# Math-123: straightedge and compass constructions

Sebastien Vasey

Harvard University

March 13, 2020

Can the following be done, using just straightedge and compass?

Can the following be done, using just straightedge and compass?

1. (Doubling the cube) Given a line segment L, construct another line segment L' so that a cube with side L' has exactly twice the volume of a cube with side L.

Can the following be done, using just straightedge and compass?

- 1. (Doubling the cube) Given a line segment L, construct another line segment L' so that a cube with side L' has exactly twice the volume of a cube with side L.
- 2. (Trisecting an angle) Given an angle  $\theta$ , construct the angle  $\theta/3$ .

Can the following be done, using just straightedge and compass?

- 1. (Doubling the cube) Given a line segment L, construct another line segment L' so that a cube with side L' has exactly twice the volume of a cube with side L.
- 2. (Trisecting an angle) Given an angle  $\theta$ , construct the angle  $\theta/3$ .
- 3. (Squaring the circle) Given a circle, construct a square with the same area as this circle.

Can the following be done, using just straightedge and compass?

- 1. (Doubling the cube) Given a line segment L, construct another line segment L' so that a cube with side L' has exactly twice the volume of a cube with side L.
- 2. (Trisecting an angle) Given an angle  $\theta$ , construct the angle  $\theta/3$ .
- 3. (Squaring the circle) Given a circle, construct a square with the same area as this circle.

Today, we will show they are all impossible to do, by translating to algebra!

Can the following be done, using just straightedge and compass?

- 1. (Doubling the cube) Given a line segment L, construct another line segment L' so that a cube with side L' has exactly twice the volume of a cube with side L.
- 2. (Trisecting an angle) Given an angle  $\theta$ , construct the angle  $\theta/3$ .
- 3. (Squaring the circle) Given a circle, construct a square with the same area as this circle.

Today, we will show they are all impossible to do, by translating to algebra!

[Note: Why is it "straightedge and compass" rather than "ruler and compass"? Because rulers have markings, and we are not allowed to use them. In fact it is possible to trisect an angle using a ruler's markings. See the book!]

Let S be a set of points in the plane. A line is S-constructed if it contains two distinct points of S. A circle is S-constructed if it contains a point in S, and its center is also in S.

Let S be a set of points in the plane. A line is S-constructed if it contains two distinct points of S. A circle is S-constructed if it contains a point in S, and its center is also in S.

#### Definition

The set of *constructible points of the plane* is the smallest subset S of  $\mathbb{R}^2$  with the following properties:

1.  $(0,0) \in S$  and  $(1,0) \in S$ .

Let S be a set of points in the plane. A line is S-constructed if it contains two distinct points of S. A circle is S-constructed if it contains a point in S, and its center is also in S.

#### Definition

The set of *constructible points of the plane* is the smallest subset S of  $\mathbb{R}^2$  with the following properties:

- 1.  $(0,0) \in S$  and  $(1,0) \in S$ .
- 2. If two non-parallel S-constructed lines intersect at the point P, then  $P \in S$ .

Let S be a set of points in the plane. A line is S-constructed if it contains two distinct points of S. A circle is S-constructed if it contains a point in S, and its center is also in S.

#### Definition

The set of *constructible points of the plane* is the smallest subset S of  $\mathbb{R}^2$  with the following properties:

- 1.  $(0,0) \in S$  and  $(1,0) \in S$ .
- 2. If two non-parallel S-constructed lines intersect at the point P, then  $P \in S$ .
- 3. If  $C_1$  and  $C_2$  are distinct S-constructed circles which intersect at a point P, then  $P \in S$ .

Let S be a set of points in the plane. A line is S-constructed if it contains two distinct points of S. A circle is S-constructed if it contains a point in S, and its center is also in S.

#### Definition

The set of *constructible points of the plane* is the smallest subset S of  $\mathbb{R}^2$  with the following properties:

- 1.  $(0,0) \in S$  and  $(1,0) \in S$ .
- 2. If two non-parallel S-constructed lines intersect at the point P, then  $P \in S$ .
- 3. If  $C_1$  and  $C_2$  are distinct S-constructed circles which intersect at a point P, then  $P \in S$ .
- 4. If C is an S-constructed circle and L is an S-constructed line which intersects C at a point P, then  $P \in S$ .

