and three-point perspective can be generalized to more complicated objects.

III. The Moon in Art

It is a well-known fact that the moon occupies only a tiny portion of the hemispherical vault of the heavens. Yet there are many artists who exaggerate the size of the moon on purely aesthetic grounds. There is a simple mechanism by which we can check if a painting has its moon painted the correct size.

· Subtending an Angle

Prior to a discussion of the approximation method, it is imperative to understand what is meant by "subtending an angle". Consider an observer shown in Fig. 48. The object under scrutiny is in the observer's pyramid or cone of vision. The angle subtended by the object is then defined as the angle θ between the two limiting rays to the eye.

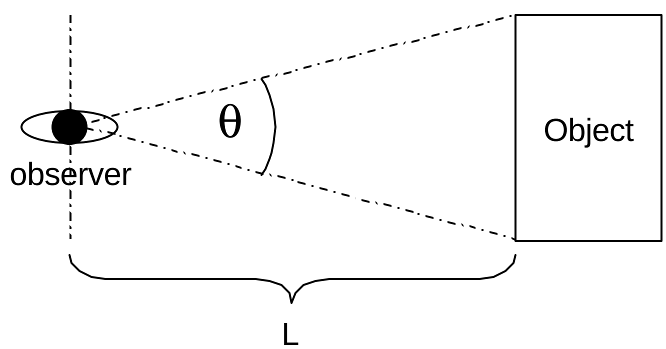


Fig. 48 – Angle subtended by an object.

It follows that as the distance L between the observer and the object decreases, the angle subtended by the object increases. Hence, θ is inversely proportional to L.

· Estimating the Moon Size

We first assume that the moon has been painted the correct size. As the distance from the moon to the earth is almost constant, the angle subtended by the diameter of the moon is a constant 0.5°. This forms the entire basis of our arguments.

Next, we select an object in the painting that we can estimate the size of in real life. We then compare it to the diameter of the painted moon. Let the diameter of the moon in the painting be D. If the width or height of the object in the painting is nD, the angle subtended by it is $n(0.5°)$.

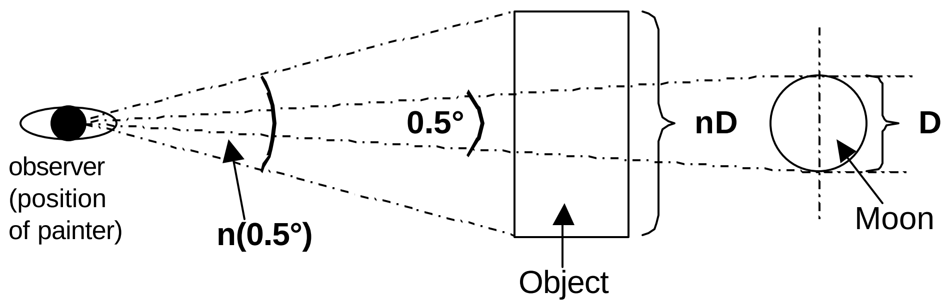
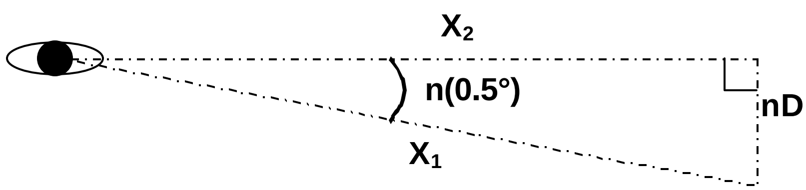


Fig. 49 – Angle subtended by an object relative to the moon.

We next make a judicious guess of nD in real life. This value of nD, along with the value of $n(0.5°)$, will give us the distance from the painter to the object (X_1 or X_2).



$$\tan[\, n(0.5°) \,] = nD / X_2$$

$$\sin[\, n(0.5°) \,] = nD / X_1$$

Fig. 50 – Calculating the distance between the painter and the object.

If the angle $n(0.5°)$ is small, $\tan[n(0.5°)] \approx \sin[n(0.5°)]$. It follows that $X_1 \approx X_2$. Hence, either X_1 or X_2 will give a reasonable approximation of the distance between the painter and the object.

If the distance from the painter to the object turns out to be ridiculously large, we can immediately suspect that the painter has represented the moon the wrong size.

· Case Study: “The Bluestocking” by H. Daumier

The previous arguments are vividly illustrated by an examination of “The Bluestocking”. It is one of many paintings by celebrated nineteenth-century French artist and caricaturist Honore Daumier, who had a reputation as a draughtsman of the highest calibre.

Fig. 51 – “The Bluestocking” by H. Daumier.

We estimate the height of the woman to be between 5 and 6 feet in real life. It follows that the window is about 3 feet wide. The width of the window in the painting is measured to be about 4 times the diameter of the painted moon. Hence, the angle subtended by the window is about 2°.

