

Polynomials and their application to ruler and compass constructions

MA213 Second Year Essay

Matt Haddock 0503624

March 4, 2008

Contents

1	Introduction	2
2	General theorems of Polynomials	2
2.1	Degrees and roots	2
2.2	Division of polynomials	4
2.3	Irreducibility	4
2.4	Theorems on irreducibility	5
3	Field Extensions and their application to polynomials	6
3.1	Field Extensions	6
3.2	Simple extensions	6
3.3	Algebraic numbers	6
4	The Degree of an extension	6
4.1	Definition	7
4.2	Minimal polynomials	7
4.3	The Tower Law	8
5	Constructions with ruler and Compasses	9
5.1	The plane and constructible numbers	9
5.2	Fields and Constructible Numbers	10
6	Impossibility Proofs	11
6.1	Squaring the Circle	11
6.2	Trisecting the angle $\pi/3$	11
6.3	Trisecting an arbitrary angle	12
6.4	Doubling the cube	13
6.5	Constructing regular polygons	13

1 Introduction

Ruler-and-compass construction are the construction of lengths or angles using a perfected version of the ruler and compass.

The ruler to be used is assumed to be infinite in length, has no markings on it and only one edge, and is known as a straightedge. The compass is assumed to collapse when lifted from the page, so may not be directly used to transfer distances.

The most famous ruler-and-compass problems have been proven impossible in several cases by Pierre Wantzel (June 5, 1814 – May 21, 1848), using the mathematical theory of fields and polynomials, along with properties of irreducible polynomials. In spite of these impossibility proofs, some mathematical novices persist in trying to solve these problems. Many of them fail to understand that many of these problems are trivially solvable provided that other geometric transformations are allowed: for example, doubling the cube is possible using geometric constructions, but not possible using ruler and compass alone.

In this essay i will explore some general properties of polynomials, then move on to ruler and compass constructions, and apply some of the properties of polynomials to the Euclidean plane.

First I begin with the definition of a univariate polynomial, which is the type of polynomial which will be used from this point on:

Definition 1.1. A *polynomial* is an equation involving sums of powers in one or more variables, multiplied by coefficients. Possibly the most commonly used polynomial is the *univariate polynomial*:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} \cdots + a_2 x^2 + a_1 x + a_0 = 0$$

where a_i usually belongs to \mathbb{R} or \mathbb{C} , and x is the variable.

2 General theorems of Polynomials

Now it will be useful to recall some simple theorems about general polynomials. I shall also include some proofs that are rarely seen as they are slightly more advanced than the 'regular' proofs employed. Here are some theorems that have trivial proofs, as well as some definitions that will come in use.

2.1 Degrees and roots

Definition 2.1. If f is a polynomial over \mathbb{C} and $f \neq 0$ then the degree of f is the largest power of x with nonzero coefficient, and is denoted δf .

Theorem 2.2. If f, g are polynomials over \mathbb{C} then

$$\delta(f + g) \leq \max(\delta f, \delta g)$$

Proof. Let f, g denote the polynomials

$$f(x) = \sum_{i=1}^p a_i x^i \quad \text{and} \quad g(x) = \sum_{i=1}^q b_i x^i$$

then

$$\begin{aligned} f + g &= \sum_{i=1}^p a_i x^i + \sum_{i=1}^q b_i x^i = \\ &\sum_{i=1}^{\max(p,q)} (a_i + b_i) x^i. \end{aligned}$$

Which has degree $\leq \max(p, q) = \max(\delta f, \delta g)$, the non strict inequality comes from the possibility that $a_p = -b_q$ with $p = q$. \square

Proposition 2.3. *Two polynomials f, g define the same function if and only if they have the same coefficients (up to multiplication by a constant)*

Proof. If a_1, a_2, \dots, a_n are distinct complex numbers, then the *Vandermonde determinant* is defined as

$$D = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_m & a_m^2 & \cdots & a_m^{n-1} \end{vmatrix}$$

First we show that D is non zero,

Consider the a_j as independent indeterminates over \mathbb{C} . Then D is a polynomial in the a_j of total degree $0 + 1 + 2 + \dots + (n - 1) = \frac{1}{2}n(n - 1)$ Moreover D vanishes whenever $a_j = a_k$ for $k \neq j$ as it then has two non identical rows. Therefore D is divisible by $a_j - a_k$ for all $k \neq j$ hence it is divisible by

$$\prod_{j < k} (a_j - a_k)$$

Now compare degrees. (*Galois Theory* Ian Stewart, p28 Exercise 2.5 [8])

