

Projective Geometry

Tom Davis

tomrdavis@earthlink.net

<http://www.geometer.org/mathcircles>

October 15, 2001

1 History: Perspective Drawing

You can learn a lot about how people view of the way the world works by examining their art. In paintings from the middle ages, the arrows of archers travel in perfectly straight lines until they reach their peak, at which point they stop, turn at a sharp angle, and drop straight to earth. This is pretty amazing, since every object that every human has thrown or shot since the beginning of time has traveled in a very smooth curve that approximates a parabola.

Similarly, it wasn't until the renaissance that drawings and paintings began to represent people further from the viewer smaller than those nearby¹. Again, it's surprising, since everyone since the dawn of man (and before, in fact) has seen people further away as smaller. Of course there are overwhelming psychological reasons for this incorrect representation—we “know” that even though the person is far away, he really remains the same size. Drawings in which things further from the viewer (be they people, buildings, or mountains) are smaller are called perspective drawings.

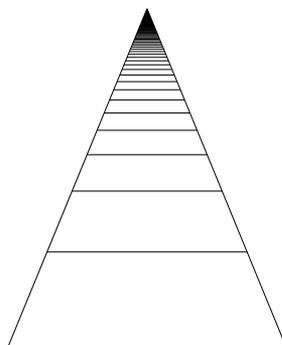


Figure 1: Railroad Tracks in Perspective

Now that we know how to draw in perspective, it is totally obvious that it is the “correct” way to draw. We know that if we look at a pair of railroad tracks on flat land going off to the horizon (see Figure 1) they will appear to meet in a point, and also that the ties under the tracks appear to get closer and closer together in the distance, even though we know that they are evenly spaced in the real world.

In fact, if you are good at the mechanics of painting, but have no sense at all of how to render a scene in perspective, there is a completely mechanical way to get a highly accurate rendition. Instead of a canvas, use a piece of glass, and keeping your head in exactly the same position², wherever you see green through the glass window, paint green at that spot on the glass. Paint red where you see red, et cetera, and it's clear that if you can match the colors exactly, you will have painted a scene on the glass that is in perfect perspective.

If you imagine the lines that light follows as it moves from the various objects to your eye through the glass, light rays from the top and bottom of an object will make an angle that basically determines the size of the object's image on the glass. If the same object is further away, the angle will be smaller, so the image on the glass will also be smaller. This

¹In fact, the painter Giotto who lived from approximately 1266 to 1337 was the first to realize that the relative sizes and shapes of things should be modified in paintings to make them appear more realistic. Of course he didn't know exactly how to do this, so some of his paintings appear a bit bizarre—there is some correction for perspective, but it's done incorrectly.

²Technically, you'll have to use just one eye, and keep that eye fixed in space.

is the basic idea behind perspective drawing.

It is also the basic idea behind projective geometry, which tells us how the drawings of objects on the glass are related to the positions of the objects in the real world, to the position of the glass, and to the position of the eye. The name “projective” comes from the fact that the scene that is being drawn is in reality being “projected” on the glass. You probably think of a projector in the opposite way, of course—a slide projector shines light from the lamp through the slide (glass) to a screen. But if you replace the lamp with your eye and imagine the light rays reversed and coming in from the objects, they would project an image of that object on the slide.

There is more to projective geometry, of course. Just to hint at a more difficult problem, imagine that you are a painter of a scene as above, but one of the subjects in your scene is another painter doing a perspective drawing on her canvas. When you draw on your canvas what she is drawing, how is your picture of her picture related to the real world, since it has undergone two projections? And if this seems too far out, consider this: the sun casts shadows on the ground, which are just projections of the objects on the “canvas” of the ground. If you’re a painter of a scene with shadows in it and you want to render the shadows correctly, you are really painting a projection of a projection.

Projective geometry is *not* just a subset of Euclidean geometry. It may seem similar since it seems to deal primarily with the projection of Euclidean objects on Euclidean planes. But that is not all it does. Think about our example of the pair of railroad tracks converging on the horizon. In your painting of the tracks, the two lines representing them meet at a point on your canvas, but what does that point represent in the real world? The answer is that it represents a point “infinitely far away” in the direction that the tracks are going (assuming, of course, that the world is really flat and extends forever). We can tell right away that something strange is going on, since Euclidean geometry is not equipped with any points that are “infinitely far away”, but this example shows that projective geometry has no problems at all representing such points (or at least their projections).

Today projective geometry is heavily used in a very practical way—every time you look at a three-dimensional drawing on your computer screen, all the calculations to produce that realistic image were calculated using the formulas of projective geometry.

1.1 Example: The Pinhole Camera

A pinhole camera provides another very nice illustration of perspective. A pinhole camera is just a light-tight box with film attached to one inside face and with a pinhole on the opposite face that is covered until you want to take a photo. To take a photo, point the pinhole in the correct direction, uncover it until the film is properly exposed, cover it again, and then remove and develop the film in a darkroom.

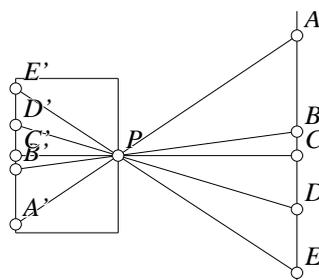


Figure 2: An Idealized Pinhole Camera

Here, of course, we’ll consider an idealized pinhole camera where the pinhole is an infinitely small point in a box with infinitely thin walls in a universe where light travels in perfectly straight lines. In the real world, the pinhole has to have some area, the camera walls some thickness, and when real light passes through a small opening, it is diffracted, or scattered slightly, depending on the size and shape of the hole.

In Figure 2, point P is the pinhole in the front of the camera, and film is attached to the side of the box opposite P . Imagine you are taking a photo of the line on the right with points A, B, \dots, E marked on it. Light scattered from each point goes off in every direction, but only rays of light aimed exactly at the pinhole will be able to reach the film. Thus the image of the point A is A' on the film, and so on. Notice that this reverses top and bottom, so if you kept track of which end of the film was up when you took the photo, things toward the bottom of the world would create

images toward the top of the film. Similarly, left and right are swapped on a two-dimensional piece of film.