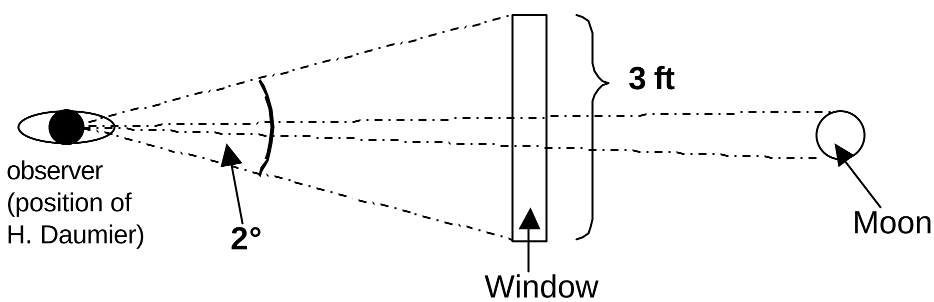


Fig. 52 – Estimation of how far Daumier is from the window.

The distance between Daumier and the window works out to be about 86 feet! The woman is sitting in a small room and it is hard to imagine that there exists 86 feet of space between the artist and the window. It appears that Daumier had made his moon much bigger than it actually was.

To find out how large the moon should be, we perform our calculations backwards. Suppose Daumier is about 20 feet away from the window. It follows that the angle subtended by the window is between 8.53° and 8.63°. The diameter of the moon should then be about 17 times smaller than the width of the window.

Fig. 53 – Suggested corrected moon size for "The Bluestocking" by H. Daumier.

The corrected painting shows that the artistic impact of the moon is lost, and it now looks very trivial indeed. One can conclude that Daumier deliberately drew a much larger moon for aesthetic reasons.

· A Relation Between Moon Size and Horizon Distance

The moon in four paintings was analyzed: H. Daumier's "The Bluestocking", Samuel Palmer's "Coming from Evening Church", Vincent van Gogh's "The Sower" and Henri Rousseau's "Carnival Evening". The moon diameters of the first three paintings were found to be 17, 25 and 100 times too big. Only Rousseau's work showed the moon in its correct size.

Interestingly, it was found that the lower the moon on the horizon, the larger the amount of exaggeration. It has been suggested that the degree of exaggeration was inspired by the famous moon illusion (discussed in Part IV). However, it should be noted that the exaggerations were first and foremost motivated by aesthetic reasons, with the moon illusion perhaps playing a part.

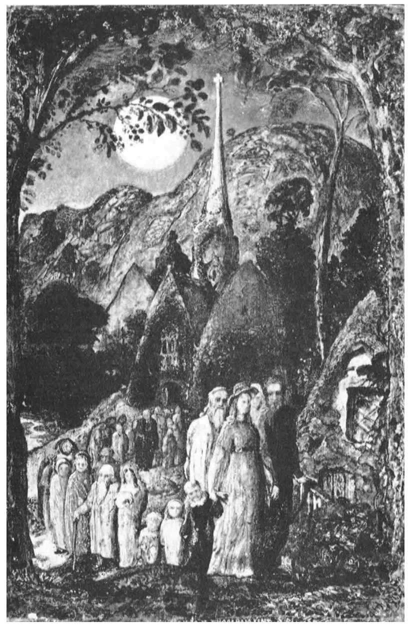


Fig. 54 – Samuel Palmer's "Coming from Evening Church".

Fig. 55 – Henri Rousseau's "Carnival Evening".

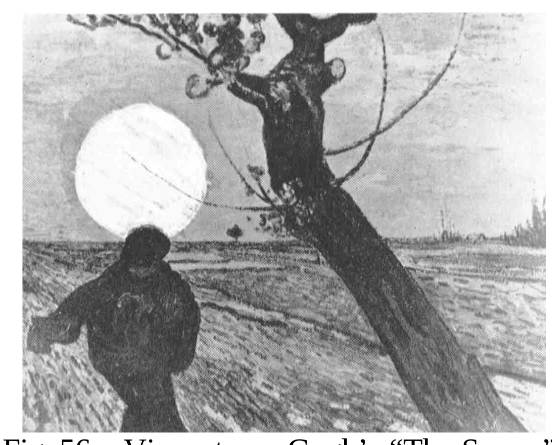


Fig. 56 – Vincent van Gogh's "The Sower".

IV. The Moon Illusion

It is a well-known phenomenon that an observer will perceive the moon closer to the horizon as being larger than one high up in the sky. However, when one photographs both the horizon and zenith moons and compares them, they turn out to be the same size. This is the famous moon illusion.

Psychologists have been studying the moon illusion for more than a hundred years, and there are many conflicting theories. A popular one suggests that we visualize the celestial dome around us as a rectangular “roof” rather than a hemispherical one. This is called the “flat sky” theory. Hence, looking at the horizon moon is akin to looking at a Ponzo illusion (Fig. 63) upside down.

