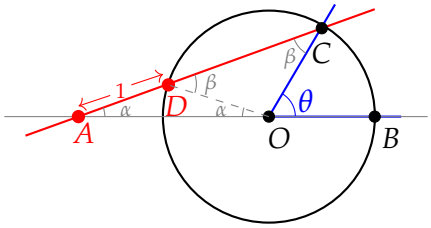
I.B. Constructible points

The impossibility of trisecting an angle with straightedge and compass is a celebrated consequence of the theory of field extensions. But for it to have meaning, we have to say what we mean by "with straightedge and compass".

Suppose we have an angle θ, as shown in blue. Use a compass to put a circle of radius[4] 1 about it; and then draw the line OB with a straightedge:



Now we do something violent. Keeping the compass set to radius 1, we jam its sharp points into the straightedge, which we lay down on some line through C. Wiggle and shift the straightedge around, keeping it so that it passes through C, until the two points of the compass lie on OB and the circle. The segment $\overline{AD}$ then has length 1, making $\triangle ODA$ isoceles. From the figure, we have

$$\pi = \theta + (\pi - 2\beta) + \alpha \quad \Longrightarrow \quad 2\beta = \alpha + \theta$$

from angles at O, and

$$\pi = \beta + (\pi - 2\alpha) \quad \Longrightarrow \quad \beta = 2\alpha$$

from angles at D. Conclude that $3\alpha = \theta$.

This is *not* a construction with straightedge and compass in the Euclidean sense, because we used a *marked straightedge*. That is forbidden! We are only supposed to use the straightedge to draw a line between two preëxisting points. With the compass, we are allowed to set its radius to the distance between any two such points, and draw a circle of that radius about a third point. We can mark new

[4] You always give yourself the points 0 and 1 in Euclidean constructions; and in any case we could just open the compass and define "1" to be the distance between its two points.

points where these lines and circles intersect, and use them as just described to create new lines and circles, and so on. That's it. Of course, what you can construct in this way will depend on what set of points you start with.

To turn this geometry into algebra, we identify the plane with $\mathbb{C}$, and give ourselves a finite subset $\mathcal{S} \subset \mathbb{C}$ containing at least $\{0,1\}$. (For instance, the situation of a given angle θ above corresponds to taking $\mathcal{S} = \{0, 1, e^{\mathbf{i}\theta}\}$.) Set $\mathcal{S}_{(1)} := \mathcal{S}$, and define inductively

$$\begin{aligned}\mathcal{S}_{(m)} := \mathcal{S}_{(m-1)} &\cup \left\{ \overline{p_1p_2} \cap \overline{p_3p_4} \;\middle|\; p_i \in \mathcal{S}_{(m-1)},\ \overline{p_1p_2} \neq \overline{p_3p_4} \right\} \\ &\cup \left\{ C_{|p_1-p_2|}(p_3) \cap C_{|p_4-p_5|}(p_6) \;\middle|\; p_i \in \mathcal{S}_{(m-1)}, p_3 \neq p_6 \right\} \\ &\cup \left\{ C_{|p_1-p_2|}(p_3) \cap \overline{p_4p_5} \;\middle|\; p_i \in \mathcal{S}_{(m-1)} \right\},\end{aligned} \tag{I.B.1}$$

where $\overline{pq}$ is the line through p, q and $C_r(p)$ the circle of radius r about p. Setting

$$\mathfrak{C}(\mathcal{S}) := \bigcup_{m>0} \mathcal{S}_{(m)},$$

we make the

I.B.2. DEFINITION. $P \in \mathbb{C}$ is **constructible from $\mathcal{S}$** if $P \in \mathfrak{C}(\mathcal{S})$. We will simply say P is **constructible** if it is constructible from $\{0,1\}$.

I.B.3. PROPOSITION. *$\mathfrak{C}(\mathcal{S})$ is the smallest[5] subfield of $\mathbb{C}$ containing $\mathcal{S}$ and closed under two algebraic operations*: *taking square roots and complex conjugates.*

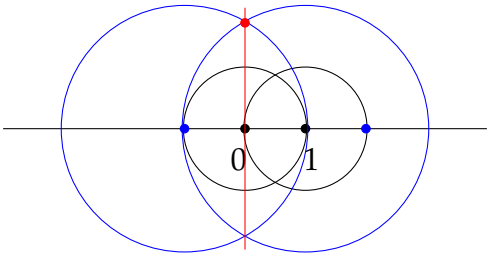
PROOF. There are two inclusions to verify:

(a) *$\mathfrak{C}(\mathcal{S})$ is contained in any subfield of $\mathbb{C}$ containing $\mathcal{S}$ and closed under the two algebraic operations.* This is checked by showing that any such field is closed under the three Euclidean operations (by which one passes from $\mathcal{S}_{(m-1)}$ to $\mathcal{S}_{(m)}$) in (I.B.1).