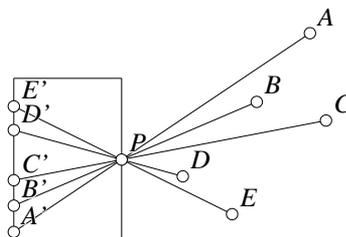


Figure 3: Pinhole Camera with a Complex Scene

In the real world, of course, the objects in the photo do not need to lie on a line parallel to the back of the camera. They can be anywhere at all in space, as shown in Figure 3. Also notice that we have drawn a one-dimensional slice of the camera, and the film is two-dimensional, and objects of interest can lie anywhere in the three-dimensional world.

It is instructive to think about restricted versions of the camera. What if the camera is taking a photo of a ruler with evenly-marked points? How will those points be spaced on the film? How does this change if the ruler is not parallel to the camera back? What if the camera back is not parallel to the front of the camera where the pinhole is?

Finally, notice that our earlier example of painting on a piece of glass is “just like” a pinhole camera where the film is in front of the pinhole (obviously a physically impossible camera, but it shows that the idea of a mathematical projection makes sense no matter where the “film” is).

1.2 Aside: Invariance Under Transformation

This is a little off the subject so it won’t hurt to skip this section, but it is interesting. As you will learn as you study more and more mathematics, there are many different kinds of geometry. Almost everyone learns Euclidean geometry first, but there is projective geometry that we’ll learn about here, as well as various other forms of non-Euclidean geometry: hyperbolic geometry, elliptic geometry, affine geometry, and so on.

One nice way to think about all the different geometries is to look at the sorts of transformations that are allowed and to see what properties of the geometric figures are preserved under those transformations. For example, if you allow objects to be rigidly moved on the plane, or possibly flipped over, you have the related idea of congruence in Euclidean geometry: two figures are congruent if one can be gotten from the other by sliding it around on the plane, perhaps rotating it in the plane, or even flipping it over.

Under these so-called “isometries”, things like lengths and angles are preserved.

In projective geometry, the main operation we’ll be interested in is projection. What sorts of things are preserved between a figure and its projection? The answer is that far less is preserved by this operation, but there are certain features that are. For example, lengths and angles are distorted, but if three points lie in a line before the projection, they will continue to do so after the projection. Similarly, if a bunch of points lie on *any* conic section (circle, parabola, ellipse, or hyperbola), their projection will also lie on a conic section, although not necessarily the same kind. Points on a circle may be projected to points on a hyperbola, for example. In fact, you’ve probably seen this happen at home—light from a bulb inside a circular lampshade makes a hyperbolic shadow on a flat wall.

What other properties are preserved under the allowed transformations? That is a central topic in projective geometry, and in fact, of any type of geometry.

2 A Taste of Projective Geometry

Let’s begin by looking at an ancient theorem of Pappus.

Theorem 1 (The Theorem of Pappus) *Let $ABCDEF$ be a hexagon with six distinct vertices such that points A , C , and E lie on one line and points B , D , and F lie on another. Let I be the intersection of AB with DE , let J be the*

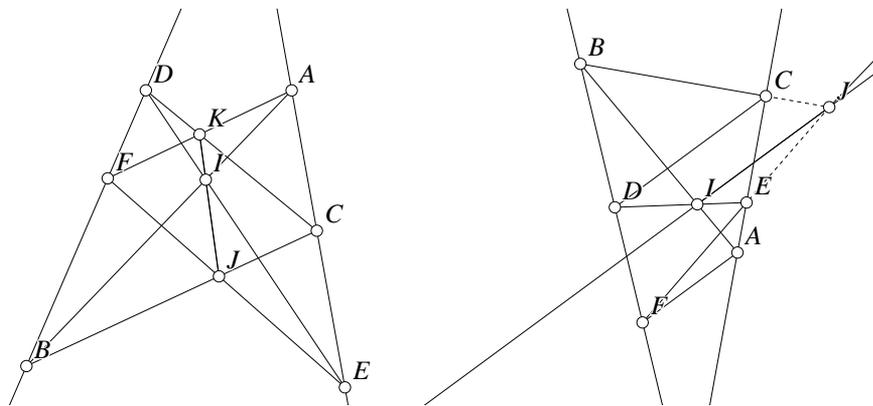


Figure 4: Theorem of Pappus

intersection of BC with EF , and let K be the intersection of CD with FA . Then I , J , and K lie in a straight line. See the left-hand drawing in Figure 4.

A couple of comments:

1. Obviously the hexagon we're talking about here is just a set of six points connected one to the other by lines. Since the vertices have to lie alternately on two lines, it's clear that our figure isn't going to look much like most of the hexagons you're used to which are usually convex, and the edges are going to cross each other. If you play with different examples of such hexagons, you'll also see that we are talking about the (infinitely long) lines that connect the six points, not just the segments that connect the lines. Sometimes the intersections mentioned in the theorems occur outside the segments, but the theorem still holds.
2. The comment above is just a technicality; more important is that *the theorem is simply not true*. At least it's not true in Euclidean geometry without some caveats. For example, there is no reason that the pairs of lines above whose intersections determine points I , J , and K intersect at all—they may be parallel. See the right-hand drawing in Figure 4 where the line AF is parallel to the line CD . In fact, it's a good exercise to draw the situation where all three sets of those lines are parallel, in pairs. But try to draw a situation where two of the pairs of lines are parallel and the third pair intersects. (Don't try too hard, though—it's impossible.)

We can fix up the theorem so that it is true in Euclidean geometry by modifying the second sentence to be this: "If none of the pairs of lines AB and DE , BC and EF , or CD and FA are parallel, then let I be the intersection. . . ." With this modification, the Theorem of Pappus becomes a perfectly good theorem in Euclidean geometry, but as we shall see, this modified version contains less information. We can add additional sentences to the statement of the theorem to recover this lost information, but the theorem will then need to be stated as a collection of cases.