The most widely accepted theory is by American psychologist Don McCready. Along with a few other prominent researchers, McCready believes that the moon illusion is due to a size illusion known as *oculomotor micropsia*. However, before we examine this theory, we need to differentiate between the various types of moon illusions.

· Linear and Angular Size

When one says the moon is of a certain “size”, it is important to clearly define what that means. In Fig. 57, the three spheres shown subtend the same angle è to the observer. Hence, they are the same angular size. However, spheres farther away from the observer are of a larger linear size.

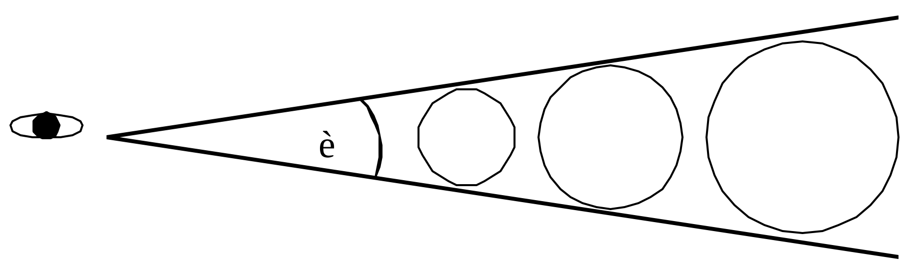


Fig. 57 – The three spheres all subtend the same angle to the observer.

In Fig. 58, the spheres are of the same linear size, but the one nearer to the observer subtends a larger angle.

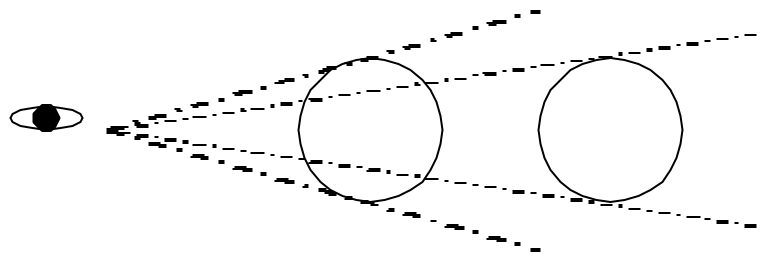


Fig. 58 – The spheres are the same linear size, but subtend different angles to the observer.

These two concepts of size are intimately related to the distance between the observer and an object, which is known as the linear distance. It is important to note that angular size is measured in degrees, while linear size is measured in metres.

To complicate matters, there is a difference between an object's absolute linear size and its perceived linear size. For instance, the moon has a diameter of about 2160 miles and this is known as its absolute diameter or absolute linear size. It is very difficult, if not impossible, for the unaided eye to know the moon's absolute linear size. Instead, an observer would make a subjective approximation known as the perceived linear size. Similarly, an approximation for the moon's angular size is known as the perceived angular size. Note that the moon always subtends 0.5° to an observer. Finally, the distance from an observer to the moon is about 238,800 miles and this is its absolute linear distance. The corresponding approximation would be the perceived linear distance.

· Types of Moon Illusions

There are three types of moon illusions known to psychologists. The first type is when the horizon moon appears angularly larger than the zenith moon, with both moons about the same distance away. Hence, the horizon moon is necessarily of a larger perceived linear size.

The second type of illusion occurs when both moons appear to be of the same linear size, but the horizon moon is perceived to be angularly larger. It follows that the horizon moon is perceived to be closer to the observer. The third type is a combination of the first and second. Studies show that most people perceive the horizon moon to be angularly larger, linearly larger and closer.

· *Oculomotor Micropsia*

The angular size illusion has defied explanation for decades. McCready suggests that the moon illusion is a special case of a lesser-known phenomenon known as *oculomotor micropsia*, which literally means “looking small”. It is intimately related to *oculomotor macropsia*, which translates as “looking large”.

It is believed that the way we focus our eyes is affected by changes in the visible patterns near an object. These patterns or secondary objects are known as distance cues. *Oculomotor macropsia* is an angular size illusion caused by changes in the activity of eye muscles. When one views the horizon moon, the details in the landscape form distance cues that sends a “very far” signal to the observer. This has the effect of making the eyes focus at a distance much greater than where the object is, called overfocusing. Hence, a smaller angular size is perceived.

Oculomotor micropsia has the opposite effect. The zenith moon offers few distance cues, and the eyes tend to adjust to a relatively near position, known as the resting focus position. This is typically about one or two metres from the face. The perceived angular size becomes smaller. A closely related phenomenon is night myopia, where the eyes tend to underfocus in dark surroundings. This causes temporary nearsightedness, also known as “night blindness”.

It must be strongly emphasized that *oculomotor micropsia* and *oculomotor macropsia* are physiological and not geometrical effects. Furthermore, they occur when viewing all objects, not just the moon.

