Constructible Points and Constructible Numbers, Part 2

Abstract

Last time, we introduced the basic constructions with straightedge and compass. We began by defining *constructible points* and *constructible numbers*. The former are points in the plane we can construct with straightedge and compass, beginning with two given points specified as $(0,0)$ and $(1,0)$. The latter are, up to sign, the distances between constructible points. We have already established that a large set of points and numbers are constructible.

This leaves an open question: are there any points in the plane that are *not* constructible? Though a complete characterization with proof of all constructible points and numbers is beyond the scope[1] of what we can prove in just this session, we shall give some examples of nonconstructible points and distances. Using this as background, we consider three classical construction problems: (1) "squaring the circle", (2) "doubling the cube", and (3) trisecting a general angle.

Background needed: Our previous session on this topic will be valuable, especially if these topics are unfamiliar, but it may not be absolutely necessary. Beyond this, we require enough background in algebra and geometry relevant to considering equations for lines or circles in the planes, as well as for computing the coordinates for points of intersection between lines, lines and circles, or two circles.

0 Review

We begin by restating the definitions of constructible point and constructible number:

Definition 0.1. In the plane, we are given two distinct, specified points P_0, P_1, and we choose coordinates such that $P_0 := (0,0)$, $P_1 := (1,0)$. (In particular, the length P_0P_1 of the segment $\overline{P_0P_1}$ is 1.) By definition, P_0 and P_1 are constructible points.

Assume that P, Q are already-constructed points. Then using our straightedge and compass, we may

(a) use the straightedge to draw the line $\overleftrightarrow{PQ}$.[2]

(b) use the compass to draw a circle $\odot P$ with center P and radius $\overline{PQ}$; and

(c) for any already-constructed line or circle that intersects another already-constructed line or circle, we may specify any point of intersection R.

[1] A comprehensive solution to this problem requires techniques from abstract algebra, especially from *Galois Theory*.

[2] For our purposes, we may imagine that our straightedge is infinitely long, so that we may construct the *line* $\overleftrightarrow{PQ}$ rather than simply the *line segment* $\overline{PQ}$;

We say that any point R distinguished as in Definition 0.1(c) is a *constructible point*. Further, we define a real number x to be a *constructible number* if and only if we can construct two points P and Q such that that $|x| = PQ$. That is, constructible numbers are precisely those arising as distances between constructible points, up to sign.

That is, R is a constructible point if and only if we locate R after finitely many constructions of the form given in Definition 0.1(a)–0.1(c), beginning with our given points P_0 and P_1.

In our first session on this topic, we established the following:

Theorem 0.2. *Each of the following constructions is possible:*

(a) Given distinct points P and Q, one can construct a point R such that $\triangle PQR$ is equilateral. That is, we can construct a point R such that $PQ = QR = RP$.

As a corollary, we can construct a 60° angle.

(b) Given distinct points P and Q, one can construct the perpendicular bisect of $\overline{PQ}$.

(c) Given distinct points P, Q, R, where R does not lie on the line $\overleftrightarrow{PQ}$, we can construct a line ℓ through R such that ℓ is parallel to $\overleftrightarrow{PQ}$.

(d) Given distinct, noncollinear points P, Q, R, one can construct a line ℓ through Q that is the angle bisector of $\angle PQR$.

(e) Let a, b be constructible numbers, where $b \neq 0$. Then $-a$, $a+b$, ab are also each constructible numbers. If $b \neq 0$, then a/b is also constructible. Finally, $\sqrt{|a|}$ is also constructible when a is constructible.

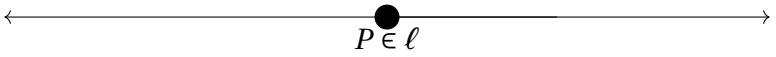
As a corollary to Theorem 0.2(e), it follows that every point of the form (r, s), where r, s are both rational numbers or square roots of rational numbers, is a constructible point.

1 Additional Constructions: Straightedge and Compass

We continue exploring different constructions:
Exercises: Many of the following are taken from the online game *Euclidea* (euclidea.xyz).