Suppose $f(x) = \sum_{j=0}^n c_j x^j$ and $g(x) = \sum_{j=0}^n d_j x^j$ be two elements in $\mathbb{C}[x]$ such that $f(x) = g(x)$ for all $x \in \mathbb{C}$. Note that by allowing the coefficients to be 0, we may assume that f and g have the same degrees. So $\sum_{j=0}^n (c_j - d_j)x^j = 0, \forall x \in \mathbb{C}$ now choose any distinct elements a_1, a_2, \dots, a_{n+1} in \mathbb{C} . then we have

$$\sum_{j=0}^n (c_j - d_j) a_k^j = 0$$

, $1 \leq k \leq n + 1$, which in terms of matrices becomes $A\underline{x} = \underline{0}$, where $\underline{0}$ is the zero column vector, \underline{x} is the column with entries $c_0 - d_0, \dots, c_n - d_n$, and A the $(n + 1) \times (n + 1)$ matrix with (i, j) entry a_i^{j-1} . so A is a Vandermonde matrix and thus its determinant is non-zero, and therefore it is invertible. thus from $A\underline{x} = \underline{0}$ we get $\underline{x} = \underline{0}$ and hence $c_j = d_j$ for all $0 \leq j \leq n$. \square

Theorem 2.4. *If a polynomial over \mathbb{C} has a root α then it is divisible by $(x - \alpha)$*

Theorem 2.5. *A polynomial has at least one root in \mathbb{C} .*

Proof. To prove this I will use an interesting argument from complex analysis, it uses some simple theorems already proven, and some which I will prove here, following closely the proof employed in 'Introduction to Complex Analysis' Rolf Nevanlinna [7]. \square

First we need:

Theorem 2.6 (Liouville's theorem). *Any bounded (ie $|f(x)| \leq M$ for some $M \in \mathbb{R}_+$) function holomorphic on the whole complex plane is a constant function.*

Proof. By the Taylor series on zero (which exists, given f is entire, i.e. holomorphic everywhere on the whole complex plane)

$$f(x) = \sum_{j=0}^{\infty} a_j z^j.$$

Which implies, from Cauchy's integral formula:

$$|a_j| = \left| \frac{1}{2\pi i} \oint_{C_r} \frac{f(z)}{z^{j+1}} dz \right| = \frac{1}{2\pi} \left| \oint_{C_r} \frac{f(x)}{z^{j+1}} dz \right| \leq \frac{1}{2\pi} \oint_{C_r} \frac{|f(x)|}{|z|^{j+1}}.$$

Where C_r is the boundary of a disc in the complex plane. Now we use the fact that f is bounded and that $|z| = r$, as it is on the circle.

$$\leq \frac{1}{2\pi} \oint_{C_r} \frac{M}{r^j} dz$$

Now integrate around the circle,

$$\frac{1}{2\pi} \frac{M}{r^{j+1}} 2\pi r = \frac{M}{r^j}$$

Now we let the radius of the circle tend to infinity so that:

$$\lim_{r \rightarrow \infty} \frac{M}{r^j} \rightarrow 0$$

i.e. all of the terms apart from a_0 in the Taylor series are zero. \square

Proof of theorem 2.5. Let $f : \mathbb{C} \rightarrow \mathbb{C}$ be a non-constant polynomial. Suppose f has no roots in \mathbb{C} , we aim to arrive at a contradiction. Define

$$g(x) = \frac{1}{f(x)}$$

Thus as $g(x)$ has no 'holes' $g(x)$ is holomorphic on \mathbb{C} However,

$$\text{As } f \rightarrow \infty \text{ } g \rightarrow 0 \text{ as } x \rightarrow \infty.$$

Then there is some M such that $|g(x)| \leq M$ so $g(x)$ is continuous and bounded on \mathbb{C} , thus by Liouville's theorem we have that $g(x)$ and therefore $f(x)$ are constant. Thus we arrive at the required contradiction and conclude that f has at least one root in \mathbb{C} . \square

Proposition 2.7. *A polynomial of degree n has at most n roots.*

Proof. By induction. The case $n = 1$ is trivial, assume true for $n = k$, name this polynomial $q(x)$, now let $p(x)$ be a polynomial of degree $k + 1$. If $p(x)$ has no roots then we are finished, but say it has at least 1 root, α . Then by Theorem 9 the polynomial is divisible by $(x - \alpha)$, thus there exists a polynomial $q(x)$ such that $p(x) = (x - \alpha)q(x)$. Then it is clear that this has at most $k + 1$ roots, so since the case $k \Rightarrow k + 1$ we are done, for all $n \geq 1$. \square

2.2 Division of polynomials

Definition 2.8. *Let f and g be polynomials over K . We say that f divides g (or f is a factor of g , or g is a multiple of f) if there exist some polynomial h over K such that $g = fh$. The notation $f|g$ will mean that f divides g , while $f \nmid g$ will mean f does not divide g .*

Definition 2.9. *A polynomial d over K is a highest common factor of f and g over K if $d|f$ and $d|g$ and furthermore, whenever $e|f$ and $e|g$ we have $e|d$.*