If we are speaking about projective geometry, however, the Theorem of Pappus is *always* true, exactly as stated above, and there simply are no special cases to consider.

Let's look at where the theorem (as stated above) goes wrong in standard Euclidean geometry. Suppose one of the pairs of lines (let's say FA and CD to be definite—see the right-hand drawing in Figure 4) are parallel, and hence point K does not exist. If you draw a figure like this, you will usually find that the other two pairs of lines meet at points in the Euclidean plane so you can draw the line IJ , but you will also find that the line IJ is parallel to FA and CD . So although FA and CD do not meet, you might like to think of them as "meeting at infinity", and not any infinity, but at an infinity in a particular direction.

Or if you don't like this, think of the situation where FA and CD are almost parallel. They meet a long ways away from the hexagon, and as they get closer to being parallel, that point of intersection moves away faster and faster. But it continues to lie on the line IJ , so the line IJ itself will get closer and closer to being parallel to FA and CD as the intersection point moves. It's sort of reasonable (although not mathematically precise) to think that when the point does finally "get to infinity", the three lines will be parallel.

Note that the discussion of a “point at infinity” is simply nonsense in Euclidean geometry, but it does provide a useful way of visualizing what’s going on.

As we stated above, it is also possible for all three pairs of lines to be parallel, so all three of the “points” I , J , and K are “at infinity”. What happens if just two of the pairs of lines are parallel? The answer is that this simply cannot happen—if any two pairs are parallel, the third pair must also be parallel.

Here is a restatement of the theorem so that it is true in Euclidean geometry, and yet does not lose any information³:

Theorem 2 (Euclidean Theorem of Pappus) *Let $ABCDEF$ be distinct points forming a hexagon such that points A , C , and E lie on one line, and points B , D , and F lie on another. Let I be the intersection of AB with DE (if it exists), let J be the intersection of BC with EF (if it exists), and let K be the intersection of CD with FA (again, if it exists). It may be that some, but not all, of points I , J , and K exist. There are three possible cases:*

1. *If all three intersections exist, then they lie on a straight line.*
2. *If any two intersections do not exist, then the third one does not exist either.*
3. *If one of them (say I) does not exist, then the line JK is parallel to AB and to DE .*

Many theorems from projective geometry follow the same pattern—they are usually also true in Euclidean geometry, except when something “special” happens, such as lines being parallel so that certain points fail to exist. But in those cases, something special also happens to the Euclidean figure, and the theorem still seems to be true if we allow some sloppy thinking about “points at infinity”.

As a good exercise, look at the following famous theorem from projective geometry that is usually true in the Euclidean plane, but not always. Try to see when and why it fails, and then try to rewrite the theorem so that it is true in all Euclidean cases, and so that it tells what happens when something “special” occurs:

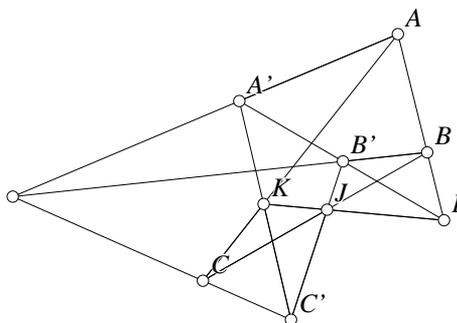


Figure 5: Desargues’ Theorem

Theorem 3 (Desargues’ Theorem) *Let $\triangle ABC$ and $\triangle A'B'C'$ be two triangles in the plane. Suppose that the lines AA' , BB' , and CC' all meet at a point. Let I , J , and K be the intersections of AB and $A'B'$, of BC and $B'C'$, and of CA and $C'A'$. Then I , J , and K all lie on the same line. See Figure 5.*

As usual, it is a very good idea to draw a few diagrams of your own and to play around with them.

3 Two Key Axioms of Projective Geometry

It is, of course, possible to make a complete list of the axioms of projective geometry, in exactly the same way as is done for Euclidean geometry. Since projective geometry is usually a full semester university course, there is no way

³In fact, it does lose a tiny bit of information, even as stated below. In projective geometry, there can be “points at infinity” and the theorem continues to hold in the projective sense even when one or two of the vertices of the hexagon lie at infinity.

to present a detailed description of all the axioms and how they all work together in a small article like this one. Many of the axioms are the same for projective and Euclidean geometry. It is, however, very useful to look at the axioms that highlight the major difference between the two geometries.

Here are two very important axioms from projective geometry. At first glance, they appear to be valid in Euclidean geometry as well:

1. Given any two distinct points in a plane, there is a unique line that lies on both of them.
2. Given any two distinct lines in a plane, there is a unique point that lies on both of them.

In Euclidean geometry, there's obviously a problem with the second axiom—what if the two lines happen to be parallel? Then there is no point that lies on both of them. The corresponding axiom for Euclidean geometry includes another phrase in the second axiom that says something like, “unless the two lines are parallel in which case there is never a point that lies on both of them.”

3.1 Models of the Axioms

If you've never seen projective geometry before, you're probably thinking, “How can axiom 2 above make sense? What about parallel lines? The author of this paper is crazy.”

The problem is that you already have a firm idea in your mind of what the words “point” and “line” mean, and that idea is almost certainly Euclidean. After all, it's probably the only way you've ever thought about those words.

But words like “point”, “line”, and “plane” in mathematics are “undefined terms”. The only things we can officially know about them are things we can prove about them from the axioms using logic. Obviously it's a good idea to have a mental model to help us think about them, but in an official mathematical proof, you can't use your mental model in the argument—you're restricted to using the axioms and pure logic.

Perhaps it would be easier if other words had been substituted. Let's look at the same axioms again:

1. Given any two distinct pratzels in a pongold, there is a unique lorper that lies on both of them.
2. Given any two distinct lorpers in a pongold, there is a unique pratzel that lies on both of them.

There! Now it's not so hard to believe, is it?

Actually, imagine that you're talking to a Martian who is extremely good at mathematics, but who only took freshman English for one semester at Diemos University. She⁴ never learned the terms for standard mathematical objects, so she just uses the Martian term when she doesn't know the English.