1.1 Let P be a point and ℓ a line through P. Construct a line $\ell_\perp$ through P that is perpendicular to ℓ.

Note: This appears as Exercise #2.7 (Beta pack) in Euclidea.



1.2 Let $ABCD$ be a rectangle. Without loss of generality, assume $BC \leq AB$. Construct points P in $\overline{AB}$ and Q in $\overline{CD}$ such that $PBQD$ is a rhombus. (A rhombus is a convex quadrilateral all of whose sides are congruent.)

This appears as Exercise #1.5 (Alpha pack) in Euclidea.

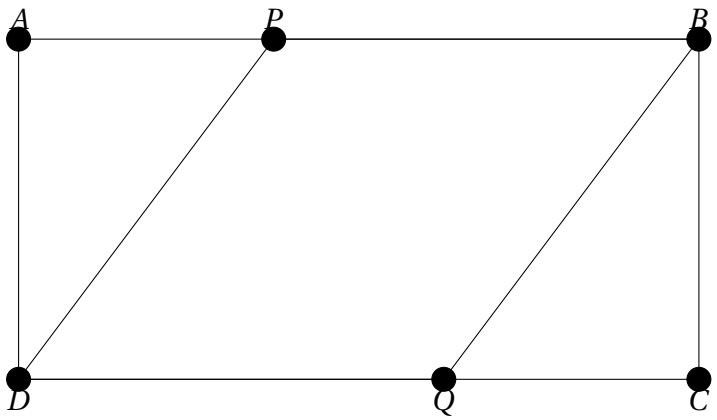
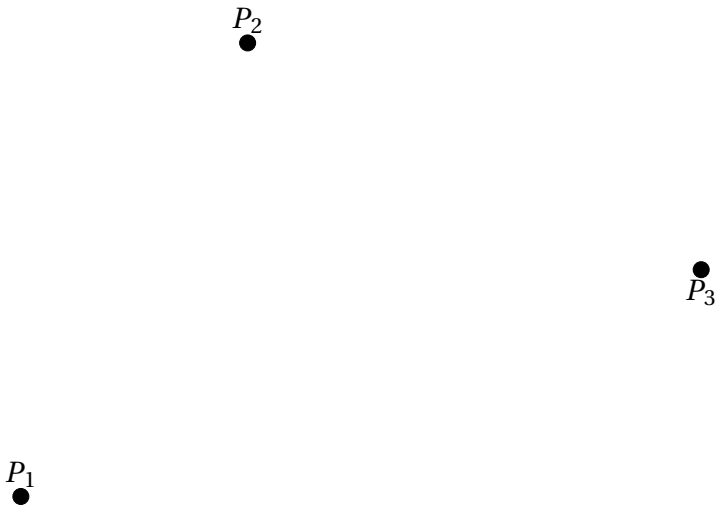


Figure 1.1: A Voronoi diagram (2048px-Euclidean_Voronoi_diagram.svg.png, *Wikipedia*) relative to the indicated black points

1.3 Assume we have finitely many distinct points $P_1, P_2, \ldots P_n$ in the plane. The *Voronoi diagram* for $\{P_1, P_2, \ldots, P_n\}$ is a decomposition of the plane into the regions of the plane closest to each P_j; see Figure #1.1. (Points equidistant from the two nearest neighbors in $\{P_1, \ldots, P_n\}$ will be on the boundary of two or more of these differently colored regions.)

Given a set of finitely many points in the plane, explain how to construct a Voronoi diagram for them. Can you produce the Voronoi diagram for the following?



1.4 From Theorem 0.2(d) (or based on your own past experience), we know that for any angle, we can in principle construct its angle bisector.

Show that we can trisect a 180° angle. Show also that we can trisect a 90° angle.

2 The Algebra of Constructible Points and Numbers

In the abstract, we mentioned that there are some points in the plane that are *not* constructible, from which it follows that there are nonconstructible numbers, too. In order to establish this, we first need to develop a bit of algebra relevant to constructible points and numbers. Recall that, by definition, a constructible point P is either one of our initially given points $(0,0)$ or $(1,0)$, or P is constructed by arising as the point of intersection of two constructible lines, two constructible circles, or a constructible line and a constructible circle.

