Galois Groups of Truncated Exponentials

James Bang

October 10, 2023

Abstract

This mathematical essay explores the Galois groups of *truncated exponential* polynomials over the field of rational numbers. We present a proof classifying the Galois groups of such polynomials using only material accessible to a student who has taken Part II Mathematics.

1 Introduction

A classic corollary of Artin's theorem in Galois Theory states that given a field k, *the general polynomial of degree n over k has Galois group S_n.*

Stated precisely, let $L = k(X_1, \dots, X_n)$ for indeterminates X_i, and $K = k(e_1, \dots, e_n)$ where e_j is the j–th *elementary symmetric polynomial* over the X_i. Then, L/K is a splitting field for the polynomial

$$f(T) = T^n - e_1 T^{n-1} + \dots + (-1)^n e_n,$$

and hence $\mathrm{Gal}(f/K) = \mathrm{Gal}(L/K) \cong S_n$.

For specific fields K satisfying the conditions of Hilbert's irreducibility theorem, this result can be strengthened to say that *almost all* polynomials of degree n over k have Galois group S_n, in the sense that such polynomials form a Zariski open subset of K^{n+1}.

However, not all specific classes of irreducible polynomials have a majority of their members with Galois group S_n. For example, polynomials of the form $T^n - a$ for $a \in k$ have Galois group contained in the product $\mathbb{Z}/n\mathbb{Z} \rtimes (\mathbb{Z}/n\mathbb{Z})^\times < S_n$.

Hence, it is of interest to determine whether the polynomials of a certain form have Galois group S_n. In this essay, we explore the Galois groups of *truncated exponential* polynomials, that is, those of the form

$$f_n(T) = \frac{T^n}{n!} + \frac{T^{n-1}}{(n-1)!} + \dots + T + 1 \in \mathbb{Q}[T].$$

Theorem 1.1. The truncated exponential polynomial f_n has Galois group A_n over $\mathbb{Q}$ if and only if $4 \mid n$, and S_n otherwise.

The method of proof is roughly as follows. Firstly, we consider the factorisation of f_n over the p–adic number field $\mathbb{Q}_p$, to show that it is irreducible and that $\mathrm{Gal}(f_n/\mathbb{Q})$ contains a large p–cycle. Using Jordan's Symmetric Group theorem, we deduce that $\mathrm{Gal}(f_n/\mathbb{Q})$ contains A_n. Finally, a computation of the discriminant will determine the Galois group completely.

2 Valued fields and p–adic numbers

To show the polynomial f_n has the aforementioned Galois groups, we need to show at the very least that f_n is irreducible. Most of the commonly known irreducibility criterions over $\mathbb{Q}$, such as factoring over $\mathbb{F}_p$, Eisenstein or Cohn's criterion depend on properties requiring prime numbers, and don't directly work here.

However, this motivates us to consider the factorisation of f_n over other number–theoretic fields, such as the *p–adic numbers* $\mathbb{Q}_p$. We will see that this technique is far more powerful than simply testing for irreducibility over a finite field $\mathbb{F}_p$, which is hopeless due to the existence of $n!$ in the denominator.

2.1 p–adic numbers

Consider the rational numbers $\mathbb{Q}$. Recall $\mathbb{Q}$ is a metric space with the usual absolute value metric $d(x, y) = |x - y|$. With this metric, $\mathbb{Q}$ is not complete; completing it with this metric gives the real numbers $\mathbb{R}$.

However, there is another number–theoretic metric on $\mathbb{Q}$. Fix a prime p. For any $q \in \mathbb{Q}^\times$, there exists a number n called the *valuation* of q, such that

$$q = p^n \cdot \frac{a}{b}$$

such that $p \nmid a, b$. We define the *p–adic norm* and the *p–adic valuation* of q to respectively be

$$|q|_p = p^{-n}, \quad \nu_p(q) = n,$$

where we also introduce the conventions $|0|_p = 0$ and $\nu_p(0) = \infty$.

It is not difficult to see that $(\mathbb{Q}, |\cdot|_p)$ is a metric space which again is not complete. Hence, denote $\mathbb{Q}_p$ the completion of $\mathbb{Q}$ with respect to $|\cdot|_p$.