Here's a system she might be talking about. It consists of a pongold, a set of three pratzels ($\{a, b, c\}$) and a set of three lorpers ($\{\alpha, \beta, \gamma\}$). We will use the symbol \vdash to mean “lies on”. Here's the complete description of the system:

$$\begin{aligned}
 \text{Pongold} &: = \{a, b, c, \alpha, \beta, \gamma\} \\
 \text{Pratzels} &: = \{a, b, c\} \\
 \text{Lorpers} &: = \{\alpha, \beta, \gamma\} \\
 \text{Relations} &: a \vdash \alpha, b \vdash \alpha, b \vdash \beta \\
 & c \vdash \beta, c \vdash \gamma, a \vdash \gamma \\
 & \alpha \vdash a, \alpha \vdash b, \beta \vdash b \\
 & \beta \vdash c, \gamma \vdash c, \gamma \vdash a
 \end{aligned}$$

It's not too hard to check to see that both axioms hold (try drawing a—surprise—picture), although the system above doesn't look very geometric. The point is that the system above is a model for the axioms, so if those were the only two axioms, then whatever can be proved from them will be true of the system above (as well as of any other system that happens to satisfy the axioms).

⁴Actually, “She” is just the closest approximation—Martians come in three sexes. But the reason her English is pretty good is that with the longer Martian year, the semesters there are correspondingly longer.

Is that the only system that satisfies our Martian friend's axioms? No—there are, in fact, an infinite number of them! Here is a system that is far more interesting (and far more geometric). We have learned from previous conversations with her that her job on Mars is as a sort of large-scale surveyor, and she designs roads that travel for very long distances on the planet. The planet is basically a sphere, and the surface is very smooth, so all roads go in as direct a line as possible between the points they are meant to connect.

On a sphere, the shortest distance (on the surface) between two points is along a “great circle route”. A great circle is the intersection of the surface of the sphere with any plane passing through the center of the sphere. On Earth, the great circles include the equator and any of the lines of longitude (the north-south lines). Other than the equator, it does not include any latitude lines, since the planes passing through those circles do not pass through the center of the Earth. Of course these are just very special great circles; any plane through the center of the Earth (at any weird angle) makes a great circle as it cuts the surface, so there are an infinite number of great circles.

In fact, if you've ever been on an airplane for a long international flight, you've probably noticed that the path the plane takes does not appear to be a straight line on the map that you find in the seat pocket in front of you. That's because your route is probably close to a great circle route. For example, if you fly from San Francisco to Rome, you'll find that the plane goes high over Greenland, which is a heck of a lot further north than either San Francisco or Rome. If you drew the path on a globe, you'd find that a plane (a Euclidean plane, not an airplane) cutting through that path would go (approximately) through the center of the Earth, since it is approximately a great circle.

Anyway, what our Martian friend means is this: a pongold is just the (spherical) surface of Mars. A lorper is just any great circle on the surface of the planet. A pratzel is where two lorpers meet. But try to visualize where two great circles meet—they meet at two Euclidean points exactly opposite each other on the planet. In other words, a line connecting the two points would pass through the center of the planet. So a pratzel is an object that contains any two diametrically opposed points on the surface of Mars. It may not be obvious (but it is true—see below) that given any two different such pratzels (pairs of points), there is exactly one lorper (great circle) passing through both of them (through all four points of the two pratzels).

This is a perfectly good model of the two axioms. If you think about it, any two different great circles *must* intersect. If the great circles are different, they must correspond to two different planes passing through the center of the planet, and we know that two planes intersect in a Euclidean line. Both planes go through the center of the planet, so the center of the planet must be on the line of intersection of the two planes. Therefore, since the line of intersection of the two planes goes through the center of the sphere, it must hit the surface in two diametrically-opposed points (or, in the Martian's words, in a pratzel).

Just to complete the argument, let's show that any two pratzels determine a great circle (a lorper). If we draw a line connecting each pair, since the points are diametrically opposite, the lines must pass through the center of the planet. So the two lines intersect at the center. But any two distinct intersecting lines determine a plane, and where that plane cuts the surface of the sphere is the required lorper (great circle).

So she is making sense after all—and in a very geometric sort of sense, too.

4 A Model for Projective Geometry

Let's see if we can come up with a useful model of projective geometry that satisfies the two axioms above, and yet allows us to use some of our intuition about Euclidean geometry.

Here's what works for me (but your mileage may vary):

- Start with the standard Euclidean plane with its normal points and lines. Every point on the Euclidean plane will be a point in our projective geometry (but we will add some additional points later).
- Every Euclidean line on the plane is parallel to an infinite collection of other lines. All these parallel lines go in the same direction (or if you prefer, all these lines have the same slope). For every direction (or for every slope), add a “point at infinity” that corresponds to this direction. Add this one point to every one of the lines having that slope to make new lines in our projective geometry.
- For every direction, there is only one “point at infinity”. It does not matter that the line seems to go in two directions; imagine that the lines “wrap around”, meeting at the point at infinity.

- Finally, take all the points at infinity, and create one new line that consists of all (any only) those points. We'll call this the "line at infinity".

That's it.

Now, we need to check that our two axioms hold.

First of all, do every two points determine a line? Well, if the two points happen to be originally standard Euclidean points, then the standard Euclidean line (plus the one additional point at infinity) serves perfectly well as a line passing through both of them. It's also clear that this is the only line that will work.

If one of the two points happens to be a point at infinity, then choose as the line the one passing through the Euclidean point, and going in the direction of that particular point at infinity.

Finally, if both of the points are points at infinity, then we know that the line at infinity passes through both of them. No other line will work, since all the others are extended Euclidean lines, which can only have one slope, and two different points at infinity would represent two different slopes or directions.

Second, do any two lines determine a point? If the two lines are both standard Euclidean lines (with an additional point each), they are either parallel or not (in the Euclidean sense). If they are not parallel, they intersect in the standard Euclidean point (which is a point in our model of projective geometry). They do not share the point at infinity since if they intersect, they have different slopes, so the two projective lines meet at one and only one point. If they are parallel (in the Euclidean sense), then they both have in common the point at infinity that we added for that particular direction.