· Explaining the Moon Illusion

Oculomotor macropsia tricks us into believing that the horizon moon looks angularly larger. If we judge the two moons to be about the same distance away, we are further tricked into believing that the horizon moon has a larger linear size. If we somehow judge that the two moons are of the same linear size, we will believe that the horizon moon looks nearer to us than the zenith moon. In most cases, a combination of these factors occurs due to several different sets of distance cues competing for perceptual dominance and hence clouding our judgement. Psychologists call this cue-conflict.

V. Optical Illusions

Optical illusions are intimately related to perspective. Many judgements made in paintings require an evaluation of size and form, and optical illusions can often impede this process. The purpose of this section is to highlight the existence of numerous subtle, but very real, optical illusions.

· The Bisection Illusion

Consider an arrangement shown in the Fig. 59. Two lines are drawn such that the horizontal line is bisected by the vertical line.

Fig. 59 – The bisection illusion. The arrangement on the left has two lines of the same length. The arrangement on the right has the vertical line much shorter than the horizontal one.

On sight, it appears that the arrangement on the right uses lines of the same length. In fact, the vertical line is about 12.5% shorter than the horizontal one. The arrangement on the left uses lines of the same length, but the vertical line appears much longer than the horizontal one.

Many explanations have been suggested. A strong argument is that this is due to the curvature of the retina. Another is that the act of halving the horizontal line makes us lose sense of its full extension. A weaker argument is that our eyes are placed in a horizontal line, so we tend to assess horizontal and vertical lengths differently.

· Convergence-divergence: the Muller-Lyer Illusion

In 1860, Zollner proposed the optical mechanism of convergence-divergence. It states that our assessment of a geometrical quantity is strongly affected by the nature of the surrounding territory. This was first illustrated in 1889, in what is known as the famous Muller-Lyer illusion.

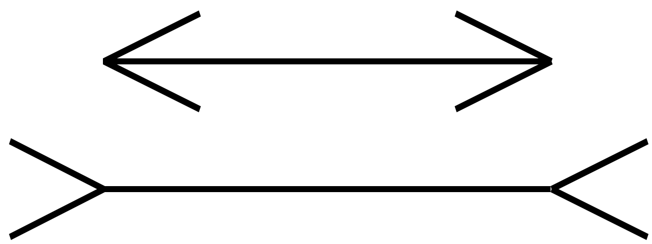


Fig. 60 – The Muller-Lyer illusion. The two horizontal lines are of equal length.

On sight, it appears that the line with the converging arrowheads is shorter than the one with the diverging arrowheads. Surprisingly, they are actually equal in length. Convergence has an effect of shortening a geometrical quantity, while divergence lengthens it. The combined effect yields the Muller-Lyer illusion.

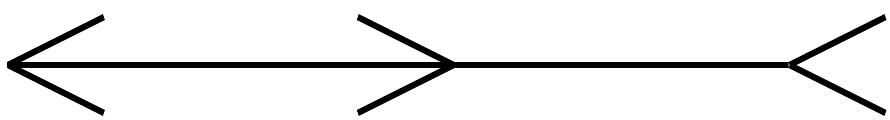


Fig. 61 – A modification of the Muller-Lyer illusion.

In Fig. 61, the line appears to be equally bisected. In fact, the portion with the diverging arrowheads is about 28% shorter.

Convergence-divergence can be exploited in dress design to give the appearance of a reduced waist. Schematically, we have the waistlines of two ladies in Fig. 62. In the upper waistline, the diverging lines on the dress make the waist appear bigger than it actually is. The lower waistline appears smaller than the upper one, due to convergence. In fact, both have the same size.

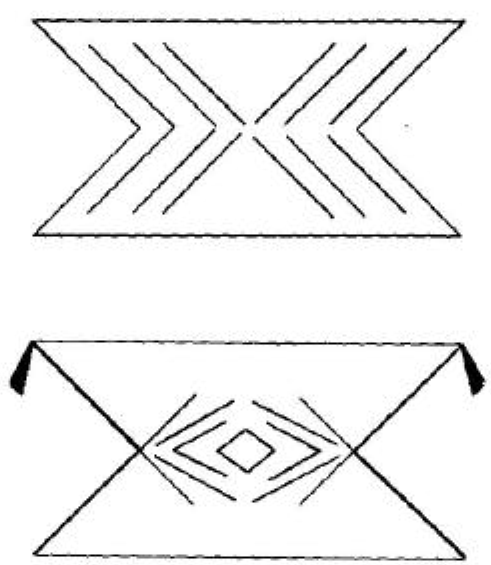


Fig. 62 – Dress patterns influence the apparent size of the waistline.

Another famous example related to convergence-divergence is the Ponzo illusion, first presented by Mario Ponzo in 1913.