2.3 Irreducibility

Definition 2.10. *A polynomial is irreducible if it cannot be expressed as a product of 2 polynomials of lesser degree.*

Example 2.11. *It seems obvious that $x^2 - 3 = 0$ is irreducible over \mathbb{Q} as the solution is $x = \pm\sqrt{3}$ which $\notin \mathbb{Q}$, but how to we prove this? We aim for a contradiction. Write:*

$$x^2 - 3 = (ax + b)(cx + d),$$

it is obvious we can assume that $a = c = 1$ then $b + d = 0$ and $bd = -3$, hence $b^2 = 3$ but we no that no element in \mathbb{Q} which is equal to $\sqrt{3}$ hence we arrive at a contradiction and the result is proven. Notice that the polynomial is reducible over \mathbb{R} or \mathbb{C} , as $\sqrt{3} \in \mathbb{R}, \mathbb{C}$ But this does not mean that because a root belongs to a field, that the polynomial is reducible.

Example 2.12.

$$2x^2 + 2 = 0 \in \mathbb{R}.$$

The ‘reduction’ $\frac{1}{2}(x^2 + 1)$ is not valid as $\frac{1}{2}(x^2 + 1)$ has degree ≥ 2 . A proof of this being irreducible is simple.

2.4 Theorems on irreducibility

Lemma 2.13 (Gauss’ Lemma). *Let f be a polynomial over \mathbb{Z} that is irreducible over \mathbb{Z} . Then f considered as a polynomial over \mathbb{Q} is also irreducible over \mathbb{Q} .*

Proof. The proof of Gauss’ lemma is simple, but long winded. A proof can be found in *Galois Theory* by Stewart [8] (p39 Lemma 3.17). \square

Theorem 2.14. (*Eisenstein’s Criterion*) *If a polynomial satisfies the following conditions, with an arbitrary prime p*

1. $p \nmid a_n$
2. $p \mid a_i, (i = 0, 1, 2, \dots, n - 1)$
3. $p^2 \nmid a_0$

Then the polynomial $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0 = 0$ is irreducible over \mathbb{Q} .

Proof. Suppose for a contradiction $f = gh$. By Gauss’ Lemma (Lemma 2.13), it is enough to show that f is irreducible over \mathbb{Z} . Therefore we can assume that g, h have integer coefficients. Set

$$g = \sum_{i=0}^{\alpha} b_i x^i \quad h = \sum_{j=0}^{\beta} c_j x^j$$

Where $\beta + \alpha = n, b_i, c_j \in \mathbb{Z}$. Since $b_0 c_0 = a_0$, then either $p \mid b_0$ or $p \mid c_0$ (we cannot have both as $p^2 \nmid a_0$). So without loss of generality we assume $p \mid b_0$. If all of b_j are divisible by p , then we have that a_n is divisible by p , which contradicts the first assumption above. Let b_r be the first coefficient of g not divisible by p , then

$$a_r = b_r c_0 + \dots + b_0 c_r$$

where $r \leq n - 1$. This implies that p divides c_0 (as p divides a_r, b_0, \dots, b_{r-1} , but not b_r), but this contradicts the fact that we assumed $p \nmid c_0$, and thus f is irreducible. (Taken from Stewart [8].) \square

Example 2.15. *The polynomial*

$$f(x) = \frac{5}{6}x^2 + \frac{1}{3}x + \frac{1}{6}$$

can become

$$6f(x) = 5x^2 + 2x + 1$$

Now Eisenstein’s theorem can be applied with $p = 2$.

Now that we have a basic understanding of polynomials, we need to apply this to a more general case, so that we may apply our results to the Euclidean plane. This requires the notion of fields and field extensions.

3 Field Extensions and their application to polynomials

3.1 Field Extensions

I define field extensions in a similar way to D.J.H.Garling [3] in A Course in Galois Theory, except simplified, as the more general cases are not really applicable, as well as lifting some helpful explanations from Stewart [8].

Definition 3.1. Let L and K be any two fields. We say that L is an extension of K if K is a subset of L . We write

$$L : K$$

to denote that L is an extension of K . Note that this requires the two fields to have the same field operations.

The elements of L are the ‘vectors’ and the elements of K are the ‘scalars’. We multiply the elements of L by the elements of K to find all of the vectors in our vector space of the field extension.

Theorem 3.2. All subfields of \mathbb{R} contain \mathbb{Q}

Proof. Let K be the subfield. By definition $0, 1 \in K$ and therefore all $0, 1, 2, \dots, n$ are in K . Now since K also has multiplicative and additive inverses so all $-n$ and $n^{-1} \in K$ and so are all products of the two, therefore all $\frac{p}{q} \in K$ so we have all of $\mathbb{Q} \in K$. \square

Here also note that if K is a subfield of \mathbb{R} then K is an extension of \mathbb{Q} , ie $K : \mathbb{Q}$.