Many of these arguments will have to be incomplete, as a fully rigorous proof might take an entire semester to present! My hope is that despite these gaps, the introduction of these ideas will still be worthwhile.

Exercises:

2.1 Let $P := (a,b)$, $Q := (c,d)$ be distinct points in the plane. Give an equation that represents the unique line ℓ through P and Q.

Note: The specific form of your line—e.g., slope-intercept form, etc—isn't important for our purposes. (One caveat: if $a = c$, then our line must be vertical, in which case it has undefined slope.)

2.2 Let $P := (a,b)$, $Q := (c,d)$, with P,Q distinct. Provide an equation of the circle C with center P and passing through Q.

2.3 Recall that for a line to be constructible, it must arise via two already-constructed points. Assume, then, we have distinct constructible points $P := (a, b), Q := (c, d)$, and another pair of distinct constructible points $P' := (a', b'), Q' := (c', d')$. Based on our results in Exercise #2.1, show that *if* the lines $\overleftrightarrow{PQ}$ and $\overleftrightarrow{P'Q'}$ intersect in precisely one point,[3] then the point of intersection (u, v) is such that u, v are both quotients of some expressions[4] in $a, b, c, d, a', b', c', d'$. (A plausibility argument here would suffice; the actual algebra here gets *very* messy, and precise solutions are unimportant for our future purposes.)

2.4 Let P, Q be distinct constructible points, and let $\ell := \overleftrightarrow{PQ}$. Further, let P', Q' be two more constructible points, distinct from each other (but not necessarily distinct from either of P and Q), and let C denote the unique circle with center P' that passes through Q'.

If (u, v) simultaneously lies in C and $\overleftrightarrow{PQ}$, show that u and v are quotients of "expressions" in the coordinates $a, b, c, d, a', b', c', d'$ (in the sense of the foonote to Exercise #2.3) and square roots of such expressions.

In other words, when constructing a new point as the intersection of a line and a circle, each arising from pairs of distinct constructible points, we can add either quotients of rational polynomials in the coordinates, or square roots of those quantities.

2.5 Let P, Q be distinct constructible points, P', Q' two more distinct constructible numbers (but not necessarily with either distinct from either of P or Q), and let C, C' denote the circles with respective centers P and P', and passing through Q and Q', respectively.

Assume that (u, v) lies in the intersection of C and C'. Following our previous exercises, what can we say about the coordinates u, v in terms of the coordinates for

[3] When would these lines intersect in more than one point? When would they fail to intersect at all?

[4] More precisely—though perhaps less usefully for those unfamiliar with this terminology—these "expressions" are rational (multivariable) polynomials in $a, b, c, d, a', b', c', d'$. In the language of abstract algebra, this means u and v lie in the *field of fractions* of the coordinates of P, Q, P', Q'.

P, Q, P', Q'?

These exercises help motivate the following:

Theorem 2.1. *If x is a constructible number, then (1) x is the root of some nonzero polynomial with integer coefficients (meaning x is an* algebraic number*), and (2) if $p(x)$ is such a polynomial that has smallest degree n, then n must be a nonnegative integer power of* 2*.*

To illustrate this polynomial property, $\sqrt[3]{2}$ is an algebraic number because it is a root of the nonzero polynomial $x^3 - 2$ which has integer coefficients.

As a corollary to the theorem, it would follow that if a real number x is either *not* an algebraic number, or if x is algebraic but has "algebraic degree" that is *not* a power of 2, then x is not a constructible number. In particular, we expect that $\sqrt[3]{2}$ has algebraic degree 3, and 3 is not a perfect power of 2.

3 Classical Impossibility Theorems

Exercises:

3.1 Let α be a real transcendental number. That is, α is a real number, but α is not the root of any nonzero polynomial $a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$, where all the a_j are *integers*.

In particular, that $\pi \approx 3.14159$ is a real transcendental numbers. *Assuming* this result about π being real transcendental, prove that $\sqrt{\pi}$ is also real transcendental.