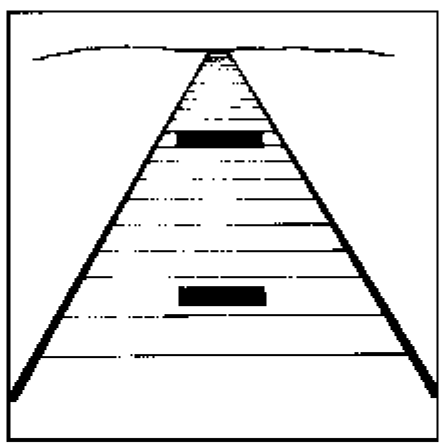


Fig. 63 – The Ponzo Illusion.

Fig. 63 shows a railway converging into the distance. Notice that the upper block appears bigger than the lower one. In fact, they are of the same size. The converging railway tracks have clouded our judgement of size.

· Hatched-line Illusions

Related to the convergence-divergence concept are the illusions produced by hatched lines. These illusions excited much interest and discussion over the last century. The earliest of them is the Zollner illusion, which dates back to 1860, and can be considered as one of the main figures which triggered off the study of optical illusions. Shown below is Zollner's illusion in its simplest form.

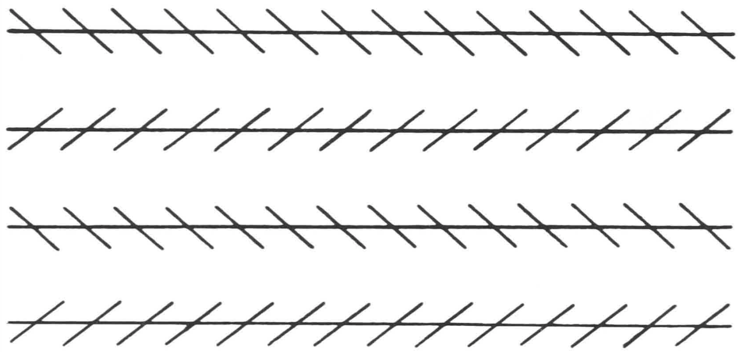


Fig. 64 – The Zollner illusion. The long horizontal lines are parallel to each other.

It has been suggested that assessing acute angles set in different directions produced the illusion. Zollner himself discovered that the illusion is the strongest when the hatching lines are inclined at 45° to the main lines. It has already been shown that convergence and divergence can affect our judgement of lengths, as shown in the Muller-Lyer illusions. Perhaps they account for the seemingly unequal distances between different portions of the main lines.

· Illusions of Curvature

One often attempts to assess the radius of curvature of an object by eye. This leads to huge degrees of error.

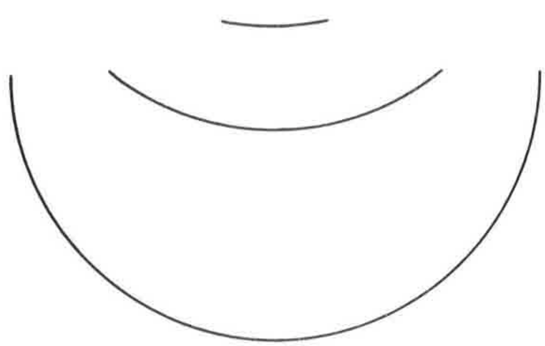


Fig. 65 – A curvature illusion. All three curves have the same radius of curvature.

In the above figure, all three arcs have the same radius of curvature, though it does not appear to be the case. They are all parts of the same circle. Yet it appears as if the smallest arc is much "flatter" than the rest, or that its radius of curvature is larger.

The estimation of the curvature of an arc is determined by the fraction of the whole circle visible to the observer. As the arc gets smaller, one tends to guess at a radius which is larger.

· Irradiation Illusions

The concept of irradiation illusions was first proposed by Helmholtz, more than a century ago. Irradiation of light from the light and dark regions on a painting tend to differ. It is suggested that bright light tends to "spill over" a little into the dark regions of an image on the retina. The effect is most marked when bright light is shone onto the image.

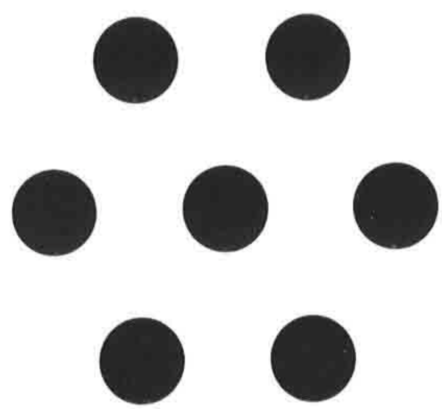


Fig. 66 – An irradiation illusion. The distance between the dark discs is equal to the disc diameter.