This metric has the following properties, which can directly be verified from the definitions:

- $|a + b|_p \leq \max\{|a|_p, |b|_p\}$, that is, $\nu_p(a + b) \leq \max\{\nu_p(a), \nu_p(b)\}$;
- $|ab|_p = |a|_p \cdot |b|_p$, that is, $\nu_p(ab) = \nu_p(a) + \nu_p(b)$;
- $|a|_p = 0 \iff \nu_p(a) = \infty \iff a = 0$.

Most of these properties are similar with that of the usual real absolute value. However, the first property makes $\mathbb{Q}_p$ into an *ultrametric space*, one which satisfies the strong triangle inequality. The following definition makes this more precise.

Definition 2.1. A *valued field* is a pair (K, ν) consisting of a field K equipped with a function $\nu : K \to \mathbb{R} \cup \{\infty\}$ such that the above three properties hold. In addition,

- The function ν is called a *valuation* on K.
- The metric $|\cdot|_\nu$ defined by $|a|_\nu = r^{-\nu(a)}$ for some $r > 0$ is called an *absolute value* on K.

Notice that specifying a valuation ν on a field K is equivalent to specifying a metric $|\cdot|_\nu$ on K by

$$|a|_\nu = r^{-\nu(a)}, \quad \nu(a) = -\log_r |a|_\nu$$

for any $r > 0$. Hence, we are free to interchange between the two notions.

For any field K, there is a *trivial absolute value* $|\cdot|_0$ given by

$$|a|_0 = \begin{cases} 1 & a \neq 0 \\ 0 & a = 0, \end{cases}$$

which makes K into a valued field. As this valuation is not very interesting and break some of the properties to be discussed, we will consider valuations to be nontrivial for the rest of this essay.

Our first result regarding valued fields are that absolute values on K which are equivalent in a topological sense are essentially the same.

Theorem 2.2. Let K be a valued field with two absolute values $|\cdot|, |\cdot|'$. Then, $|\cdot|$ and $|\cdot|'$ induce the same topology on K if and only if there exists $r > 0$ such that $|x|' = |x|^r$ for all $x \in K$.

Simply stated, equivalent absolute values are the same up to a change of base. In particular, the equivalent valuations ν, ν' corresponding to $|\cdot|$ and $|\cdot|'$ are the same up to a constant factor.

Proof. The reverse direction is trivial, so consider the forward direction. Suppose $|\cdot|$ and $|\cdot|'$ induce the same topology on K. Then, there is $C > 0$ such that $|x|' \leq C|x|$ for any $x \in K$.

If $|x| < 1$, then $C^{-1} \cdot |x^n|' \leq |x^n| = |x|^n < 1$. Hence, $|x^n|' < C$, or $|x|' < C^{1/n}$ for all $n \geq 1$, and hence by taking the limit $n \to \infty$ we have $|x|' < 1$ as well.

It suffices to show that the quantity $\frac{\log|x|}{\log|x|'}$ is independent of $x \neq 0$. Suppose not; then, we can find $x, y \in K$ such that

$$\frac{\log|x|}{\log|x|'} < \frac{\log|y|}{\log|y|'}.$$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $n, m \in \mathbb{N}$ such that

$$\frac{\log|x|}{\log|y|} < \frac{n}{m} < \frac{\log|x|'}{\log|y|'} \implies \left|\frac{x^n}{y^m}\right| < 1 < \left|\frac{x^n}{y^m}\right|'.$$

This contradicts the earlier observation that $|x| < 1 \implies |x|' < 1$. □

2.2 Hensel's Lemma

So far, we haven't used the fact that the absolute value $|\cdot|$ turns K into an ultrametric space. One interesting fact from this property is that the set

$$\mathcal{O}_K \overset{\text{def}}{=} \{x \in K : |x| \leq 1\}$$

forms a *ring*. Indeed, for any $a, b \in \mathcal{O}_K$, we have $a + b, ab \in \mathcal{O}_K$ for any $a, b \in \mathcal{O}_K$, since

$$|a + b| \leq \max\{|a|, |b|\} \leq 1 \quad \text{and} \quad |ab| = |a| \cdot |b| \leq 1.$$

Note that this is highly wrong for the usual absolute value on $\mathbb{Q}$, where $|1 + 1| = |2| > 1$.