Finally, if one of the two lines happens to be the line at infinity which contains all the points at infinity, then in particular it will contain the point at infinity that we added for any other Euclidean line. Since every augmented Euclidean line is augmented with only a single point at infinity, the intersection of the line at infinity with any other line is restricted to this single point at infinity.

So our model satisfies the two projective axioms, and plane projective geometry is a valid system to consider.

It's a good idea now to go back to the Theorem of Pappus and Desargues' Theorem to see that under this model, the simple statements of both theorems are exactly theorems in projective geometry. Do this now.

5 Duality

Let's return to our Martian friend's pair of axioms. We've decided that a lorper means "line" (or at least our new funny kind of line), and pratzel means "point" (in the same funny sense).

Or does it? Maybe lorper means "point" and pratzel means "line". If you look carefully at the axioms, and replace every instance of "lorper" with "pratzel" and vice-versa, both axioms read exactly the same, although the first and second axioms are reversed!

So how can we tell which is which? The answer is, we cannot! There is absolutely nothing in projective geometry that makes a point behave differently from a line (from the point of view of the axioms). If you begin with any valid theorem about points and lines, and you go through and cross out every instance of "point" and write "line" and do the reverse process, replacing "line" with "point", you'll obtain another valid theorem.

So projective geometry is much cooler than Euclidean geometry, since every time you manage to prove a theorem, you have in fact proven two of them. This is called "duality".

We will look at one pair of dual theorems, but keep in mind that there is nothing special about these two—every theorem in projective geometry has a dual theorem that's obtained from the first one by swapping the words for "point" and "line".

5.1 Pascal's Theorem

The following theorem is to be interpreted in the projective sense—parallel lines meet at the appropriate point at infinity, et cetera. (In fact, the careful student will ask, "What in the heck is a circle in projective geometry? A circle certainly isn't preserved under projection—a circle can look like an ellipse, parabola, or hyperbola." Well, the fact is that Pascal's theorem in fact applies as long as all six vertices of the hexagon lie on any conic section, not just a circle.)

Also note that this (and its dual) is a projective theorem. It can be written in a Euclidean form, but care must be taken to describe what happens in the case of parallel lines, et cetera.

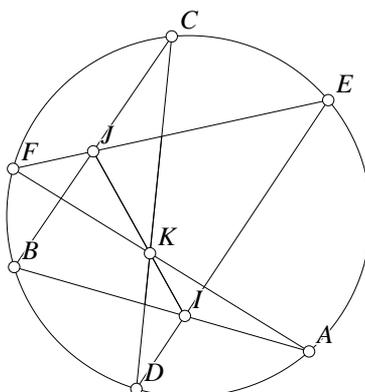


Figure 6: Pascal's Theorem

Theorem 4 (Pascal's Theorem) *Let $ABCDEF$ be any hexagon such that the distinct vertices $A, B, C, D, E,$ and F lie on a circle. Let AB and DE intersect at point I , let BC and EF intersect at J , and let CD and FA intersect at K . Then $I, J,$ and K all lie on the same line. See Figure 6.*

We won't prove this theorem, but it seems to be true from the figure, and if you draw other examples, you'll see that it seems to hold in those cases as well.

5.2 Brianchon's Theorem

So what does the dual of Pascal's Theorem look like? The dual is called "Brianchon's Theorem", and here's the direct "translation". To make it slightly easier to read, we use lower-case letters for lines, and if a and b are two lines, we'll use ab to indicate the point that lies at the intersection of the two.

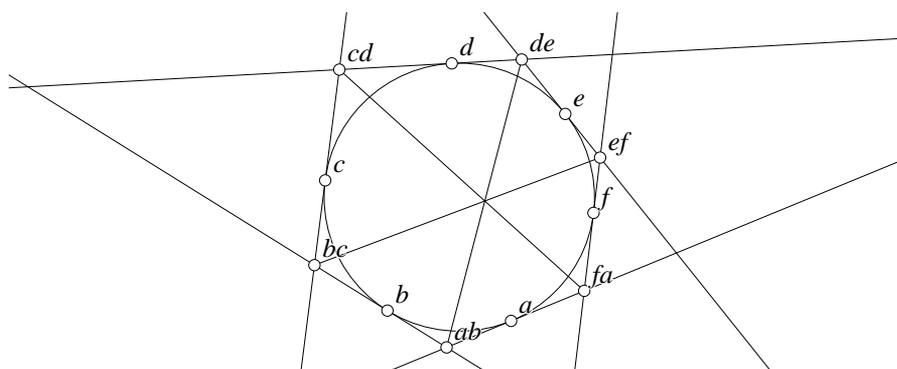


Figure 7: Brianchon's Theorem

Theorem 5 (Brianchon's Theorem) *Let $abcdef$ be any hexagon such that the distinct lines $a, b, c, d, e,$ and f are tangent to a circle. Let i be the line connecting points ab and de , let bc and ef lie on the line j , and let cd and fa lie on the line k . Then $i, j,$ and k all meet at a point. See Figure 7.*

Check that Brianchon's theorem is in fact the dual of Pascal's Theorem, and it's also a good idea to draw a few examples to convince yourself that it's true. Finally, check to see that it's true in the pure projective sense—when some of the lines intersect at points at infinity.

If \mathcal{A} and \mathcal{B} are two theorems, and if \mathcal{A} is the dual of \mathcal{B} , then \mathcal{B} is the dual of \mathcal{A} . This is probably obvious to you, but we'll state it here just to make sure you're aware of it.

Remember that there is *nothing* special about the fact that Pascal's (or Brianchon's) Theorem has a dual—every theorem in projective geometry has a dual. Try to find the duals of the Theorem of Pappus and of Desargues' Theorem. Draw some pictures to convince yourself that you've stated the dual theorems correctly.