3.2 Simple extensions

Definition 3.3. A simple extension is a field extension such that we adjoin a single element to the subfield L . If α is the element we denote the extension

$$K(\alpha).$$

Example 3.4. \mathbb{R} with the element i ($= \mathbb{C}$) is a simple field extension.

3.3 Algebraic numbers

Definition 3.5. Let K be a subfield of \mathbb{C} and let $\alpha \in \mathbb{C}$ we say that α is algebraic if there exists a nonzero polynomial over K such that it has α as a root. Otherwise we say α is transcendental.

Theorem 3.6. The set of all algebraic numbers is countably infinite.

Proof. Let $\mathbb{Z}[x]$ be the set of all polynomials p with all coefficients $\in \mathbb{Z}$, and $\delta p \leq n$ then the map from $\mathbb{Z}[x]$ to \mathbb{Z}^{n+1} forms an injection. Therefore each $\mathbb{Z}[x]$ is countable, and so is the set of all polynomials with integer coefficients. Each polynomial has at most n roots so by a similar argument all of the algebraic numbers are countable. \square

Definition 3.7. A simple field extension is an **algebraic extension** if the single element adjoined to it is algebraic. Otherwise it is a **transcendental extension**.

4 The Degree of an extension

I use definitions and proofs from Swallow [9] to prove the important tower law, and its generalization to arbitrarily many extensions.

Theorem 4.1. If $L : K$ is a field extension then using the usual operations, then the vector space axioms hold.

Proof. This follows from the fact that the extension is a subfield of \mathbb{C} . \square

4.1 Definition

Definition 4.2. The degree of a field extension is the degree of the vector space, i.e. the number of elements in the basis, denoted

$$[L : K].$$

Example 4.3. The extension $\mathbb{Q}(i, \sqrt{3})$ has degree 4 as the elements $(1, \sqrt{3}, i, i\sqrt{3})$ span the vector space.

Theorem 4.4. All finite extensions are algebraic extensions.

Proof. A finite field means the degree of the associated vector space is finite, and thus there are a finite number of elements that span the vector space associated with that extension. Let $L \supseteq K$ be the field extension. We are trying to show that for every $\alpha \in K$ there exists a non zero polynomial $p(x)$ such that $p(\alpha) = 0$. Let n denote the dimension of the extension. Consider the following set in L :

$$S = \{1, \alpha, \alpha^2, \dots, \alpha^n\}$$

there are $n + 1$ vectors in this set, if we consider L as a vector space they must be linearly dependent. Thus there exist k_0, k_1, \dots, k_n , not all zero, such that

$$k_0 + k_1\alpha + k_2\alpha^2 + \dots + k_n\alpha^n = 0.$$

Now we simply define

$$p(x) = k_0 + k_1x + k_2x^2 + \dots + k_nx^n = 0$$

and therefore $p(\alpha) = 0$ as desired. \square

Definition 4.5. A polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

is monic over $K \subseteq \mathbb{C}$ if $a_n = 1$ [9].

4.2 Minimal polynomials

Definition 4.6. The minimal polynomial over a field K and an element $\alpha \in K$ is the unique monic polynomial, m , of minimal degree such that

$$m(\alpha) = 0.$$

Example 4.7. As in example 2.13

$$x^2 - 3 = 0$$

over \mathbb{Q} this time is the minimal polynomial of $\alpha = \sqrt{3}$, as it satisfies the required properties.

Theorem 4.8. The minimal polynomial of an element $\alpha \in K \subseteq \mathbb{C}$ then the minimal polynomial of α is irreducible over K . It also divides every polynomial of which α is a zero.

Here I use some ideas from the Galois Theory paper by Artin [1] on division of polynomials, and a similar contradiction argument from the section on polynomials (page 78).

Proof. Suppose for a contradiction $m = fg$. Thus as $m(\alpha) = 0$ then either $g(\alpha) = 0$ or $f(\alpha) = 0$. But $\delta g, \delta f \leq \delta m$, but this contradicts the definition of m , proving that m is of least degree. Now suppose that p is an arbitrary polynomial such that $p(\alpha) = 0$ by the division algorithm there exist polynomials q, r such that

$$p = mq + r$$

and $\delta r < \delta m$ Now $p(\alpha) = 0 + r(\alpha) = 0$. But again we have that $r(\alpha) = 0$ contradicting the definition, therefore $p = qm$ and the second result is proven. \square

Lemma 4.9. *A polynomial, $p(x)$ over \mathbb{C} cannot be the minimal polynomial of more than element $\in \mathbb{C}$.*

Proof. Suppose α, β are the two roots, then by Theorem 8 $p(x)$ is divisible by $(x-\alpha)(x-\beta)$, with $(x-\alpha) = 0, (x-\beta) = 0$ this shows that obviously the two minimal polynomials are $x-\alpha = 0$ and $x-\beta = 0$. \square

4.3 The Tower Law

I use the proof employed by Lang [5] taking parts from the section on algebraic extensions from Stewart [8], changing parts to in keep with my notation.