In addition, there is a unique maximal ideal $\mathfrak{m}$ of $\mathcal{O}_K$, given by

$$\mathfrak{m} = \{x \in K : |x| < 1\},$$

giving rise to a *residual field* $k = k_K \overset{\text{def}}{=} \mathcal{O}_K/\mathfrak{m}_K$.

The ring $\mathcal{O}_K$ is called the *ring of integers* of K, and plays a similar role to the ring $\mathbb{Z}$ of integers in $\mathbb{Q}$. In particular, the notions of a polynomial being *primitive* can be generalised to polynomials over $\mathcal{O}_K$.

Definition 2.3. Let (K, ν) be a valued field. A polynomial $f(T) = a_n T^n + \cdots + a_0 \in K[T]$ is *primitive* if $|a_i| \le 1$ for all i, and there is some j with $|a_j| = 1$.

This implies that primitive polynomials are in $\mathcal{O}_K[T]$.

In Part II Number Theory, we learnt that, under certain conditions, we can lift the roots x of an integer polynomial f modulo p^n to roots modulo p^{n+1}. This is a key reason to study the p–adic numbers $\mathbb{Q}_p$ – while it is impossible to make sense of the *classic* Euclidean limit of a sequence of potential unbounded integers, we can make sense of it in the alternative topology of $\mathbb{Q}_p$.

Hence, we can lift the root x all the way to a root of f in $\mathbb{Q}_p$ by taking said limit. This makes such arithmetic more suitable to study in $\mathbb{Q}_p$ where we don't need to worry about the presence of the integer n in p^n. The following theorem is a generalisation of this result to valued fields, with the same proof given in the Part II Number Theory lectures.

Theorem 2.4 [Hensel's Lemma]. Let K be a complete valued field, and $k = \mathcal{O}_K/\mathfrak{m}_K$. Suppose $f \in \mathcal{O}_K[T]$ and there is $\overline{a} \in k$ such that the reductions $\overline{f}$ of f in k satisfy

$$\overline{f}(\overline{a}) = 0 \quad \text{but} \quad \overline{f'}(\overline{a}) \neq 0.$$

Then, there exists a unique lift $a \in \mathcal{O}_K$ of $\overline{a}$ such that $f(a) = 0$.

In our use case, we are interested in lifting not just roots, but entire *polynomials* from $k[T]$ to $\mathcal{O}_K[T]$. The above theorem can be seen as lifting the factor $T - a$ of f from $k[T]$ to $\mathcal{O}_K[T]$. The following result is a generalisation of this, allowing lifts of entire factorisations of polynomials.

Theorem 2.4′ [Hensel's Lemma, polynomial version]. Let K be a complete valued field, and $f(T) \in K[T]$ primitive. Suppose $\overline{f} \equiv f \pmod{\mathfrak{m}_K} \in k[T]$ has a factorisation $\overline{f} = \overline{g} \cdot \overline{h}$ with $\overline{g}, \overline{h}$ coprime. Then, there exists $g, h \in \mathcal{O}_K[T]$ such that $\deg g = \deg \overline{g}$ and

$$f = gh, \quad \overline{g} = g \pmod{\mathfrak{m}_K}, \quad \overline{h} = h \pmod{\mathfrak{m}_K}.$$

Notice that we *do not* guarantee that $\deg h = \deg \overline{h}$. This is because $f, \overline{f}$ don't necessary have the same degree, so we can only guarantee one of g, h to have the same degree as $\overline{g}, \overline{h}$.

As we will be considering the factorisation of f_n over $\mathbb{Q}_p$ in the hope of showing it is irreducible, such a result is very useful to prove results regarding the factorisation of polynomials over $\mathbb{Q}_p$.