In words, a point of the plane is constructible if it can be obtained from (0,0) and (1,0) using straightedge and compass operations:

In words, a point of the plane is constructible if it can be obtained from (0,0) and (1,0) using straightedge and compass operations:

1. Drawing a line between already constructed points.

In words, a point of the plane is constructible if it can be obtained from (0,0) and (1,0) using straightedge and compass operations:

- 1. Drawing a line between already constructed points.
- 2. Drawing a circle with center a constructed point and also passing through a constructed point.

In words, a point of the plane is constructible if it can be obtained from (0,0) and (1,0) using straightedge and compass operations:

- 1. Drawing a line between already constructed points.
- 2. Drawing a circle with center a constructed point and also passing through a constructed point.
- 3. Finding points of intersections of two non-parallel lines, or of two distinct circles, or of a line and a circle.

In words, a point of the plane is constructible if it can be obtained from (0,0) and (1,0) using straightedge and compass operations:

- 1. Drawing a line between already constructed points.
- 2. Drawing a circle with center a constructed point and also passing through a constructed point.
- 3. Finding points of intersections of two non-parallel lines, or of two distinct circles, or of a line and a circle.

#### Example

The point (2,0) is constructible: draw the line *L* between (0,0) and (1,0), find the intersection with the circle with center (1,0) passing through (0,0). Similarly, (n,0) is constructible for  $n \in \mathbb{Z}$ .

A line is *constructible* if it passes through two distinct constructible points, and a circle is *constructible* if its center is constructible and it passes through a constructible point.

A line is *constructible* if it passes through two distinct constructible points, and a circle is *constructible* if its center is constructible and it passes through a constructible point.

```
Lemma (Exercise)
```

Assume L is a constructible line and P is a constructible point. Then:

► The line L<sup>⊥</sup> perpendicular to L and passing through P is constructible.

A line is *constructible* if it passes through two distinct constructible points, and a circle is *constructible* if its center is constructible and it passes through a constructible point.

### Lemma (Exercise)

Assume L is a constructible line and P is a constructible point. Then:

- ► The line L<sup>⊥</sup> perpendicular to L and passing through P is constructible.
- ► The line L<sup>||</sup> parallel to L and passing through P is constructible.

A line is *constructible* if it passes through two distinct constructible points, and a circle is *constructible* if its center is constructible and it passes through a constructible point.

#### Lemma (Exercise)

Assume L is a constructible line and P is a constructible point. Then:

- ► The line L<sup>⊥</sup> perpendicular to L and passing through P is constructible.
- ► The line L<sup>||</sup> parallel to L and passing through P is constructible.

It follows that (0,1) is constructible: draw a line L from (0,0) to (1,0), draw the perpendicular line  $L^{\perp}$  through (0,0), and find its intersection with a circle of radius 1 centered at (0,0).

### Definition

A real number r is *constructible* if |r| is the length of a segment between two constructible points.

### Definition

A real number r is *constructible* if |r| is the length of a segment between two constructible points.

Lemma (Exercise)

• r is a constructible number if and only if (r, 0) is constructible.

### Definition

A real number r is *constructible* if |r| is the length of a segment between two constructible points.

### Lemma (Exercise)

- r is a constructible number if and only if (r, 0) is constructible.
- ► A point (x, y) is constructible if and only if both x and y are constructible.

### Definition

A real number r is *constructible* if |r| is the length of a segment between two constructible points.

### Lemma (Exercise)

- r is a constructible number if and only if (r, 0) is constructible.
- ► A point (x, y) is constructible if and only if both x and y are constructible.
- A circle is constructible if and only if its center is a constructible point and its radius is a constructible number.

#### Lemma

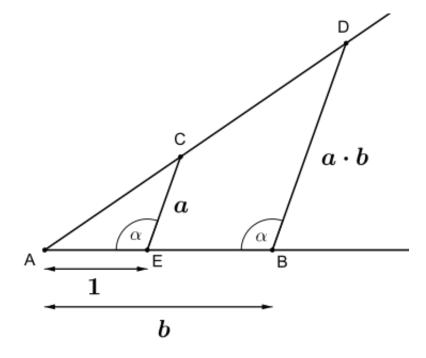
If a and b are constructible numbers, then a + b, -a,  $a \cdot b$ ,  $a^{-1}$   $(a \neq 0)$ , and  $\sqrt{a}$   $(a \ge 0)$  are constructible.