Theorem 4.10. *The degrees of finite extensions are multiplicative. i.e.*

$$[E : K] = [E : F][F : K].$$

Proof. Suppose F has degree n and E has degree m . Let $(e_i)_{1 \leq i \leq n}$ be a basis for E and $(f_j)_{1 \leq j \leq m}$ be a basis for F . We show that $(e_i f_j)_{1 \leq j \leq m, 1 \leq i \leq n}$ is a basis for L over F . Every element in E can be expressed as follows:

$$z = \sum_{i=1}^n e_i l_i \quad \text{for some } l_i \in F$$

and every element in F as

$$\sum_{j=1}^m f_j k_j \quad \text{for some } k_j \in K.$$

Combining these two gives:

$$z = \sum_{j=1}^m e_i \sum_{i=1}^n f_j \alpha_{ij} = \sum_{j=1}^m \sum_{i=1}^n e_i f_j \alpha_{ij}$$

for some $\alpha \in K$. This shows that $(e_i f_j)$ form a basis for E over K , which is clearly linearly independent as $(e_i), (f_j)$ are, and hence the basis of $E : K$ has $mn = [E : F][F : K]$ elements, and the theorem is proved. \square

Corollary 4.11. *If $K_0 \subseteq K_1 \subseteq K_2 \subseteq \dots \subseteq K_n$ are subsets of \mathbb{C} then*

$$[K_n : K_0] = [K_n : K_{n-1}][K_{n-1} : K_{n-2}] \cdots [K_1 : K_0].$$

Proof. This follows trivially from theorem 5.10 by induction. \square

Lemma 4.12. *Let $K(\alpha)$ be a simple field extension. Then $\delta m = [K(\alpha); K]$ where m is the minimal polynomial of the element α .*

Proof. It is simple as the elements $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ form a basis where $n = \delta m$ [8]. \square

Theorem 4.13. *Let α, β be two elements of an extension field of a given field K . Both these elements are algebraic over K if and only if both $\alpha + \beta$ and $\alpha\beta$ are algebraic over K .*

Proof. Assume first that α and β are algebraic. because

$$[K(\alpha, \beta) : K] = [K(\alpha, \beta) : K(\alpha)] [K(\alpha) : K]$$

and both factors here are finite, then $[K(\alpha, \beta) : K]$ is finite. So we have a finite field extension $K(\alpha, \beta) : K$ which thus is also algebraic, and therefore the elements $\alpha + \beta$ and $\alpha\beta$ of $K(\alpha, \beta)$

are algebraic over K . Secondly suppose that $\alpha+\beta$ and $\alpha\beta$ are algebraic over K . The elements α and β are the roots of the quadratic equation $x^2 - (\alpha+\beta)x + \alpha\beta = 0$ with the coefficients in $K(\alpha+\beta, \alpha\beta)$. Thus

$$[K(\alpha, \beta) : K] = [K(\alpha, \beta) : K(\alpha+\beta, \alpha\beta)] [K(\alpha+\beta, \alpha\beta) : K] \leq 2[K(\alpha+\beta, \alpha\beta) : K].$$

Since $[K(\alpha+\beta, \alpha\beta) : K]$ is finite, then also $[K(\alpha, \beta) : K]$ is, and in the finite extension $K(\alpha, \beta) : K$ the elements α and β must be algebraic over K (Stewart [8]). \square

5 Constructions with ruler and Compasses

Constructions with ruler and compasses have been around for years, in its early days it was simply done by eye. For proper impossibility proofs, it is simply not accurate enough, the modern approach is done by field theory, in a much more abstract way.

5.1 The plane and constructible numbers

We carry out ruler and compass constructions in the plane i.e.

$$F \subset \mathbb{R} \times \mathbb{R}$$

We use a more perfected version of the ruler and compasses, definitions taken from Stewart [8] and Dickson [2].

Definition 5.1.

1. **The compass** can be opened arbitrarily wide, but (unlike most real compasses) it has no markings on it. It can only be opened to widths that have already been constructed.
2. **The ruler** is infinitely long, but it has no markings on it and has only one edge, unlike ordinary rulers. It can only be used to draw a line segment between two points or to extend an existing line.

With these we can make construct 2 different shapes in the plane

Definition 5.2.