(b) *$\mathfrak{C}(\mathcal{S})$ is such a subfield, so in particular contains the smallest one.* For this, we need to demonstrate that $\mathfrak{C}(\mathcal{S})$ is closed under all field theory operations, square roots, and complex conjugation.

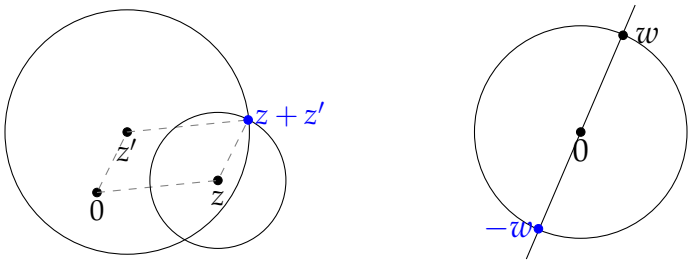
[5]Equivalently, it is the *intersection* of all subfields containing $\mathcal{S}$ and closed under square roots and complex conjugation.

We carry out (b) first. Note that, given $\{0,1\}$ we can construct the real and imaginary axes: for the latter, use the picture



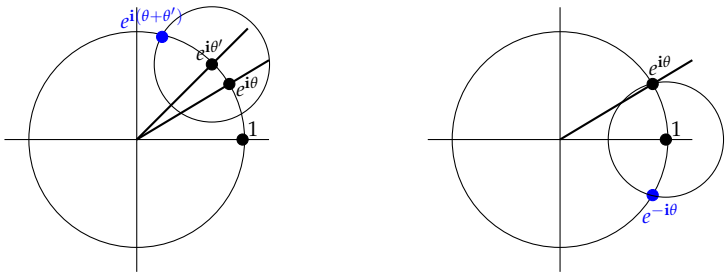
where blue points are in $\mathcal{S}_{(2)}$ and red in $\mathcal{S}_{(3)}$. Notice that the red point is a square root of -2, and we also get $\mathbf{i} = \sqrt{-1}$ (in $\mathcal{S}_{(4)}$).

Here are pictures for closure under addition and additive inversion:

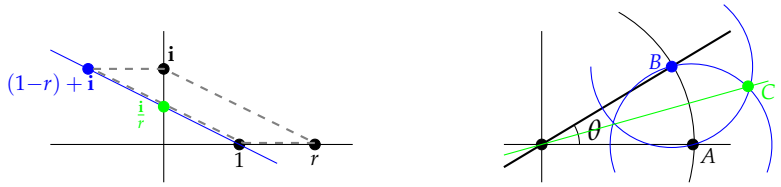


where the circles on the left are $C_{|z|}(z')$ and $C_{|z'|}(z)$. The point is that since $|(z+z')-z| = |z'|$ and $|(z+z')-z'| = |z|$, $z+z'$ is one of the intersection points.

For multiplication, inversion, complex conjugation, and square roots, we can split $z = re^{\mathbf{i}\theta}$ into modulus r and argument θ by constructing $C_{|z|}(0)$ resp. $\overline{0z}$; and we can reassemble it from these. Multiplication and square root of moduli are left to the exercises. Here are diagrams for multiplying and inverting the $e^{\mathbf{i}\theta}$ parts:



In the left picture, the smaller circle is $C_{|e^{\mathbf{i}\theta}-1|}(e^{\mathbf{i}\theta'})$, which works because $|e^{\mathbf{i}(\theta+\theta')} - e^{\mathbf{i}\theta'}| = |e^{\mathbf{i}\theta} - 1|$. Finally, for $1/r$ and $(e^{\mathbf{i}\theta})^{\frac{1}{2}}$ (angle bisection) we have



where on the left, we have $\mathbf{i}$ from above, $(1-r)+\mathbf{i}$ is obtained by closure under addition (and additive inversion), and it is enough to get any point of modulus $\frac{1}{r}$. On the right, $\triangle AOC$ and $\triangle BOC$ are similar. This completes (b).