Proof. We know that $\overline{f} \equiv f \pmod{\mathfrak{m}_K}$. There are finitely many coefficients in $f - \overline{f}$, which have absolute value less than 1. Hence, we can replace $\mathfrak{m}_K$ with some $\pi_K \in \mathfrak{m}_K$ with $|\pi_K|_K$ large enough (i.e. close to 1) such that $f \equiv \overline{f} \pmod{\pi_K}$.

It is enough to show that there is a factorisation $f = g_n h_n \pmod{\pi_K^n}$ for each n such that $g_{n+1} \equiv g_n$ and $h_{n+1} \equiv h_n \pmod{\pi_K^n}$. Indeed, this would imply that $g = \lim g_n$ and $h = \lim h_n$ are the desired polynomials.

We construct the factorisation by induction on n. For $n = 1$, we may just take $g_1 = g$ and $h_1 = h$, which trivially work. Suppose we have found $g_n, h_n \in \mathcal{O}_K[T]$ with $g_n h_n \equiv f \pmod{\pi_K^n}$. Let

$$g_{n+1} = g_n + \pi_K^n \cdot a \quad \text{and} \quad h_{n+1} = h_n + \pi_K^n \cdot b$$

for some $a, b \in k[T]$ to be chosen later. Then, $g_{n+1} h_{n+1} \equiv f \pmod{\pi_K^{n+1}}$ is equivalent to

$$\overline{a} \cdot \overline{h} + \overline{b} \cdot \overline{g} = c \overset{\text{def}}{=} \pi_K^{-n}(f - g_n h_n) \tag{$\star$}$$

where $\overline{a}, \overline{b}$ are the reductions of a, b modulo π_K. If we denote P_n the k–vector space of polynomials whose degrees are at most n, then the map

$$\varphi : P_{\deg g} \times P_{\deg h} \to P_{\deg g + \deg h}, \quad \varphi(a, b) = a \cdot \overline{h} + b \cdot \overline{g}$$

has kernel spanned by $(-\overline{g}, \overline{h})$ since $\overline{g}, \overline{h}$ are coprime, and hence has nullity $\deg g \deg h$. It follows that

$$\operatorname{rank}(\varphi) = (\deg g + 1)(\deg h + 1) - \deg g \deg h = \deg g + \deg h + 1$$

and hence φ is surjective, leading to the existence of $\overline{a}, \overline{b}$ as wanted in $(\star)$. This completes the induction.

Corollary 2.5. Let K be a complete valued field, and $f = a_0 + a_1 T + \cdots + a_n T^n \in K[T]$ is an irreducible polynomial with $a_0, a_n \neq 0$. Then,

$$|a_i|_K \leq \max\{|a_0|_K, |a_n|_K\} \quad \text{for all} \quad i \leq n.$$

Proof. Assume f is primitive by scaling the coefficients. By definition, we need $\max\{|a_0|_K, |a_n|_K\} = 1$. Otherwise, take a minimal m such that $|a_m|_K = 1$; by assumption, we have $0 < m < n$, and moreover

$$f(x) \equiv x^m(a_m + a_{m+1}x + \cdots + a_n x^{n-m}) \pmod{\mathfrak{m}_K}.$$

By Hensel's lemma, this lifts to a factorisation of f in $K[T]$, and thus f has a nontrivial factor of degree $n - m$, which contradicts the irreducibility of f. □

It turns out this result can be vastly generalised by introducing the notion of a *Newton polygon*, which we will discuss later when we have proved enough results about valuations. Roughly speaking, an irreducible polynomial corresponds to a single straight line segment in the Newton polygon, and the above corollary is a special case where all other points are above this line segment.

Finally, we have one more application of Hensel's Lemma, instrumental in extending the absolute $|\cdot|_K$ of a complete valued field K to a finite extension L/K of such fields while preserving the absolute value properties and completeness of L with respect to this extended metric.

Corollary 2.6. Let K be a complete valued field, and L/K be a finite extension. Then, the set

$$\mathcal{O}_L \overset{\text{def}}{=} \{x \in L : |N_{L/K}(x)|_K \leq 1\}$$

is closed under addition by 1.