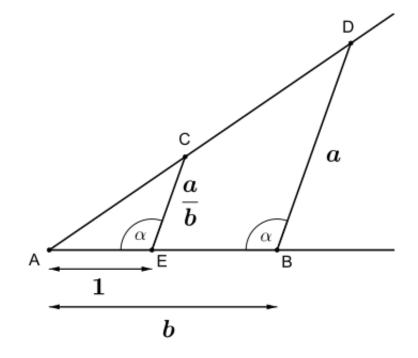
#### Lemma

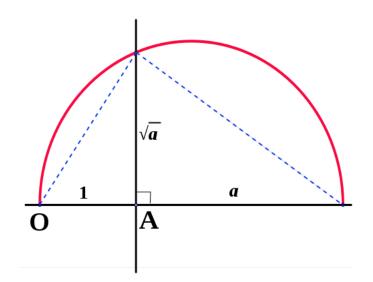
If a and b are constructible numbers, then a + b, -a,  $a \cdot b$ ,  $a^{-1}$   $(a \neq 0)$ , and  $\sqrt{a}$   $(a \ge 0)$  are constructible.

#### Proof.

|-a| = a, and a + b is the distance between (b, 0) and (-a, 0). For products and inverses, use similar triangles. For roots, use a circle of diameter 1 + a (pictures to follow).







Thus a real number is constructible if and only if it can be obtained from 0 and 1 using field operations and square roots.

Thus a real number is constructible if and only if it can be obtained from 0 and 1 using field operations and square roots. In field-theoretic terminology:

#### Theorem

A real number x is constructible if and only if it is contained in an extension K of  $\mathbb{Q}$  which is an iteration of quadratic extensions: there are subfields  $\mathbb{Q} = F_1 \subseteq F_2 \subseteq \ldots \subseteq F_k = K$  such that  $F_1 = \mathbb{Q}$  and for all i < k,  $F_{i+1} = F_i(\sqrt{\alpha_i})$ , for  $\alpha_i$  a non-negative real number in  $F_i$ .

Thus a real number is constructible if and only if it can be obtained from 0 and 1 using field operations and square roots. In field-theoretic terminology:

#### Theorem

A real number x is constructible if and only if it is contained in an extension K of  $\mathbb{Q}$  which is an iteration of quadratic extensions: there are subfields  $\mathbb{Q} = F_1 \subseteq F_2 \subseteq \ldots \subseteq F_k = K$  such that  $F_1 = \mathbb{Q}$  and for all i < k,  $F_{i+1} = F_i(\sqrt{\alpha_i})$ , for  $\alpha_i$  a non-negative real number in  $F_i$ .

#### Proof sketch.

We just saw that numbers arising from iterated quadratic extensions are constructible. Conversely, any constructible number arises from taking sums, products, and finding root of degree two polynomials using the quadratic formula (to find intersection points of lines and circles).

#### Theorem

A real number x is constructible if and only if it is contained in an extension K of  $\mathbb{Q}$  which is an iteration of quadratic extensions: there are subfields  $\mathbb{Q} = F_1 \subseteq F_2 \subseteq \ldots \subseteq F_k = K$  such that  $F_1 = \mathbb{Q}$  and for all i < k,  $F_{i+1} = F_i(\sqrt{\alpha_i})$ , for  $\alpha_i$  a non-negative real number in  $F_i$ .

#### Theorem

A real number x is constructible if and only if it is contained in an extension K of  $\mathbb{Q}$  which is an iteration of quadratic extensions: there are subfields  $\mathbb{Q} = F_1 \subseteq F_2 \subseteq \ldots \subseteq F_k = K$  such that  $F_1 = \mathbb{Q}$  and for all i < k,  $F_{i+1} = F_i(\sqrt{\alpha_i})$ , for  $\alpha_i$  a non-negative real number in  $F_i$ .

Because the degree of an extension must divide the degree of a bigger extension, we have:

Corollary

If x is constructible, then  $[\mathbb{Q}(x) : \mathbb{Q}] = 2^k$  for some k.

### Corollary

If x is constructible,  $[\mathbb{Q}(x) : \mathbb{Q}] = 2^k$  for some k.

### Corollary

If x is constructible,  $[\mathbb{Q}(x) : \mathbb{Q}] = 2^k$  for some k.