Turning to (a), let $\mathbb{F} \subset \mathbb{C}$ be a subfield closed under square roots and complex conjugation. Clearly $\mathbf{i} \in \mathbb{F}$. So for any $z = x + \mathbf{i}y \in \mathbb{F}$, we have $x = \frac{1}{2}(z + \bar{z})$ and $y = \frac{1}{2\mathbf{i}}(z - \bar{z})$ in $\mathbb{F}_0 := \mathbb{F} \cap \mathbb{R}$. Conversely, any ordered pair $(x, y) \in \mathbb{F}_0^2 (= \mathbb{F}_0 \times \mathbb{F}_0)$ gives rise to an element $x + \mathbf{i}y \in \mathbb{F}$. (Indeed, we have $\mathbb{F} = \mathbb{F}_0[\sqrt{-1}]$.) This identification means that we can test the Euclidean closure conditions on ordered pairs in $\mathbb{F}_0^2 \subset \mathbb{R}^2$; all we need to know about $\mathbb{F}_0$ is that it is a subfield of $\mathbb{R}$ which is closed under taking square roots *that remain real*.

Now consider the three Euclidean operations described in (I.B.1):

- Lines through pairs of points in $\mathbb{F}_0^2$ have equations with coefficients in $\mathbb{F}_0$. So by Cramer's rule, the solution lies in $\mathbb{F}_0^2$.
- The circle about a point in $\mathbb{F}_0^2$ with radius the distance between two points in $\mathbb{F}_0^2$ has equation $x^2 + y^2 + 2dx + 2ey + f = 0$, with $d, e, f \in \mathbb{F}_0$. Consider a line $ax + by + c = 0$ with $a, b, c \in \mathbb{F}_0$; we may assume $a \neq 0$ (otherwise swap x, y). Substituting into the conic equation gives a quadratic equation in y. If the roots are non-real (unreal?) then the line and circle don't meet. If they are real, the roots are in $\mathbb{F}_0$ since they are obtained via square roots.
- Given two circles of the above form, subtracting their equations gives a linear equation, and we repeat the last bullet.

This completes the proof of (a). □

Write $\mathcal{S} = \{0, 1, z_1, z_2, \ldots, z_n\}$ and $F := \mathbb{Q}(z_1, \ldots, z_n, \bar{z}_1, \ldots, \bar{z}_n)$. A *square-root tower* over $F =: F_1$ is a field $E =: F_m$ that is reached by a sequence of extensions $F_{i+1} := F_i(u_i)$ with $u_i^2 \in F_i$.

I.B.4. THEOREM. *$z \in \mathbb{C}$ is constructible from $\mathcal{S}$ if and only if it is contained in a square-root tower over F.*

PROOF. By I.B.3, $\mathfrak{C}(\mathcal{S})$ contains F and any square-root tower E over it, hence the union $\mathfrak{F}$ of all such towers. I claim that $\mathfrak{F}$ *is a field closed under conjugates and square roots* (and containing F), hence contains the smallest such field $\mathfrak{C}(\mathcal{S})$. So $\mathfrak{F} \supset \mathfrak{C}(\mathcal{S}) \supset \mathfrak{F}$ and they are equal. This is exactly the statement of the Theorem: that the set of numbers constructible from $\mathcal{S}$ equals the set of numbers contained in square-root towers over F.

We just need to check the claim. Given $z, z' \in \mathfrak{F}$, we have $z \in E = F(u_1, \ldots, u_r)$ and $z' \in E' = F(u_1', \ldots, u_t')$; but then the compositum $E'' = F(u_1, \ldots, u_r; u_1', \ldots, u_t')$ is a square-root tower containing both z and z', and thus their product, sum, inverses, etc. Clearly (by definition) $\mathfrak{F}$ is closed under square roots; and since the complex conjugate of a square-root tower over F is a square root tower over $\bar{F} = F$, $\mathfrak{F}$ is closed under conjugation. □

I.B.5. COROLLARY. *If $z \in \mathbb{C}$ is constructible, then $[\mathbb{Q}(z):\mathbb{Q}] = 2^s$ for some $s \in \mathbb{N}$.*

PROOF. By I.B.4, $z \in \mathbb{Q}(u_1, \ldots, u_r) =: E$, with $u_i^2 \in F_i$ and (we may assume) $u_i \notin F_i$. The Tower Law gives $[E:\mathbb{Q}] = \prod_{i=1}^{r} [F_{i+1}:F_i] = 2^r$ since the degree of each minimal polynomial is 2. Since $\mathbb{Q} \subseteq \mathbb{Q}(z) \subseteq E$, we have $[\mathbb{Q}(z):\mathbb{Q}] \,|\, [E:\mathbb{Q}]$ (by I.A.9). □

I.B.6. REMARK. (i) An immediate consequence of I.B.5 is that the constructible numbers $\mathfrak{C} := \mathfrak{C}(\{0,1\})$ are contained in $\bar{\mathbb{Q}}$. (By I.B.3, we also know that $\mathfrak{C}$ is a field.)