Proof. Clearly, $0 + 1 = 1 \in \mathcal{O}_L$, so it suffices to show that for nonzero $\alpha \in \mathcal{O}_L$, we have $\alpha + 1 \in \mathcal{O}_L$. Suppose α has the minimal polynomial

$$f(T) = T^n + a_{n-1}T^{n-1} + \cdots + a_0 \in K[T].$$

From properties of $N_{L/K}$ from Part II Galois Theory, we know there is $m = \frac{1}{n}[L : K] \in \mathbb{N}$ such that $N_{L/K}(\alpha) = a_0^m$. Since by assumption $|N_{L/K}(\alpha)|_K \leq 1$, we have $|a_0|_K \leq 1$.

By Hensel's lemma, for each $0 < i < n$, we have

$$|a_i|_K \leq \max\{1, |a_0|_K\} = \max\{1, |N_{L/K}(\alpha)|_K^{1/m}\} \leq 1.$$

Hence, $|a_i|_K \leq 1$ for each i. It follows $f(T) \in \mathcal{O}_K[T]$, and hence the minimum polynomial of $\alpha + 1$ is also in $\mathcal{O}_K[T]$. Hence, $|N_{L/K}(\alpha + 1)| \in \mathcal{O}_K$, and hence

$$|\alpha + 1|_L = |N_{L/K}(\alpha + 1)|_K^{1/n} \leq 1. \quad \square$$

In fact, with a bit more work, it is possible to show that $\mathcal{O}_L$ is, as suggested by the name, the ring of integers associated with L, and is the integral closure of $\mathcal{O}_K$ in L. The above proof shows one direction, that is, any $\alpha \in \mathcal{O}_L$ is integral over $\mathcal{O}_K$.

3 Finite extensions of $\mathbb{Q}_p$

In the case of the usual Euclidean metric $|\cdot|$ on $\mathbb{Q}$, the completion $\widehat{\mathbb{Q}} = \mathbb{R}$ has only one nontrivial algebraic extension, $\mathbb{C}$. However, in the case of the p–adic metric $|\cdot|_p$, there are many more such extensions; in fact, there are infinitely many of them.

For instance, for any integer $n \geq 2$, we have $\sqrt[n]{p} \notin \mathbb{Q}_p$; otherwise, if there exists $x \in \mathbb{Q}_p$ such that $x^n = p$, then taking the valuation of both sides gives

$$n \cdot \nu_p(x) = 1$$

which is absurd as ν_p takes values in $\mathbb{Z} \cup \{\infty\}$. It follows that $\mathbb{Q}_p(\sqrt[n]{p})$ is a proper extension of $\mathbb{Q}_p$, for each $n \geq 2$.

However, it seems plausible that we may be able to *extend* ν_p in a natural way to the finite extension $\mathbb{Q}_p(\sqrt[n]{p})$ by allowing $\nu_p(\sqrt[n]{p}) = 1/n$. We would hope that such an extension is possible in general, as it would allow us to talk about the valuation of roots of polynomials.

Indeed, this is possible by the following theorem, which also states that such an extension is unique.

Theorem 3.1. Let $(K, |\cdot|_K)$ be a complete valued field, and L/K be a finite extension with $n = [L : K]$. Then, there exists an extension of absolute values $|\cdot|_L$, such that $(L, |\cdot|_L)$ is a complete valued field and

$$|\alpha|_L^n = |N_{L/K}(\alpha)|_K.$$

Moreover, $|\cdot|_L$ is unique.

In particular, since any element α in the algebraic closure $\overline{\mathbb{Q}}_p$ lies in a finite extension $\mathbb{Q}_p(\alpha)$ over $\mathbb{Q}_p$, there is a unique extension of $|\cdot|_p$ to $\overline{\mathbb{Q}}_p$. By the theorem above, the corresponding extended valuation ν_p to $\overline{\mathbb{Q}}_p$ now takes values in $\mathbb{Q} \cup \{\infty\}$.

It is interesting, though not so important here, to note that $(\overline{\mathbb{Q}}_p, |\cdot|_p)$ is *not* complete as a consequence of the Baire category theorem. The completion with respect to this metric is denoted $\mathbb{C}_p$, which turns out is also algebraically closed.