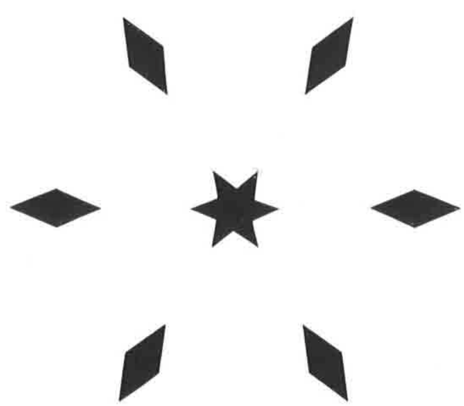


Fig. 67 – The star illusion. The distance from the star-tip to the lozenge-tip is equal to the length of the lozenge.

In Fig. 66, the dark discs are each surrounded by a large white area. The light area therefore encroaches into each of the dark discs, making it appear smaller than it actually is.

In Fig. 67, in addition to the irradiation effect are the mechanisms of convergence and divergence. Convergence makes each lozenge appear shorter than its true length. Divergence makes the space between the star-tip and lozenge-tip appear much bigger than it is. Irradiation "spills" the light over into the dark regions, "increasing" the separation. The overall effect is very convincing indeed.

· Crossed-Bar Illusions

The crossed-bar illusions were invented over a century ago by the well-known physicist Poggendorff. They are sometimes referred to as Poggendorff's illusions, and were the subject of much discussion in the late nineteenth century.

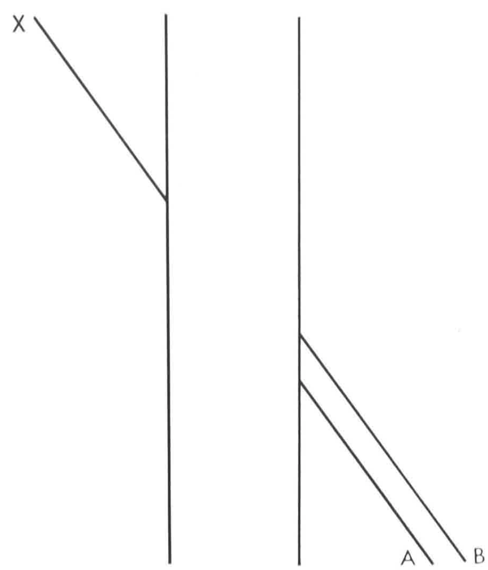


Fig. 68 – Poggendorff's illusion. The lines A and X are actually continuous.

The lines *B* and *X* in Fig. 68 appear to be continuous, and it appears that the line *A* is placed below the line *BX*. However, if a ruler is placed on *A*, it is discovered that *A* and *X* actually form a continuous line.

It is not well understood why the illusion arises, but we do know that the degree of illusion is strongly related to the degree of obliquity, as shown in Fig. 69.

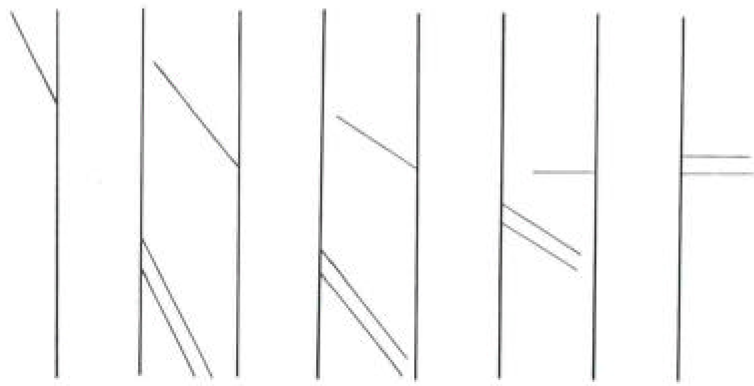


Fig. 69 – The degree of obliquity strongly affects Poggendorff's illusion.

Fig. 69 clearly shows that as the inclination of the lines is diminished, the illusion progressively weakens. It disappears when the lines are perpendicular to the bar. Hence, to produce a convincing crossed-bar illusion, a steep crossing angle is required. Furthermore, a steeper inclination will tolerate a greater separation distance between the two lines, as shown in Fig. 70.

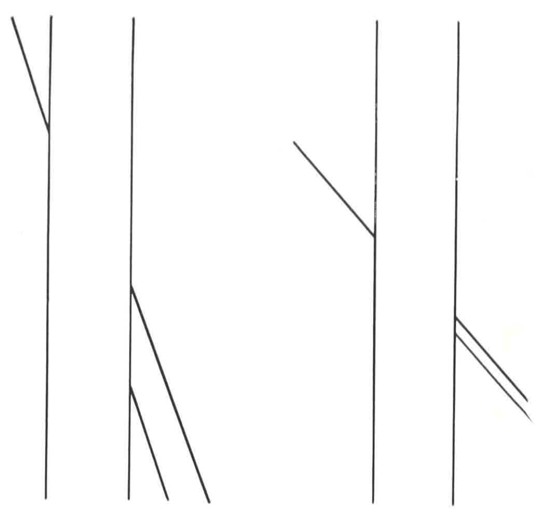


Fig. 70 – The steeper the inclination the greater the separation tolerated.