1. **The line** given by equation

$$L = ax + by + c = 0 \quad a, b, c \in F$$

2. **The circle** given by equation

$$C = (x - a)^2 + (y - b)^2 = c^2 \quad a, b, c \in F.$$

Lemma 5.3. With the above definitions of circles and lines

1. $L \cap L = \emptyset$ or consists of 1 point in F
2. $L \cap C = \emptyset$ or consists of 1 or 2 points in F
3. $C \cap C = \emptyset$ or consists of 1 or 2 points in F .

Proof. These equations can be easily solved, and it is obvious that we will be solving an equation of either degree 1 or 2. \square

Definition 5.4. We say that a point is constructible if it is the intersection of lines and circle, or it is the point 0 or 1.

Definition 5.5. If α, β are given points then the line from α to β , and the circle, centre α , radius $|\beta|$ are denoted

$$L(\alpha, \beta) \quad C(\alpha, \beta).$$

To start with we simply have our blank 'page' F , with two numbers 0 and 1 marked on it (We simply choose any arbitrarily long line to count as our 'unit'). To begin we can construct the x axis easily by extending our unit line, then we can draw a perpendicular line to the x axis, using the well known theorem. We also have all of the natural numbers, and the intersection of $C(0, 1)$ with the x axis intersect at -1, now we can construct all of \mathbb{Z} . Now the circles $C(0, 1)$ and $C(1, 1)$ intersect at $2 \pm \frac{\sqrt{3}}{2}i$, thus it is advantageous for us to work in the complex plane [8]. If we write a point p as

$$p = x + iy$$

then it simplifies notation, again it is simple to show that all of $[\mathbb{Q}(i) : \mathbb{Q}]$ is constructible.

Lemma 5.6. If $p, q \in \mathbb{R}$ are constructible, then so are

$$p + q \quad p - q \quad pq \quad p/q \quad \sqrt{q}$$

Proof. The proof of these are simple and belong more in the field of geometry than in the study of impossibility proofs. But however it is useful as it proves that the set of all constructible numbers forms a field. \square

5.2 Fields and Constructible Numbers

Theorem 5.7. The set of all constructible numbers forms a field.

Proof. Follows directly from Lemma 5.6. \square

When we are adding new constructible numbers to a current set of numbers, we are essentially adjoining these new elements to the base field. These adjunctions form a tower of subfields, if we define

$$F_j = F_{j-1}(x_j, y_j)$$

where here we are adjoining the elements x_j, y_j to our original field to create a new one. We then have

$$F_0 \subseteq F_1 \subseteq F_2 \subseteq \dots \subseteq F_n \subseteq \mathbb{R}$$

and the tower law can now be applied [8].

Theorem 5.8. Let F be a subset $\subset \mathbb{R}$ of points. Then the simple field extension brought about by ruler and compass constructions to find a new element α , then adjoining it to F , has the property

$$[F(\alpha) : F] = 1 \text{ or } 2.$$

Proof. The element α will be the root of a possibly reducible quadratic or linear equation (see Lemma 5.3). Thus its minimal polynomial will be a polynomial of degree either 1 or 2. Then by Lemma 4.12 we are done [9]. \square

Thus by the tower law we have

$$[F_n : F_0] = [F_n : F_{n-1}][F_{n-1} : F_{n-2}] \dots [F_1 : F_0] = 2^{n-p}$$

i.e. any set of n constructible numbers has degree 2^n . Where p is the number of extensions with degree 1 [9].

6 Impossibility Proofs

6.1 Squaring the Circle

Now we are in a position where we have the tools to prove that squaring the circle is in fact impossible. The problem concerns constructing a square the same size as a given circle. We have the radius of the circle then by Lemma 5.6 we can compute r^2 . Thus this problem is equivalent to:

Theorem 6.1. *It is impossible to construct the number $\sqrt{\pi}$ [2].*

Proof. Since π is transcendental (proving this is very difficult I will not attempt it here, a full proof can be found in Stewart [8].) so is $\sqrt{\pi}$ this follows from Theorem 4.13; if we assume $\sqrt{\pi}$ to be algebraic, then its square must be algebraic, since π is not algebraic we arrive at a contradiction as π is transcendental, and therefore $\sqrt{\pi}$ is. Therefore it has no minimal polynomial and by Theorem 5.8 and by Theorem 4.12 the degree of the field extension created by adjoining $\sqrt{\pi}$ is ∞ , which is not a power of 2. \square

6.2 Trisecting the angle $\pi/3$

Next comes another well known impossibility proof: trisecting the angle $\frac{\pi}{3}$. Again it is impossible, but uses a slightly different approach to the last theorem. First the following lemma is needed [2]:

Lemma 6.2.