(ii) Another variant of I.B.5 is that if (x, y) (i.e. $x + \mathbf{i}y$) is a constructible point, then $[\mathbb{Q}(x, y):\mathbb{Q}]$ is a power of 2. Argue as follows: constructibility of $z = x + \mathbf{i}y$ is equivalent to that of x and y ($\mathfrak{C}$ is a field containing $\mathbf{i}$ and closed under conjugation); and if x and y are

both contained in square-root towers, then concatenating the towers gives a tower.

(iii) The converse of I.B.5 (and (ii)) is completely false: e.g., if z is the root of a general quartic polynomial over $\mathbb{Q}$, then $[\mathbb{Q}(z):\mathbb{Q}] = 4$ but z does not lie in a square-root tower.

Geometric applications. The only reason I don't call the next three results Corollaries (of Theorem I.B.4, which they are) is that they resolve problems of classical antiquity. They were all proved by Wantzel in 1837 (in essentially this way).[6]

I.B.7. THEOREM. *A general*[7] *angle cannot be trisected with* (*unmarked*) *straightedge and compass.*

PROOF. Assume otherwise: that given $\mathcal{S} = \{0, 1, e^{\mathbf{i}\theta}\}$ (any θ), we have $e^{\mathbf{i}\theta/3} \in \mathfrak{C}(\mathcal{S})$. Then this should be true in particular for $\theta = \pi/3$. Since $\zeta_6 = e^{\mathbf{i}\pi/3} \in \mathfrak{C}$, in this case we have $\mathfrak{C}(\mathcal{S}) = \mathfrak{C}$. So our assumption implies $\zeta_{18} \in \mathfrak{C}$, hence $\xi := 2\cos(\pi/9) = \zeta_{18} + \bar{\zeta}_{18} \in \mathfrak{C}$. By I.B.5, we conclude that $[\mathbb{Q}(\xi):\mathbb{Q}]$ is a power of 2.

But now we recall from I.A.2 that ξ is a root of the irreducible cubic polynomial $f(x) = x^3 - 3x - 1$. Indeed,

$$\begin{aligned}(\zeta_{18} + \bar{\zeta}_{18})^3 - 3(\zeta_{18} + \bar{\zeta}_{18}) - 1 &= \zeta_6 + \bar{\zeta}_6 + 3\zeta_{18} + 3\bar{\zeta}_{18} - 3(\zeta_{18} + \bar{\zeta}_{18}) - 1 \\ &= \zeta_6 + \bar{\zeta}_6 - 1 = \tfrac{1+\sqrt{-3}}{2} + \tfrac{1-\sqrt{-3}}{2} - 1 = 0.\end{aligned}$$

This gives $[\mathbb{Q}(\xi):\mathbb{Q}] = 3$, a contradiction. □

I.B.8. THEOREM. *The cube cannot be duplicated* (*doubled in volume*) *with straightedge and compass.*

PROOF. $[\mathbb{Q}(\sqrt[3]{2}):\mathbb{Q}] = 3$. □

Next, the constructibility of a regular p-gon (p prime) hinges on that of $\zeta_p = e^{2\pi\mathbf{i}/p}$. For now, we limit ourselves to the following necessary condition and three examples. A *Fermat prime* is any member

[6] Wantzel proved a stronger "if and only if" statement than our I.B.9; we'll get to that later.

[7] We have to say "general", because of angles like $\pi/2$: $e^{\mathbf{i}\pi/6}$ is quadratic, hence constructible, and trisects it.

of the sequence $1+2^{2^s}$ ($s \in \mathbb{N}$) which is prime; indeed, it begins with the five prime numbers $3, 5, 17, 257, 65537$. It is conjectured, on good evidence, that these are the only Fermat primes.