· Illusions Involving Oscillation of Attention

There are curious geometrical illusions that are not concerned with false assessments of sizes or angles. Instead, they arise from a difficulty in distinguishing what might be termed positive and negative interpretations of an image pattern.

Fig. 71 and 72 show photographs of the surface of a diamond, taken by physicist S. Tolansky. Using chemical processes to attack the diamond surface, an etched surface was produced. In Fig. 71, the microscopic blocks resemble mountainous chunks sticking out from the surface. However, in Fig. 72 the features resemble caverns or hollows in a solid wall, with no blocks standing out. Strangely, Fig. 72 is merely Fig. 71 turned upside down.

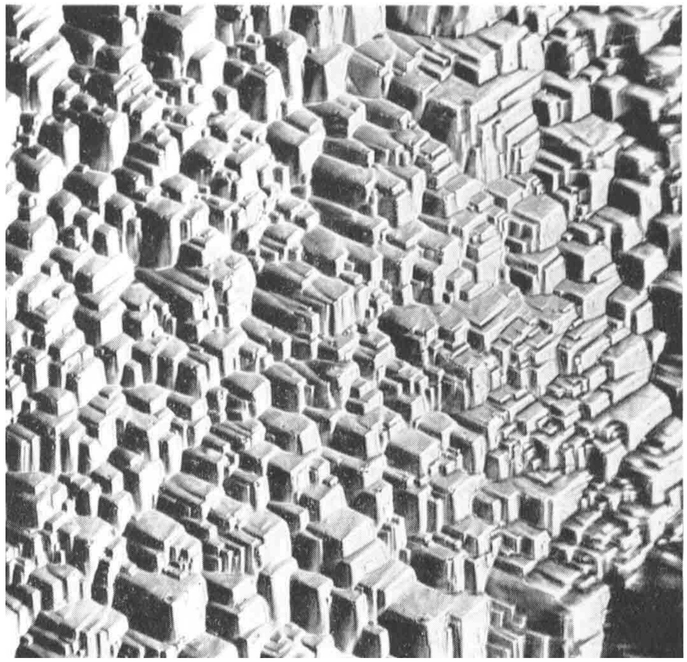


Fig. 71 – The etched surface of a diamond.

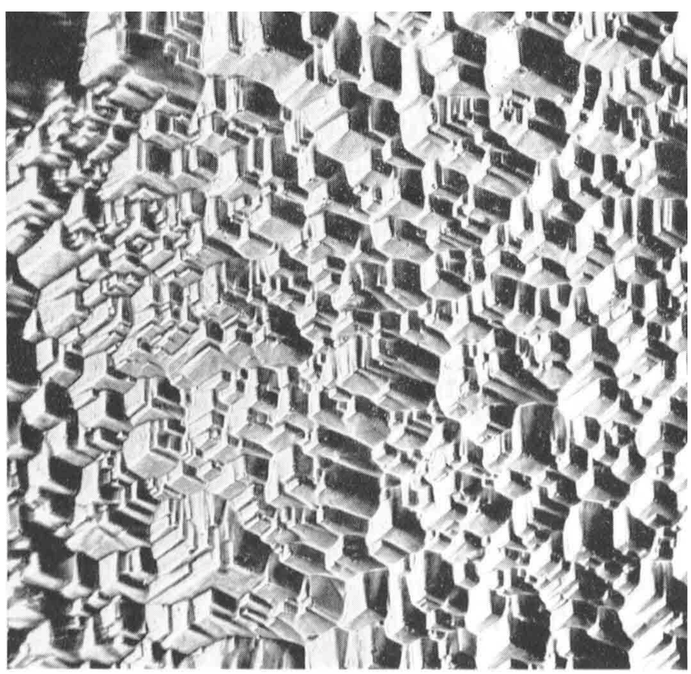


Fig. 72 – Another etched surface of a diamond.
This figure is actually Fig. 70 turned upside down.

The most interesting point is that there exists a turning point between the two interpretations, where the illusions morph into each other. If one turns Fig. 71 or Fig. 72 slowly, there will be some intermediate position where the picture suddenly oscillates. It jumps rapidly from blocks to caverns and back. There will be two positions where this takes place.

The illusion causes a serious scientific problem. Which surface are we actually looking at? Are we examining blocks or caverns? Tolansky reported that he had to use additional optical techniques to find out.

Shown below are a few classical examples of illusions due to the oscillation of attention. An artist particularly famous for such paintings is Maurits Cornelis Escher. As a child, he had an intensely creative side and an acute sense of wonder. Escher often claimed to see shapes that he could relate to in the clouds. Though people expressed the opinion that he possessed a mathematical brain, he never excelled in the subject at any stage during his schooling.