To prove theorem 3.1, we first need a primer on norms on vector spaces. The idea of the proof is to consider a weaker notion of an absolute value, and show that in fact any two norms on a finite dimensional vector space over a complete valued field are equivalent. This will allow us to prove the uniqueness and completeness of $|\cdot|_L$.

3.1 Vector space norms

As in the theorem, we would like to show that there is an absolute value on L extending $|\cdot|_K$ for which it is complete, and show it's unique and equal to the norm given by the theorem.

To do this, we first consider absolute values over L as regular norms on *vector spaces* over K. We derive intuition from a similar argument in Part II Linear Analysis, where we showed that any two norms on a finite–dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$ induce the same topology. It turns out that the same theorem holds for complete valued fields, though the argument is different.

Definition 3.2. A *norm* on a vector space V over a valued field K is a function $\|\cdot\| : V \to \mathbb{R}$ such that

- $\|v\| \geq 0$ for all $v \in V$, and $\|v\| = 0$ if and only if $v = 0$;

- $\|\lambda v\| = |\lambda| \cdot \|v\|$ for all $\lambda \in K$ and $v \in V$, with the usual absolute value $|\cdot|$ on K;
- $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.

Two norms $\|\cdot\|, \|\cdot\|'$ are considered *equivalent*, if there are constants $C, C' > 0$ such that

$$C \cdot \|v\| \leq \|v\|' \leq C' \cdot \|v\|$$

for all $v \in V$.

An example of a norm is the *max* (or *sup*) norm. Given a finite dimensional vector space V over a valued field K with a basis $\{x_i\}_{i \leq n}$, the max norm is defined as

$$v = \sum_i v_i x_i \implies \|v\|_\infty = \max_{i \leq n} |v_i|.$$

This norm is convenient to work with, as it also plays nicely with ultrametric spaces such as $\mathbb{Q}_p$.

Theorem 3.3. Let K be a complete valued field, and V is a finite dimensional K-vector space. Then, any norm $\|\cdot\|$ is equivalent to $\|\cdot\|_\infty$.

In Part II Linear Analysis, the corresponding theorem for real or complex vector spaces came down to the Heine–Borel theorem which states the unit ball in the supremum norm is compact. However, we no longer have as much control over the topology of K which we took for granted.

Proof. One inequality is easy; for fixed basis $\{x_i\}$, notice that for $C' = \max_{i \leq n} \|x_i\|$,

$$\|x\| = \left\| \sum_i a_i x_i \right\| \leq \max_{i \leq n} |a_i| \cdot \|x_i\| \leq \left(\max_{i \leq n} |a_i| \right) \cdot \max_{i \leq n} \|x_i\| = C' \cdot \|x\|_\infty.$$

For the other direction, we use induction on n. For $n = 1$, take a basis $\{x_1\}$. Notice that

$$\|x\| = \|a_1 x_1\| = |a_1| \cdot \|x_1\|,$$

and so $C = \|x_1\|$ will work. For the induction step, denote

$$V_i = \bigoplus_{j \neq i} K x_j, \quad \text{and} \quad W = \bigcup_{i \leq n} (x_i + V_i) \subsetneq V.$$

By induction, any norm on V_i is equivalent to the max norm. Notice that V_i is complete with respect to the max norm, due to K being complete, and pointwise convergence implying norm convergence in finite dimensions. Hence V_i is complete with respect to any norm and so V_i is a closed subspace of V under the norm on V. In particular, since W is a finite union of closed sets, it is also closed.

Notice that $0 \notin W$ by construction, and so by definition of a closed subset, we can find $C > 0$ such that $B(0, C) \cap W = \emptyset$. Equivalently, any $x \in W$ has $\|x\| \geq C$.

It suffices to show that $\|x\| \geq C \cdot \|x\|_\infty$, which would complete the inductive step. Pick k such that $|a_k| = \max_{i \leq n} |a_i| = \|x\|_\infty$. Then

$$\|x\|_\infty^{-1} \|x\| = \left\| \sum_{i \neq r} \frac{a_i}{a_r} x_i + x_r \right\| \geq C$$

since the right side is an element of V_r. □

3.2 Proof of the Extension Theorem

We now have all the tools to prove the extension theorem, which involves piecing together all the results we have proved so far.