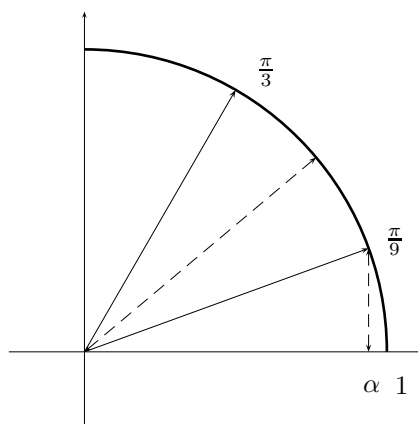
$$\cos(3\theta) = 4 \cos^3 \theta - 3 \cos \theta$$

Proof.

$$\begin{aligned} \cos 3\theta &= \cos(2\theta + \theta) \\ &= \cos 2\theta \cos \theta - \sin 2\theta \sin \theta \\ &= \cos \theta (2 \cos^2 \theta - 1) - 2 \sin^2 \theta \cos \theta \\ &= 2 \cos^3 \theta - \cos \theta - 2 \cos \theta (1 - \cos^2 \theta) \\ &= 4 \cos^3 \theta - 3 \cos \theta. \end{aligned}$$

\square

Theorem 6.3. *It is impossible to trisect the angle $\frac{\pi}{3}$ using ruler and compass constructions.*



Proof. This is equivalent to starting from $(0,0)$ and $(1,0)$ and constructing a number α such that $\alpha = \cos\left(\frac{\pi}{9}\right)$ then we can obviously from Lemma 5.6 construct $\beta = 2\cos\left(\frac{\pi}{9}\right)$. Using Lemma 6.2 we have

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

now put $\theta = \frac{\pi}{9}$ and $\cos 3\theta = \frac{1}{2}$ and we get

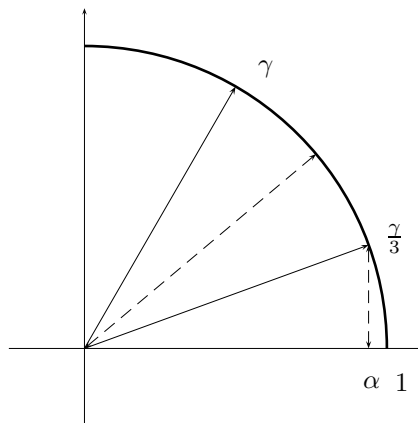
$$\beta^3 - 3\beta - 1 = 0.$$

Now this polynomial is irreducible (Stewart [8]), and its degree of the field extension

$$[\mathbb{Q}(\beta) : \mathbb{Q}] = 3$$

so by Theorem 5.8, page 10 we are done. □

6.3 Trisecting an arbitrary angle



We now show the criteria for being able to trisect an arbitrary angle. Similarly to trisecting the angle $\frac{\pi}{3}$, this is equivalent to constructing the length α . We can construct multiples of $\frac{\pi}{2}$ and add them to any other angles easily, thus we only need to prove this for angles γ , $0 \leq \gamma \leq \frac{\pi}{2}$. Similarly to trisecting the angle $\frac{\pi}{3}$, $\alpha = \cos\left(\frac{\gamma}{3}\right)$, and we have that

$$\cos(\gamma) = 4\cos^3\left(\frac{\gamma}{3}\right) - 3\cos\left(\frac{\gamma}{3}\right)$$

Using this and defining $\beta = 2\cos(\gamma)$, we have that we can only construct α if

$$4\beta^3 - 3\beta - \cos\gamma$$

is reducible over $\mathbb{Q}(\cos\gamma)$. Thus we have proved the following theorem:

Theorem 6.4. *The angle θ can only be trisected if*

$$x^3 - 3x - \cos\theta$$

is reducible over $\mathbb{Q}(\cos\theta)$.

6.4 Doubling the cube

Theorem 6.5. *If given any cube, it is impossible to create a cube twice its volume.*

Proof. By a similar argument to that used in squaring the circle, it is equivalent problem to constructing the number [8]

$$\sqrt[3]{2}$$

the minimal polynomial of this is evidently

$$x^3 - 2 = 0.$$

By Eisenstein this is irreducible, and with degree 3

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

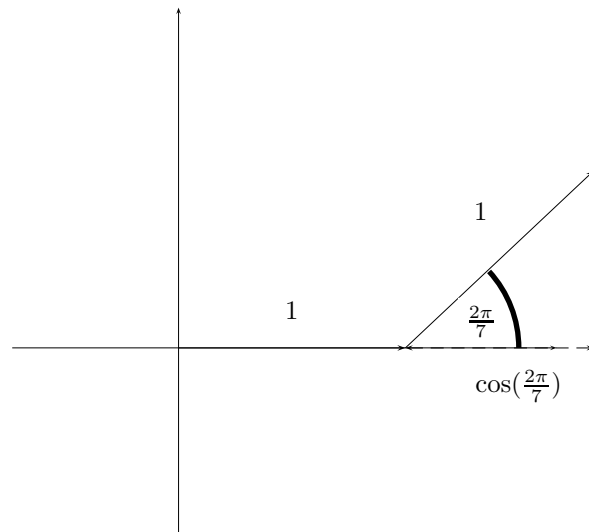
we are done, since this degree is $\neq 1, 2$ and thus the number $\sqrt[3]{2}$ cannot be constructed by ruler and compass. \square

6.5 Constructing regular polygons

Now we move onto a another impossibility proof, this time concerning the construction of a 7 sided polygon.