Thus we cannot double the cube: take a cube of side length 1, which has volume 1. A cube of double that volume would have side length  $\sqrt[3]{2}$ , which has degree 3, which is not a power of 2.

#### Corollary

## If x is constructible, $[\mathbb{Q}(x) : \mathbb{Q}] = 2^k$ for some k.

Thus we cannot double the cube: take a cube of side length 1, which has volume 1. A cube of double that volume would have side length  $\sqrt[3]{2}$ , which has degree 3, which is not a power of 2.

We also cannot square the circle: a circle of unit radius has area  $\pi$ , and since  $\pi$  (so also  $\sqrt{\pi}$ ) is transcendental, its degree is infinite, hence not a power of 2.

By "an angle  $\theta$  can be constructed" we mean (by definition) that a point in the plane with  $\theta$  as part of its polar coordinates can be constructed.

By "an angle  $\theta$  can be constructed" we mean (by definition) that a point in the plane with  $\theta$  as part of its polar coordinates can be constructed.

Exercise: an angle  $\theta$  can be constructed if and only if  $\cos(\theta)$  is a constructible real number.

By "an angle  $\theta$  can be constructed" we mean (by definition) that a point in the plane with  $\theta$  as part of its polar coordinates can be constructed.

Exercise: an angle  $\theta$  can be constructed if and only if  $\cos(\theta)$  is a constructible real number.

Thus the problem of trisecting the angle asks whether  $cos(\theta)$  constructible implies  $cos(\theta/3)$  constructible.

By "an angle  $\theta$  can be constructed" we mean (by definition) that a point in the plane with  $\theta$  as part of its polar coordinates can be constructed.

Exercise: an angle  $\theta$  can be constructed if and only if  $\cos(\theta)$  is a constructible real number.

Thus the problem of trisecting the angle asks whether  $\cos(\theta)$  constructible implies  $\cos(\theta/3)$  constructible.

Note that  $\cos(\theta) = 2\cos^2(\theta/2) - 1$  (double angle formula), so if we know  $\cos(\theta)$ , we can get  $\cos(\theta/2)$  by solving a quadratic. Thus bisecting the angle is possible (try to find a geometric construction!)

## Trisecting the angle is impossible

By contrast,  $\cos(\theta) = 4\cos^3(\theta/3) - 3\cos(\theta/3)$  (triple angle formula). Set  $\theta = \pi/3$  (60 degrees). Recall that  $\cos(\theta) = \frac{1}{2}$ , so if  $\beta = \cos(\theta/3)$ , then we must have  $4\beta^3 - 3\beta - \frac{1}{2} = 0$ , or  $8\beta^3 - 6\beta - 1 = 0$ .

## Trisecting the angle is impossible

By contrast,  $\cos(\theta) = 4\cos^3(\theta/3) - 3\cos(\theta/3)$  (triple angle formula). Set  $\theta = \pi/3$  (60 degrees). Recall that  $\cos(\theta) = \frac{1}{2}$ , so if  $\beta = \cos(\theta/3)$ , then we must have  $4\beta^3 - 3\beta - \frac{1}{2} = 0$ , or  $8\beta^3 - 6\beta - 1 = 0$ .

Set  $\alpha = 2\beta$  to clarify things. Then the above equation becomes  $\alpha^3 - 3\alpha - 1 = 0$ . The left hand side is an irreducible polynomial in  $\alpha$  (it has no rational roots, check it!), so  $\alpha$ , hence  $\beta$ , has degree 3, hence is not constructible.

# Trisecting the angle is impossible

By contrast,  $\cos(\theta) = 4\cos^3(\theta/3) - 3\cos(\theta/3)$  (triple angle formula). Set  $\theta = \pi/3$  (60 degrees). Recall that  $\cos(\theta) = \frac{1}{2}$ , so if  $\beta = \cos(\theta/3)$ , then we must have  $4\beta^3 - 3\beta - \frac{1}{2} = 0$ , or  $8\beta^3 - 6\beta - 1 = 0$ .

Set  $\alpha = 2\beta$  to clarify things. Then the above equation becomes  $\alpha^3 - 3\alpha - 1 = 0$ . The left hand side is an irreducible polynomial in  $\alpha$  (it has no rational roots, check it!), so  $\alpha$ , hence  $\beta$ , has degree 3, hence is not constructible.

We just saw that  $\cos(\frac{1}{9}\pi)$  is not constructible. Coming up eventually: for which rational numbers r is  $\cos(r\pi)$  constructible?