Fig. 73 – M. Escher's "Convex and Concave".

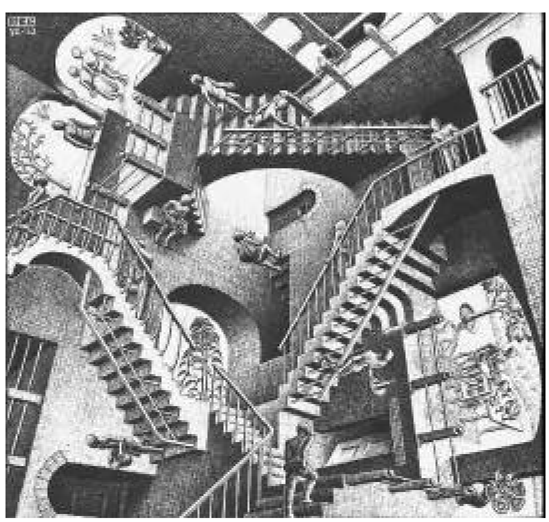


Fig. 74 – M. Escher's "Relativity".

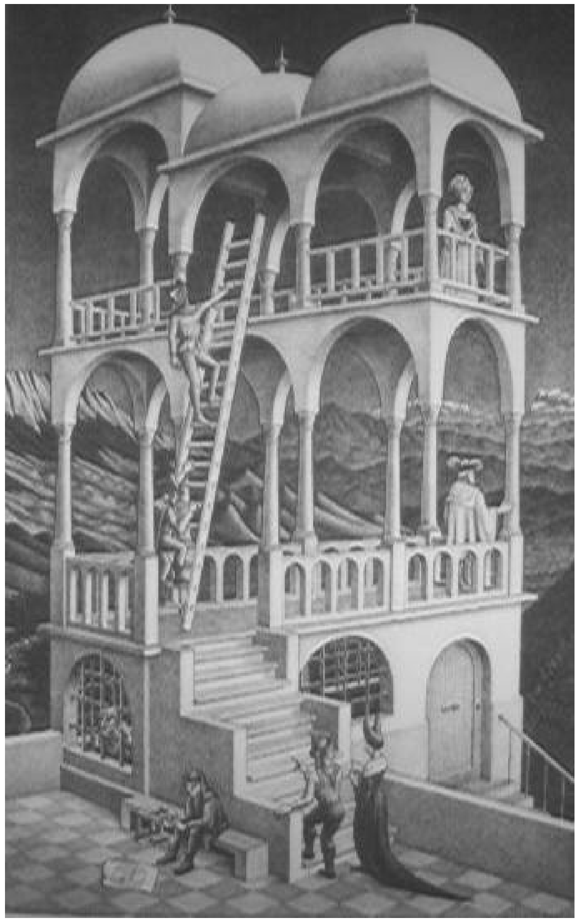


Fig. 75 – M. Escher's "Belvedere".

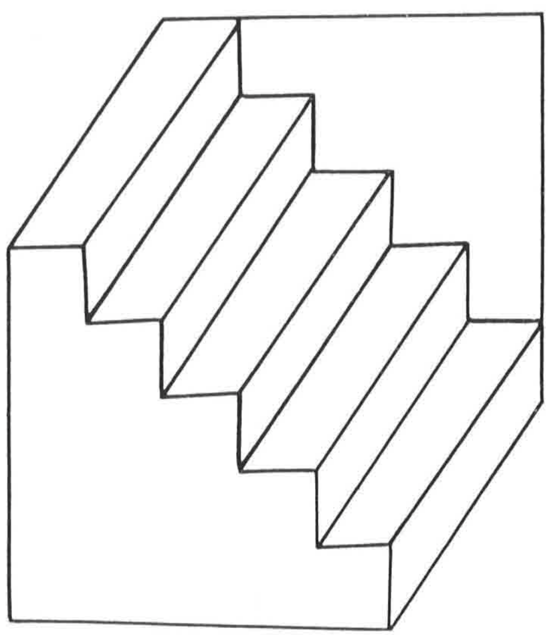


Fig. 76 – A classic example: Schroder's staircase.
Which is front and which is back?

VI. References

1. Judith V. Fields, ‘The Invention of Infinity’, Oxford University Press, 1997.

2. William M. Ivins Jr, ‘On the Rationalization of Sight’, Da Capo, 1975.

3. Dan Pedoe, ‘Geometry and the Visual Arts’, Dover Science Books, 1976.

4. M. H. Pirenne, ‘Optics, Painting and Photography’, Cambridge University Press, 1970.

5. S. Tolansky, ‘Optical Illusions’, Pergamon Press, 1964.

6. http://facstaff.uww.edu/mccreadd/