Theorem 6.6. *It is impossible to construct a regular heptagon using ruler and compass constructions.*

Again this problem, like trisecting the angle, can be done using a marked ruler using a *Neusis construction*[4], however is impossible without, as i shall now show.



Proof. The exterior angle of a heptagon is clearly $\frac{2\pi}{7}$ (see figure), furthermore using this fact that constructing a regular heptagon is equivalent to solving the polynomial

$$f(t) = t^6 + t^5 + t^4 + t^3 + t^2 + t + 1$$

(Stewart [8], page 83, Exercise 83) Which has roots $\cos\left(\frac{2\pi}{7}\right) + i \sin\left(\frac{2\pi}{7}\right)$, then

$$f(t) = t^6 + t^5 + t^4 + t^3 + t^2 + t + 1 = t^3 + t^2 + t^1 + 1 + t^{-1} + t^{-2} + t^{-3} = 0$$

now we substitute $x = t + \frac{1}{t}$ to find that

$$f(t) = x^3 + x^2 - 2x - 1$$

Suppose that $f(x)$ is reducible, then

$$f(x) = (ax + b)(cx^2 + dx + e) = acx^3 + (dc + bc)x^2 + (ea + bd)x + bc$$

Then $ac = 1$, $(dc + cb) = 1$, $(ea + bd) = -2$ and $bc = -1$. These equations have no solution, thus f is irreducible, and we cannot construct the number $\cos\left(\frac{2\pi}{7}\right) + i \sin\left(\frac{2\pi}{7}\right)$, and thus we cannot construct a regular heptagon. \square

Theorem 6.7. *It is impossible to construct a regular nonagon using ruler and compass constructions*

Proof. The interior angle of a nonagon is $\pi/9$, which we have already shown is impossible to construct. \square

In fact Carl Friedrich Gauss in 1796 showed that a regular n -sided polygon can be constructed with ruler and compass if the odd prime factors of n are distinct Fermat primes, positive integers of the form

$$2^{2^n} + 1.$$

(Michal Křížek [6])

Example 6.8. *Constructible regular polygons*

- $9 = 3 \times 3$ so although these are Fermat primes, they are not distinct so a nonagon is non constructible.
- $7 \neq (2^2)^n + 1$ for any $n \in \mathbb{N}$ which confirms a heptagon is non constructible.
- $5 = 2^2 + 1$, so a pentagon is constructible. I will not prove it here as it lies more in the realm of geometry.
- $3 = 2 + 1$, 2 is a Fermat prime, and clearly an equilateral triangle is constructible.

Furthermore so is any polygon with $3a$ or $5a$ sides, where a is any even number, or 1.

Gauss conjectured that this condition was also necessary, but he offered no proof of this fact, which was proven by Pierre Wantzel in 1837 (Michal Křížek [6] *17 Lectures on Fermat Numbers: From Number Theory to Geometry*).

Now we have proved the impossibility of many problems posed by ancient geometers, using interesting results from algebra, rather than purely the field of geometry. However, as stated at the start of the essay, still to this day many mathematicians still continue to try to find geometric solutions despite the proofs outlined here. For many of the problems solutions have thought to be found, but are in fact just approximations to the problems, of which there are many [8]. The results included here on the seemingly unrelated field of field extensions and irreducibility have had far reaching applications into Galois theory, abstract algebra and geometry, to name a few.

References

1. Emil Artin. *Exposition: by Emil Artin: A selection*. American Mathematics Society, first edition, 2007.
2. Leonard Eugene Dickson. *New First Course in the Theory of Equations*. John Wiley and Sons, Inc, fifteenth edition, 1964.
3. D.J.H.Garling. *A Course in Galois Theory*. Cambridge University Press, second edition, 1988.
4. T. L. Heath. *A history of Greek Mathematics*. Oxford, 1921.
5. Serge Lang. *Algebra*. Springer, third edition, 2002.
6. Lawrence Somer Michal Křížek, Florian Luca. *17 Lectures on Fermat Numbers: From Number Theory to Geometry*. Springer, 2001.
7. V.Paatero Rolf Nevanlinna. *Introduction to Complex Analysis*. Birkhauser Verlag Basel, first edition, 1964.
8. Ian Stewart. *Galois Theory*. Chapman and Hall/CRC, third edition, 2003.
9. John Swallow. *Exploratory Galois Theory*. Cambridge University Press, first edition, 2004.