# More field theory

If F is a field and  $p(x) \in F[x]$ , we saw we can find an extension of F with a root for p. If p is irreducible, the minimal such extension is unique. We now generalize this:

# More field theory

If F is a field and  $p(x) \in F[x]$ , we saw we can find an extension of F with a root for p. If p is irreducible, the minimal such extension is unique. We now generalize this:

## Definition

An extension K of a field F is called a *splitting field* for a polynomial  $p(x) \in F[x]$  if p(x) factors into linear factors in K[x] (we say that p(x) *splits completely in* K[x]), and does not split completely over any proper subfield of K containing F.

# More field theory

If F is a field and  $p(x) \in F[x]$ , we saw we can find an extension of F with a root for p. If p is irreducible, the minimal such extension is unique. We now generalize this:

## Definition

An extension K of a field F is called a *splitting field* for a polynomial  $p(x) \in F[x]$  if p(x) factors into linear factors in K[x] (we say that p(x) *splits completely in* K[x]), and does not split completely over any proper subfield of K containing F.

After adding enough roots, we see that every polynomial has a splitting field (13.4.25 in DF). We will also see that splitting fields are unique up to isomorphism, so we talk about "the" splitting field of a polynomial p(x) over F.

• The splitting field of  $x^2 - 2$  over  $\mathbb{Q}$  is just  $\mathbb{Q}(\sqrt{2})$ .

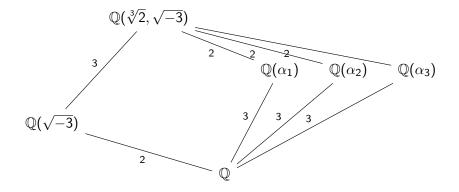
- The splitting field of  $x^2 2$  over  $\mathbb{Q}$  is just  $\mathbb{Q}(\sqrt{2})$ .
- The splitting field of  $(x^2 2)(x^2 3)$  is  $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ .

- The splitting field of  $x^2 2$  over  $\mathbb{Q}$  is just  $\mathbb{Q}(\sqrt{2})$ .
- The splitting field of  $(x^2 2)(x^2 3)$  is  $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ .
- The splitting field of x<sup>3</sup> − 2 is not Q(<sup>3</sup>√2): it is missing the complex roots! In fact the roots are α<sub>1</sub> = <sup>3</sup>√2, α<sub>2</sub> = <sup>3</sup>√2e<sup>2πi/3</sup>, α<sub>3</sub> = <sup>3</sup>√2e<sup>4πi/3</sup> and the splitting field K is the smallest containing all these roots.

- The splitting field of  $x^2 2$  over  $\mathbb{Q}$  is just  $\mathbb{Q}(\sqrt{2})$ .
- The splitting field of  $(x^2 2)(x^2 3)$  is  $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ .
- The splitting field of x<sup>3</sup> − 2 is not Q(<sup>3</sup>√2): it is missing the complex roots! In fact the roots are α<sub>1</sub> = <sup>3</sup>√2, α<sub>2</sub> = <sup>3</sup>√2e<sup>2πi/3</sup>, α<sub>3</sub> = <sup>3</sup>√2e<sup>4πi/3</sup> and the splitting field K is the smallest containing all these roots.
- We could describe it as K = Q(α<sub>1</sub>, α<sub>2</sub>, α<sub>3</sub>), but we have seen in a previous class that really K = Q(α<sub>1</sub>, α<sub>2</sub>), and even that description does not tell us the degree of K over F.

- The splitting field of  $x^2 2$  over  $\mathbb{Q}$  is just  $\mathbb{Q}(\sqrt{2})$ .
- The splitting field of  $(x^2 2)(x^2 3)$  is  $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ .
- The splitting field of x<sup>3</sup> − 2 is not Q(<sup>3</sup>√2): it is missing the complex roots! In fact the roots are α<sub>1</sub> = <sup>3</sup>√2, α<sub>2</sub> = <sup>3</sup>√2e<sup>2πi/3</sup>, α<sub>3</sub> = <sup>3</sup>√2e<sup>4πi/3</sup> and the splitting field K is the smallest containing all these roots.
- We could describe it as K = Q(α<sub>1</sub>, α<sub>2</sub>, α<sub>3</sub>), but we have seen in a previous class that really K = Q(α<sub>1</sub>, α<sub>2</sub>), and even that description does not tell us the degree of K over F.
- Instead, observe that α<sub>2</sub>/α<sub>1</sub> = e<sup>2πi/3</sup> ∈ K. In cartesian coordinates, it is -<sup>1</sup>/<sub>2</sub> + <sup>√3</sup>/<sub>2</sub>i, so √3i = √-3 ∈ K. This is a root of x<sup>2</sup> + 3, so has degree 2, hence in fact K = Q(<sup>3</sup>√2, √-3) has degree 6.