Hint: Consider proof by contradiction, and assume that $p(x)$ is a nonzero polynomial with integer coefficients such that $p(\sqrt{\pi}) = 0$. What can we say about the polynomial $q(x) := p(x) \cdot p(-x)$?

3.2 One of the problems from antiquity is "squaring the circle": given a circle C of radius 1, construct a square S with side length r such that the areas of the circle and square are the same.

Prove that if we can square the circle in this sense, then $\sqrt{\pi}$ is a constructible number. If we further *assume*[5] that if (u, v) is a constructible point, then neither u nor v can be real transcendental numbers. Show that under this assumption, $\sqrt{\pi}$ is *not* a constructible number, so that we *cannot* square the circle.

3.3 Another problem from antiquity is "doubling the cube": given a cube with side lengths 1, construct a cube whose volume is double that of the unit cube.

Show that a cube with double the volume of a unit cube must have side length $\sqrt[3]{2}$. Note that $\sqrt[3]{2}$ is a root of the polynomial $x^3 - 2$, which has degree 3. Prove that $\sqrt[3]{2}$ is not the root of any nonzero polynomial with integer coefficents that has degree 1 or 2. From Theorem 2.1, deduce that $\sqrt[3]{2}$ is *not* a constructible number.

3.4 In our session on September 23 on irrational numbers, we proved that $\cos 20^\circ = \cos \frac{\pi}{9}$ (measuring the angle in radians) is an irrational number. Today, we show this value is not constructible, either. This exercise is adapted from Chapter 5 of [1].

As a reminder, let us recall the Rational Root Theorem:

Theorem 3.1 (The Rational Root Theorem)**.** *Let n be a positive integer, and let $a_n, a_{n-1}, \ldots, a_1, a_0$ be integers with $a_n \neq 0$. Consider the polynomial*

$$p(x) := a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0. \tag{3.1}$$

Let q be a rational number $q = a/b$ such that a, b are integers, $b \neq 0$, and $\gcd(a, b) = 1$.

If q is a root *or* zero *of $p(x)$—i.e., if $p(q) = 0$, meaning we get* 0 *when we "plug q into $p(x)$"—then*

[5] This result can indeed be proven, to be clear, but a full, rigorous proof is beyond the scope of what we can accomplish in the time available. That said, the results in Exercises#2.3–2.5 make this plausible: in building a constructible point, we're (at most) adding square roots from terms in previous constructible points. If that's all we're doing, it seems unlikely we could produce a transcendental number in this way.

- *a, the numerator of q, is a divisor*[6] *of the constant term a_0, and*
- *b, the denominator of q, is a divisor of the leading term a_n.*

That is, the only rational numbers a/b (with $\gcd(a,b)=1$*) that are even* candidates *to be roots of $p(x)$ must be such that the numerator is a divisor of the constant term, and the denominator is a divisor of the leading coefficient.*

(a) What is $\cos 60^\circ = \cos \pi/6$ (in radians)?

(b) Using the identity

$$\cos 3\theta = 4(\cos\theta)^3 - 3\cos\theta$$

for all θ, conclude that $\cos 20^\circ$ is a root of the polynomial

$$p(x) = 8x^3 - 6x - 1.$$

(c) Using the Rational Root Theorem, prove that $\cos 20^\circ$ is irrational.

(d) Since $\cos 20^\circ$ is the root of an *prime* polynomial of degree 3—which is not a power of 2—deduce that $\cos 20^\circ$ is not a constructible number. Show that it follows that, in particular, we cannot trisect a $60^\circ = \frac{\pi}{3}$ (in radians) angle, meaning it is impossible to have a construction method to trisect a general angle.

[6] Note that we are simply requiring that $\gcd(a,b)=1$, *not* that a/b is lowest terms in the sense of the denominator being strictly positive. In particular, we may have either of a and b is negative, not our lowest-terms convention that b must be strictly positive.

References

[1] Ivan Niven. *Numbers: Rational and Irrational*. New Mathematical Library. Random House and The L. W. Singer Company, New York, 1961.