I.B.9. THEOREM. $\zeta_p \in \mathfrak{C} \implies$ *p is a Fermat prime.*

PROOF. Recall that the p^{th} **cyclotomic polynomial**

$$\Phi_p(x) := \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + 1$$

is irreducible in $\mathbb{Q}[x]$. (Apply Gauss and Eisenstein to $\Phi_p(y+1) = \frac{(y+1)^p-1}{(y+1)-1} = y^{p-1} + py^{p-2} + \binom{p}{2}y^{p-3} + \cdots + \binom{p}{p-2}y + p$.) This means that $[\mathbb{Q}(\zeta_p):\mathbb{Q}] = p - 1$. So by I.B.5, constructibility implies $p - 1 = 2^n$, i.e. $p = 1 + 2^n$, for some $n \in \mathbb{N}$.

Now given any odd integer $2k + 1 > 1$, we can factor

$$A^{2k+1} + 1 = (A + 1)(A^{2k} - A^{2k-1} + \cdots - A + 1).$$

If $2k + 1$ divides n, then (writing $m := \frac{n}{2k+1}$ and $A = a^m$) this

$$\begin{aligned}
&\implies \quad a^n + 1 = (a^m + 1)(a^{2km} - a^{(2k-1)m} + \cdots - a^m + 1) \\
&\implies \quad 2^n + 1 = (2^m + 1)(2^{2km} - 2^{(2k-1)m} + \cdots - 2^m + 1) \\
&\implies \quad 2^n + 1 \text{ is not prime,}
\end{aligned}$$

a contradiction. So n must be a power of 2. □

Unfortunately this doesn't prove that any p-gon *is* constructible. For now, we will content ourselves with establishing that for the first three Fermat numbers.

Clearly ζ_3, being quadratic, satisfies I.B.4; and the construction of a regular triangle is really easy.

For $p = 5$, you are asked to do the geometric construction in the HW; on the algebraic side, one verifies that $\zeta_5 =: x + \mathbf{i}y$ is contained

in a square-root tower as follows:

$$\begin{aligned} \zeta_5^3 &= \overline{\zeta}_5^2 \\ x^3 + 3\mathbf{i}yx^2 - 3y^2x - \mathbf{i}y^3 &= x^2 - 2\mathbf{i}yx - y^2 \\ yx^2 - y^3 &= -2yx \qquad \text{[take imaginary parts]} \\ 4x^2 - 1 = 3x^2 - y^2 &= -2x \qquad \text{[use } x^2 + y^2 = 1\text{]} \\ (x, y) &= \left(\tfrac{-1 \pm \sqrt{5}}{4}, \pm\sqrt{\tfrac{5 \pm \sqrt{5}}{8}} \right). \end{aligned}$$

For $p = 17$, things get much more complicated, since four iterated square-roots are required. But we can summarize the algebra as follows: we are after $\sin(\theta)$ and $z' := \cos(\theta)$, where $\theta = \frac{2\pi}{17}$. Partition $\mathbb{Z}_{17}^* = P_1 \sqcup P_2 \sqcup P_3 \sqcup P_6$, where $P_k := \{\pm k, \pm 4k\}$; and set $y_k := \sum_{j \in P_k} \zeta_{17}^j$. Defining $x_+ := y_1 + y_2$ and $x_- := y_3 + y_6$, we compute

$$y_1 y_2 = y_3 y_6 = x_+ + x_- = \textstyle\sum_{j=1}^{16} \zeta_{17}^j = -1$$

and $x_+ x_- = 4 \sum_{j=1}^{16} \zeta_{17}^j = -4$. This gives

$$(x - x_+)(x - x_-) = x^2 + 4x - 1 \quad \Longrightarrow \quad x_\pm = \tfrac{1}{2}(-1 \pm \sqrt{17})$$

and

$$(y - y_1)(y - y_2) = y^2 - x_+ y - 1, \quad (y - y_3)(y - y_6) = y^2 - x_- y - 1,$$

so that extending $\mathbb{Q}$ first by $\sqrt{17}$, then by y_1 and by y_3 gives a square-root tower (of degree 8 over $\mathbb{Q}$).[8] Now write $z'' = \cos(4\theta)$, so that $z' + z'' = y_1$ and $z'z'' = y_3$; taking a final quadratic extension by the root z' of

$$(z - z')(z - z'') = z^2 - y_1 z + y_3$$

produces a degree 16 square-root tower containing z' and $\sin(\theta) = 1 - (z')^2 = 1 + y_3 - y_1 z'$. It was actually Gauss (1796, at 19) who first showed a regular 17-gon is constructible.

[8] Note that adjoining one root of a quadratic equation gives you both, by the quadratic formula. So the degree-8 tower actually contains $x_\pm$ and all the $\{y_j\}$; and the degree-16 tower contains z''.