Picture of the splitting field of  $x^3 - 2$ , and known subfields:



Lemma

The splitting field of a polynomial p(x) of degree n has degree at most n!.

#### Lemma

The splitting field of a polynomial p(x) of degree n has degree at most n!.

### Proof.

Add a root  $\alpha$  to p. This is an extension of degree at most n. Dividing p(x) by  $(x - \alpha)$ , we get a polynomial of degree n - 1. Add a root to this one, and continue. At the end we have built an extension of degree at most  $n(n - 1) \dots 1 = n!$ .

#### Lemma

The splitting field of a polynomial p(x) of degree n has degree at most n!.

### Proof.

Add a root  $\alpha$  to p. This is an extension of degree at most n. Dividing p(x) by  $(x - \alpha)$ , we get a polynomial of degree n - 1. Add a root to this one, and continue. At the end we have built an extension of degree at most  $n(n - 1) \dots 1 = n!$ .

We could have figured out the degree d of the splitting field of  $x^3 - 2$  this way:

#### Lemma

The splitting field of a polynomial p(x) of degree n has degree at most n!.

## Proof.

Add a root  $\alpha$  to p. This is an extension of degree at most n. Dividing p(x) by  $(x - \alpha)$ , we get a polynomial of degree n - 1. Add a root to this one, and continue. At the end we have built an extension of degree at most  $n(n - 1) \dots 1 = n!$ .

We could have figured out the degree d of the splitting field of  $x^3 - 2$  this way: We know  $d \le 3! = 6$ , and must strictly contain  $\mathbb{Q}(\sqrt[3]{2})$  which has degree 3, so  $3 < d \le 6$ , and 3 divides d: d = 6 is the only possibility.

Splitting fields can be smaller than expected.

Example

Consider the polynomial  $p(x) = x^4 + 4$ .

Splitting fields can be smaller than expected.

#### Example

Consider the polynomial  $p(x) = x^4 + 4$ . It factors as  $(x^2 + 2x + 2)(x^2 - 2x + 2)$ . The roots are  $\pm 1 \pm i$ , so in fact the splitting field is  $\mathbb{Q}(i)$ , of degree 2!

We discussed two topics, straightedge and compass constructions, and splitting fields.

► The numbers that can be constructed using straightedge and compass are precisely those that can be obtained from Q by iterating quadratic extensions (that is, iterating taking square roots and closing under field operations). In particular they have degree a power of 2.

- ► The numbers that can be constructed using straightedge and compass are precisely those that can be obtained from Q by iterating quadratic extensions (that is, iterating taking square roots and closing under field operations). In particular they have degree a power of 2.
- This shows immediately that doubling the cube and squaring the circle is impossible.

- ► The numbers that can be constructed using straightedge and compass are precisely those that can be obtained from Q by iterating quadratic extensions (that is, iterating taking square roots and closing under field operations). In particular they have degree a power of 2.
- This shows immediately that doubling the cube and squaring the circle is impossible.
- ► Trisecting an angle θ is equivalent to constructing cos(θ/3). This is impossible for θ = π/3, because the triple angle formula shows cos(θ/3) must be the root of an irreducible polynomial of degree 3.

- ► The numbers that can be constructed using straightedge and compass are precisely those that can be obtained from Q by iterating quadratic extensions (that is, iterating taking square roots and closing under field operations). In particular they have degree a power of 2.
- This shows immediately that doubling the cube and squaring the circle is impossible.
- ► Trisecting an angle θ is equivalent to constructing cos(θ/3). This is impossible for θ = π/3, because the triple angle formula shows cos(θ/3) must be the root of an irreducible polynomial of degree 3.
- The splitting field of a polynomial (over a base) is the smallest field extension containing all the roots of that polynomial. If the polynomial has degree n, the splitting field has degree at most n!.