What is a straight line?

Dan Chazan   Bill McCallum

August 5, 2008

1 Straight lines in Euclidean geometry

Although Euclid purports to define straight line, his definition is enigmatic: "a straight line is a line that lies evenly with the points on itself." Straight lines are treated in Euclidean geometry as primitive, undefined objects satisfying certain axioms. However, the theorems resulting from these axioms give us many ways of characterizing straight lines. For example, we know that if a straight line cuts two parallel lines transversally, then corresponding angles are equal (Book I, Proposition 29). In this essay, when we talk about straight lines, we always mean straight lines in Euclid's sense. However, we want to related it to two other ways of describing straight lines.

2 Straight lines in analytic geometry

In modern analytic geometry we coordinatize straight lines and represent them on a Cartesian plane by equations in two variables. In fact, many school children think that this is the definition of straight line: it is the set of solutions (x, y) to the equation

$$Ax + By = C, \tag{1}$$

where A, B, and C are constants (with at least one of A and B being non-zero). In this section we aim to show two things:

   I. all solutions to (1) lie on a straight line

  II. all straight lines can be described as the set of solutions to an equation of the form (1).

Suppose we have a solution (x_0, y_0) to the equation (1). Additional solutions can be obtained from (x_0, y_0) by moving uB in the x direction and $-uA$ in the y direction. This new point, say (x_1, y_1), will have coordinates $(x + uB, y - uA)$. Thus

$$Ax_1 + By_1 = A(x + uB) + B(y - uA) = Ax + uAB + By - uAB = Ax + By,$$

so if $Ax + By = C$, then $Ax_1 + By_1 = C$. So if the point (x_0, y_0) is a solution to the equation, then so is $(x_1, y_1) = (x_0 + uB, y_0 - uA)$.
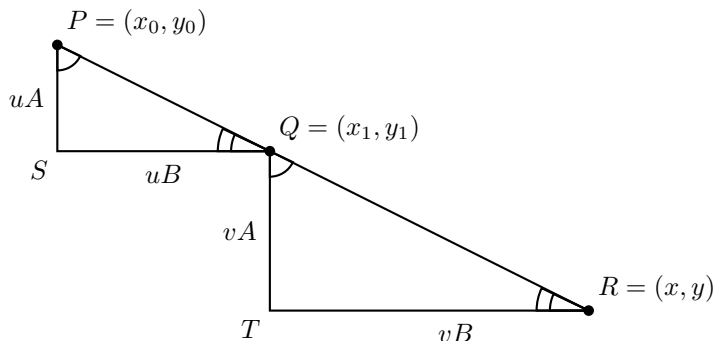
Figure 1: Two triangles formed from solutions to $Ax + By = C$

This can be represented graphically by drawing a triangle starting at P, moving $-uA$ in the y direction and uB in the x direction, for some non-zero real number u. In a similar fashion, a second triangle can be formed from this new point (which we already have shown to be a solution to the equation) by moving an additional $-vA$ in the y direction and vB in the x direction. In fact, every solution to the equation (1) can be constructed in this way. See Figure 1.

We now show that R is on the straight line through P and Q, in the sense of Euclidean geometry. Since R represents an arbitrary solution to the equation (1), this will demonstrate fact I above, that all solutions lie on the same straight line.

Since the x-axis is perpendicular to the y-axis $\angle PSQ$ and $\angle QTR$ are right angles. Furthermore, if A, B, and v are non-zero, then the ratios

$$\frac{|SQ|}{|TR|} = \frac{|uB|}{|vB|} = \left|\frac{u}{v}\right| \quad \text{and} \quad \frac{|PS|}{|QT|} = \frac{|uA|}{|vA|} = \left|\frac{u}{v}\right|$$

are the same. Since the two triangles have two pairs of corresponding sides that are proportional and the included angles are congruent, $\triangle PSQ$ is similar to $\triangle QTR$, by the SAS similarity theorem.[1] Therefore, $\angle SPQ \simeq \angle TQR = \theta$. On the other hand, the straight line through P and Q forms the same angle θ with each of the parallel lines PS and QT, so R is indeed on that straight line, as required.

The same argument works in reverse to prove fact II above. Indeed, suppose ℓ is the straight line through P and Q, and let R be an arbitrary point on ℓ. Construct right angled triangles with legs parallel to the axes and vertices at P, Q, and R, as in Figure 1. Since PS and QT are parallel to the vertical axis, the corresponding angles $\angle SPQ$ and $\angle TQR$ are congruent, and since SQ and TR are parallel to the horizontal axis, the corresponding angles $\angle PQS$ and $\angle QRT$ are congruent. Thus the triangles $\triangle SPQ$ and $\triangle TQR$ are similar, by the AA

---

[1] If A or B is zero, we have similar degenerate triangles that are either horizontal or vertical line segments. If $v = 0$, then $R = Q$ and there is nothing to prove.

similarity theorem. Hence the lengths of corresponding sides are proportional:

$$\frac{x - x_1}{x_1 - x_0} = \frac{y - y_1}{y_1 - y_0}.$$

This equation can be rewritten in the form (1), as required:

$$\left(\frac{1}{x_1 - x_0}\right) x + \left(\frac{-1}{y_1 - y_0}\right) y = \frac{x_1}{x_1 - x_0} - \frac{y_1}{y_1 - y_0}.$$

3 Relation with the physical notion of straightness

We have shown that the straight lines of Euclid's geometry are the same as the lines defined by linear equations in analytic geometry. However, it is unsatisfying to leave open the question of how the mathematical notion of a straight line relates to the physical experience of straightness.