6 Homogeneous Coordinates

The model of projective geometry in the previous section may be nice for visualization, but it is not too useful for making calculations, which is required if the goal is to find an exact mapping between the coordinates of objects before and after a projection.

Obviously we cannot just use the standard pairs of cartesian coordinates that we use in the Euclidean plane since every possible pair of real numbers exactly cover the plane, but there are no extra pairs left over to identify the newly added points at infinity. We can, however, make some progress by looking at those cartesian coordinates as points move "toward infinity" along straight lines in a given direction.

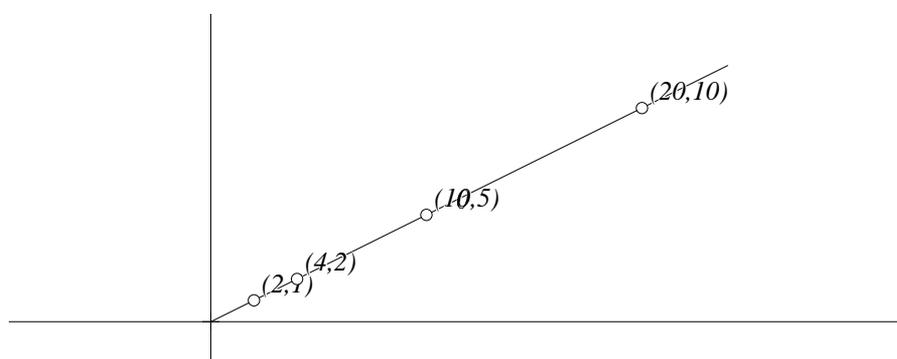


Figure 8: Cartesian Coordinates

As an example, consider the following points: $(2, 1)$, $(4, 2)$, $(6, 3)$, $(20, 10)$, $(200, 100)$, and so on, where the second coordinate is always double the first. If we plot these points on the standard plane (see Figure 8), as the coordinates get larger and larger, the points move toward infinity along the line $2y = x$. We'd like to have a point at infinity something like $(2\infty, \infty)$, but even if there were an ∞ in the real numbers, we'd need a 2∞ as well (and lots of other sizes of infinity to correspond to every other possible direction).

Here's another approach: All the points along the line above can be written $(2 \cdot \alpha, 1 \cdot \alpha)$, where α is a real number. As α gets larger, the points move off toward infinity. At first, this doesn't seem to make much progress since we won't get to the point we want until α gets to infinity, which it never will, quite.

But we can do something similar—consider points of the form $(2/c, 1/c)$, and this time, let c move toward zero. Again, we can't let c get to zero since division by zero is undefined, but at least c is headed toward zero—a real number that exists rather than ∞ —a number that does not.

6.1 Definition of Homogeneous Coordinates

Here is an idea that *does* work: use three coordinates. Every point on the projective plane is represented by a set of three numbers where at least one of the three numbers is non-zero. As long as c is not zero, we will think of the point (a, b, c) as representing the standard Euclidean point with coordinates $(a/c, b/c)$.

The first problem we encounter is that the same point can have more than one set of coordinates associated with it. For example, the three points $(2, 1, 1)$, $(4, 2, 2)$, and $(200, 100, 100)$ all correspond to the same Euclidean point $(2, 1)$. Is this going to cause serious problems?

It shouldn't—after all, we do exactly the same thing all the time with rational numbers. Nobody sees any problem with the fact that $1/2 = 2/4 = 100/200$, even though each of them has different “coordinates”. (We usually call the “coordinates” of rational numbers the numerator and denominator.) It is just a convention that we use to the “/” symbol to represent rational numbers. We could have written fractions as follows: $1/2 = (1, 2)$, $2/4 = (2, 4)$, et cetera.

In other words, if a/b is any normal fraction and c is any non-zero real number, a/b and $(ac)/(bc)$ represent exactly the same number. In the same way, if (x, y, w) are the homogeneous coordinates⁵ for any point in the projective plane, and α is any non-zero real number, then $(\alpha x, \alpha y, \alpha w)$ represents exactly the same point. In the same way that we normally reduce our rational numbers to lowest terms, we generally choose a simple representative for homogeneous coordinates when we can. If w is not equal to zero, we can multiply every coordinate by $1/w$ to obtain the equivalent point $(x/w, y/w, 1)$. Similarly, if you have a Euclidean point with coordinates (x, y) , you can find the corresponding homogeneous coordinates for that point by adding a 1 as the final coordinate: (x, y) (Euclidean) corresponds to $(x, y, 1)$ (homogeneous).

Now let's look again at what we were trying to do above. We noticed that the point $(x/w, y/w)$ moved out toward infinity along a line as w moved toward zero. But the Euclidean point $(x/w, y/w)$ corresponds to the projective point $(x/w, y/w, 1)$ which represents exactly the same point as (x, y, w) . In projective homogeneous coordinates, we can let w go all the way to zero, and we obtain a “point at infinity” with homogeneous coordinates $(x, y, 0)$.

In fact, any point with zero as the third coordinate will correspond to a point at infinity, and all those points together will lie on the “line at infinity”. The two pairs of homogeneous coordinates $(x_1, y_1, 0)$ and $(x_2, y_2, 0)$ on the line at infinity represent the same point only if there exists a non-zero number α such that $x_1 = \alpha x_2$ and $y_1 = \alpha y_2$.

Remember that at least one of the homogeneous coordinates *must* be non-zero: $(0, 0, 0)$ is *not* a valid set of homogeneous coordinates and it does not represent a point in our projective plane.

6.2 Calculations with Homogeneous Coordinates

OK, so we know how to relate our new homogeneous coordinates with points on the old Euclidean plane and with the new points at infinity. So what is the equation of a line? How can we tell if a point is on a line? How can we find the intersection of two lines? How can we tell if three points lie on the same line?

In high school algebra you probably learned that the equation of a line is given by the formula:

$$y = mx + b, \tag{1}$$

where m and b are real numbers. m is the slope, and b is the y -axis intercept. A point (x, y) lies on that line if and only if equation 1 holds.

The only serious drawback of equation 1 is that there are an infinite number of lines that it is unable to represent! Any line that is parallel to the y -axis has no representation because the slope is “undefined”. (And of course now that we're learning about projective geometry, whenever we see the word “undefined”, we are thinking to ourselves, “infinite”.)

Here is an alternate equation of a (Euclidean) line that works a lot better:

$$Ax + By + C = 0, \tag{2}$$

where A , B , and C are real numbers, and either A or B is non-zero.

This form of the equation for the line has a huge advantage in that it also allows vertical lines. If $B = 0$ and $A \neq 0$ then the equation becomes $x = -C/A$, and this is the equation of the line parallel to the y -axis having x -coordinate $-C/A$ everywhere.

It's not quite right for our homogeneous coordinates, however, since in general a projective point has three coordinates: (x, y, w) . A tiny modification fixes everything. Here is the equation for a completely general line in the projective plane:

$$Ax + By + Cw = 0, \tag{3}$$

⁵Notice that we tend to use w for the extra coordinate rather than z . This is because we don't want to get confused by the fact that (x, y, z) usually represents a point in normal three-dimensional Euclidean space. When we look at three-dimensional projective space (which will require four coordinates), we will write general points in that space as: (x, y, z, w) .

where A , B , and C are real numbers, and not all three of them are zero.

First of all, equation 3 works fine for the standard Euclidean points. The Euclidean (x, y) is normally represented in homogeneous coordinates as $(x, y, 1)$, and since the w coordinate is 1, equation 3 reduces to $Ax + By + C = 0$ which is just what we had in the Euclidean plane. If we happen to use a non-standard representation for a point (say we use $(\alpha x, \alpha y, \alpha)$, where $\alpha \neq 0$ for the Euclidean point (x, y) , it works just fine: $\alpha Ax + \alpha By + \alpha C = 0$ is the same as $Ax + By + C = 0$ since (being non-zero), we can divide through by α .

But equation 3 also works perfectly for points at infinity and can represent the line at infinity as well.

First let's check that the points at infinity work. If $Ax + By + C = 0$ is the equation for a standard Euclidean line, then the point at infinity that we have to add to it to make a projective line is this: $(-B, A, 0)$ (check that this is the case). If we plug $x = -B$, $y = A$, and $w = 0$ into equation 3, we obtain: $A \cdot (-B) + B \cdot A + C \cdot 0 = 0$ which is satisfied, so $(-B, A, 0)$ lies on the projective line, as it should.

Second, what is the equation for the line at infinity? The answer is: $0 \cdot x + 0 \cdot y + 1 \cdot w = 0$. Every point at infinity will have its w coordinate equal to zero, so it will look something like this: $(a, b, 0)$. Substitute this into the equation for our proposed line at infinity, and we obtain: $0 \cdot a + 0 \cdot b + 1 \cdot 0 = 0$. Again, it is satisfied.

To be completely sure our system makes sense, we should also check that the point at infinity for a line does not satisfy the equation for any other line with a different slope, and that points not at infinity do not lie on the line at infinity. These are very easy to do, but make sure you know how to do them.

6.3 Duality, Again

So a projective line is specified by exactly three real numbers, at least one of which is non-zero. We can thus write the equation of a line as: $[A, B, C]$. If a general projective point has coordinates (x, y, w) , where at least one of the three numbers is non-zero, then we know that the condition that must be satisfied for point (x, y, w) to lie on line $[A, B, C]$ is that $Ax + By + Cw = 0$. For now we're using the square brackets and upper-case letters to indicate coordinates of a line, and regular parentheses and lower-case letters to represent points.

Now, suppose I had gotten mixed up in the previous paragraph, and reversed the words "point" and "line". Would the paragraph still be true? Read it again and see that it is absolutely unchanged in meaning if you reverse the two words. If you are not told which objects are points and which are lines, there is no way you can figure out which is which from the axioms. This is another way to see that duality works.

Actually, the only problem with exchanging the words "point" and "line" is the appearance of the square brackets and parentheses. To do the exchange, you have to think of points as being written with square brackets and lines with parentheses. Because of the duality, it seems pointless to write points with parentheses and lines with square brackets. If we just write everything with parentheses, the duality is even more obvious.

And by the way, we could have just pulled the definition in the first paragraph in this subsection out of the air and said, "Here is a model for projective geometry." You can check it against all the axioms of projective geometry and see that these triples of numbers interpreted as points and lines satisfy all the axioms.

6.4 A few more relationships

If you know anything about three-dimensional vectors, it is quite interesting to learn that various vector and matrix operations on the homogeneous coordinates of a projective point have very interesting geometric interpretations.

If you don't know anything about three-dimensional vectors, here are the definitions of two of the more interesting operations, described only in terms of arithmetic operations on the coordinates. If $V_1 = (x_1, y_1, z_1)$ and $V_2 = (x_2, y_2, z_2)$, then the dot-product of V_1 and V_2 is defined to be:

$$V_1 \bullet V_2 = x_1x_2 + y_1y_2 + z_1z_2,$$

and the cross-product is defined as follows:

$$V_1 \times V_2 = (y_1z_2 - z_1y_2, z_1x_2 - z_2x_1, x_1y_2 - x_2y_1).$$

Note that the dot-product takes two three-dimensional vectors and returns a real number (usually called a "scalar", as opposed to a "vector" quantity). The cross-product always returns a vector quantity.

Now, to apply these definitions to projective geometry, let $P_1 = (x_1, y_1, w_1)$ and $P_2 = (x_2, y_2, w_2)$ be two projective points and let $L_1 = (A_1, B_1, C_1)$ and $L_2 = (A_2, B_2, C_2)$ be two projective lines.

We have already seen that if point P_1 lies on L_1 that is equivalent to $P_1 \bullet L_1 = 0$.

Here's an interesting exercise: Show that the equation of the line that passes through P_1 and P_2 is given by $P_1 \times P_2$. It isn't difficult—just list out the coordinates and make certain that P_1 and P_2 lie on the resulting line (and the previous paragraph shows you how to check that).