How do we decide if a piece of wood or a wall is straight? One way is to line it up along our line of sight and see if we can detect any wobbles. This method depends on our belief that light travels in straight lines. We can make this intuition more mathematical by remembering Fermat's principle: the path taken between two points by a ray of light is the path that can be traversed in the shortest time. Assuming the medium through which the light is travelling is homogeneous, so that light travels at the same speed at every point, Fermat's principle of shortest time is equivalent to saying that light takes the shortest distance between two points (this is not true when the medium changes, for example, from air to water, and in fact if you line your sight along a straight piece of wood as it enters water you will see that it does not in fact look straight, but appears to bend).

Using these considerations as a guide, we can formulate a notion of straight line that corresponds to physical experience: a line in $\mathbb{R}^2$ is straight if it takes the shortest path between any two points on it.

The Cartesian measuring stick

In order to even talk about the shortest path, we need some way of measuring distance between two points—we need a measuring stick. In the Cartesian plane our measuring stick is the distance formula: if $P = (a, b)$ and $Q = (x, y)$, then the distance between P and Q is

$$d(P, Q) = \sqrt{(x - a)^2 + (y - b)^2}.$$

This formula is really Pythagoras's theorem in disguise [insert figure and explanation of this]. By the way, if you were wondering in the previous section about the possibility of light following curved paths, as it does in general relativity, you can stop worrying. By adopting the Cartesian measuring stick we have put ourselves very firmly into flat space.

How do we tell if a path is following the shortest distance?

Now that we have a way of measuring distance, we want a way of deciding whether a path is the shortest distance between two points. Although we know how to measure the distance between two points, that's not quite the same thing as being able to measure the length of a path. Imagine, for example, trying to measure the length of a curving brick pathway in your garden. You wouldn't be able to use a long, stiff, measuring stick to do this. You would have to make a lot of small measurements with a short ruler, and add them up. Since even along small distances the path might curve, this would yield only an approximation to the length. The problem of defining the length exactly is quite difficult, and requires the ideas from calculus. Fortunately, there is a way of telling whether a path is straight without actually measuring its length. The idea is that we can detect non-straightness quite easily. Suppose we have three points in order along the path, P, Q, and R. If P, Q, and R are not lined up, the distance from P to R is less than the sum of the distances from P to Q and from Q to R. In other words, the shortest distance from P to R deviates from the path. So our physical definition of straightness can be phrased as follows:

> a line ℓ is straight if for any three points P, Q, and R on ℓ, with Q between P and R, we have

$$d(P, R) = d(P, Q) + d(Q, R). \tag{2}$$

We now show that points satisfying this condition satisfy an equation of the form (1). Suppose that $P = (a, b)$, $Q = (x, y)$, and $R = (c, d)$. Our geometric condition is

$$d(P, R) = d(P, Q) + d(Q, R), \tag{3}$$

or

$$\sqrt{(c-a)^2 + (d-b)^2} = \sqrt{(x-a)^2 + (y-b)^2} + \sqrt{(c-x)^2 + (d-y)^2}.$$

We expect to be able to derive the equation of a straight line from this. A brute force approach to this might go as follows. First, square both sides of the equation and get

$$(c-a)^2 + (d-b)^2 = (x-a)^2 + (y-b)^2 +$$
$$2\sqrt{((x-a)^2 + (y-b)^2)((c-x)^2 + (d-y)^2)} + (c-x)^2 + (d-y)^2.$$

Then, isolate the square root term on the right, square both sides again, and get

$$((c-a)^2 + (d-b)^2 - (x-a)^2 - (y-b)^2 - (c-x)^2 - (d-y)^2)^2 =$$
$$4((x-a)^2 + (y-b)^2)((c-x)^2 + (d-y)^2),$$

which, the intrepid reader will verify, is equivalent to

$$-4((ad - bc) - (d-b)x + (c-a)y)^2 = 0,$$

or, finally, the linear equation

$$(d - b)x - (c - a)y = (ad - bc), \tag{4}$$

which is in the required form (1).

Note that equation (4) is not necessarily equivalent to equation (3), since we squared some things along the way. All we can say is that the solutions to (3) are contained in the solutions to (4). In fact, (3) describes the segment between P and Q. The other two rays of the line, starting from P and heading away from Q, and starting from Q and heading away from Q, would be described by equations similar to (3) but with different signs. The reader can verify that all these equations lead to the same equation (4).

Back to Euclid

How do we link Euclid's straight lines to the condition (3)? One of the theorems in Euclid relates to this condition. Book I, Proposition 20, states that in any triangle, the sum of any two sides is greater than the remaining one. So, if Q is not on the line through P and R, then (3) is not satisfied, or, in other words, (3) implies that Q is on the line through P and R. The converse is taken for granted in Euclid.