Here's another interesting exercise: Show that the coordinates of the point that lies on the intersection of two lines L_1 and L_2 is given by $L_1 \times L_2$. Don't work too hard—you already proved it if you solved the problem in the previous paragraph. Remember duality?

6.5 Matrix Transformations

Finally, and this is a bit beyond what is covered in the rest of this paper, here's a clean way to list all possible projective transformations. It requires that you know how to multiply a vector by a matrix. If you don't, here is the formula for multiplying a three-dimensional vector by a 3×3 matrix:

$$(x, y, w) \begin{pmatrix} A & B & C \\ D & E & F \\ G & H & I \end{pmatrix} = (xA + yD + wG, xB + yE + wH, xC + yF + wI). \quad (4)$$

If the 3×3 matrix above is non-singular (if it has an inverse, or equivalently, if its determinant is non-zero), then the operation above represents a valid projective transformation.

What is truly amazing is that such a huge collection of "standard" transformations are just special cases of projective transformations, including rotation, translation, scaling, shearing, and perspective projections. If you have a matrix that represents a translation and another that represents a rotation, the product of the two matrices represents the operation of rotation followed by translation. The details are beyond the scope of this paper; see any book on computer graphics for very practical details.

7 One Dimensional Projective Geometry

In this paper, we've only looked at the projective plane. One can consider projective spaces in any number of dimensions, and the equations will all be similar. Three-dimensional computer graphics uses three-dimensional projective space (with 4 homogeneous coordinates), and higher dimensional spaces are also reasonable to consider. Things do get complicated in three and more dimensions, and to learn about them, it is probably a good idea to look both in mathematics books and in books on computer graphics, since the computer graphics books take a very practical view of the subject.

But it is also interesting (and very easy to do) to analyze a one-dimensional projective space (with only 2 homogeneous coordinates). That's what we'll do in this section.

The nicest thing about one-dimensional projective geometry is that you can draw pictures on a piece of paper that completely describe the projections. For example, figure 9 illustrates a projection of a one-dimensional line onto another.

In this particular example, the line with evenly-spaced numbers on it on the left is projected on the numbered line on the right. The center of projection is the point O . You can see that the point 1 is projected to the point 0, that 2 is projected on 2, that 3 is projected to 5, and so on. Of course most of the other integer points are projected to points without integer coordinates.

If you think about it, this is exactly the one-dimensional analog of two-dimensional projections—of (2-dimensional) planes onto planes in three-space. In this case, (1-dimensional) lines are projected onto lines in two-space.

If you are given the equations of the two lines and the coordinates of the point of projection, it's clear that the projection is completely determined, and one of the things we'll look at in this section is the class of projections that are possible given arbitrary locations of the lines and of the point of projection. In fact, we'll even allow the point of projection to

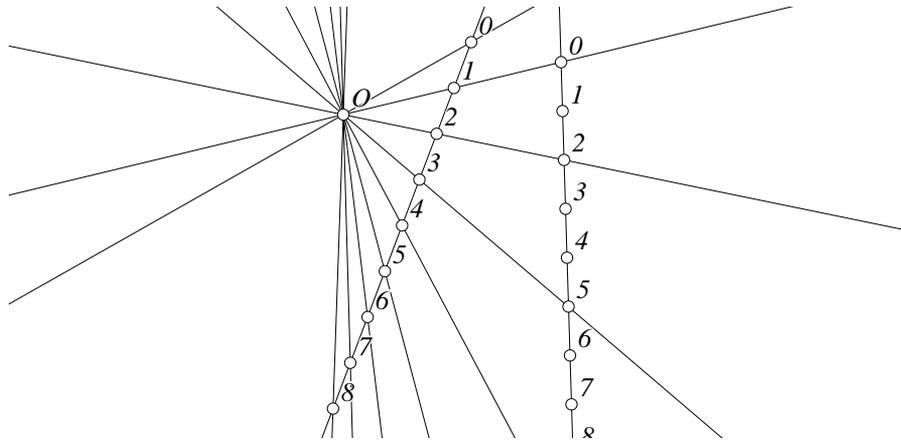


Figure 9: One Dimensional Projection

be a point at infinity, as we should⁶.

Just as we did in the two-dimensional case, we'll have to extend the Euclidean line to a projective line, but in this case, we only have one direction to go, so we only need to add a single "point at infinity" to turn a Euclidean line into a projective line. Looking back at figure 9, you can see that the line from O through the point 7 on the left line is approximately parallel (well, actually exactly parallel) to the right line, and so will never hit it. In fact, this projection maps 7 to the "point at infinity".

Using some techniques that we'll derive later, here is the formula for the projection above:

$$P(x) = \frac{10x - 10}{7 - x}. \quad (5)$$

It's easy to test that $P(1) = 0$, $P(2) = 2$, and that $P(3) = 5$. You can also see that if you try to evaluate $P(7)$ you get a denominator of zero which doesn't exist in the real numbers. But if you examine what happens to $P(x)$ as x gets close to 7, you can see that $P(x)$ gets very large, so it seems to "get close to infinity".

If you look back at equation 4 and see what it amounts to for standard Euclidean points with coordinates (x, y) (and therefore homogeneous coordinates $(x, y, 1)$), we obtain:

$$\begin{aligned} x' &= \frac{Ax + Dy + G}{Cx + Fy + I} \\ y' &= \frac{Bx + Ey + H}{Cx + Fy + I} \end{aligned}$$

If we toss out the y coordinate from the equations above, we obtain something very similar to equation 5. In fact, the most general possible projective transformation of a line into a line is given by the formula:

$$P(x) = \frac{Ax + B}{Cx + D},$$

as long as $AD - BC \neq 0$. This final condition that $AD - BC \neq 0$ is necessary; otherwise numerator and denominator are just multiples of each other and the projection will take every point on the line to a constant point.

⁶If we don't allow projections from the point at infinity, we can't even have an identity projection. With the point at infinity, it's easy—make the lines to be projected to each other both parallel to the y -axis and the point at infinity to be in the direction of the x -axis. Then all the projection lines will be parallel to the x -axis, and we'll obtain the desired result.