Proof of Theorem 3.1. If a norm $|\cdot|_L$ exists which extends $|\cdot|_K$, then it is automatically complete since L is finite dimensional over the complete valued field K.

Suppose $|\cdot|_L$ and $|\cdot|_L'$ are two extensions of $|\cdot|_K$ to L satisfying the conditions of the theorem. Then, as we have previously shown, these two norms are equivalent. However, by Theorem 2.2, the equivalence of absolute values implies there is $r > 0$ such that $|x|' = |x|^r$ for any $x \in L$. Since they agree nontrivially on K, we must have $r = 1$, and hence $|\cdot| = |\cdot|'$.

It remains to show that such an absolute value exists, i.e. the function

$$|\alpha|_L \stackrel{\text{def}}{=} |N_{L/K}(\alpha)|_K^{1/n}$$

is a norm on L. All three properties are easy to verify with the results we have proved:

- If $|\alpha| = 0$, then $|N_{L/K}(\alpha)| = 0$, and hence $\alpha = 0$.
- Since $N_{L/K}$, $|\cdot|$ and $\sqrt[n]{\cdot}$ are multiplicative, so is $|\cdot|_L$.
- For strong triangle inequality, let $\alpha, \beta \in L$ with $|\alpha|_L \le |\beta|_L$. If $\beta = 0$, then clearly $|\alpha + 0|_L \le |\alpha|_L$. Otherwise,
$$|\alpha + \beta|_L \le |\beta|_L \iff |\alpha/\beta + 1|_L \le 1.$$
Hence, it suffices to show that $|\alpha|_L \le 1 \implies |\alpha + 1|_L \le 1$. To do this, consider the ring
$$\mathcal{O}_L \stackrel{\text{def}}{=} \{x \in L : |x|_L \le 1\} = \{x \in L : |N_{L/K}(x)|_K \le 1\}.$$
By Corollary 2.6, $\mathcal{O}_L$ is closed under addition by 1, which is exactly what we want. □

4 Newton Polygons

So far, we have come across an object $\mathbb{Q}_p$, which is a complete field with an exotic metric $|\cdot|_p$ which also gives rise to a corresponding valuation ν_p taking values in $\mathbb{Q} \cup \{\infty\}$.

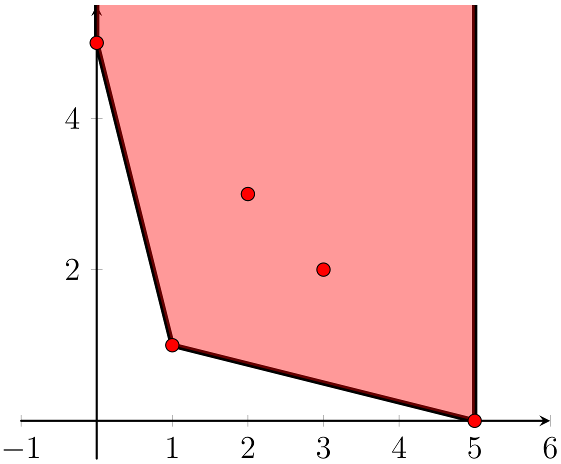
In this section, we present a tool called the *Newton polygon*, which is a useful visual tool in studying the roots of polynomials in $\mathbb{Q}_p$ and the degrees of their irreducible factors over $\mathbb{Q}_p$.

Definition 4.1. Let $f(T) = a_nT^n + \cdots + a_0 \in K[T]$ be a polynomial over a complete valued field K with valuation ν. The *Newton polygon* of f is the lower convex hull of the points $\{(i, \nu(a_i)) : a_i \neq 0\}$.

By the *lower convex hull* of the points $(i, \nu(a_i))$, we actually mean taking the convex hull of *all* points $(i, \nu(a_i))$ for $i \in \mathbb{Z}$ with the convention that $\nu(0) = \infty$. In particular, the leftmost and rightmost sides of the polygon will be vertical lines.

However, for our purposes, we will only consider segments of this convex hull with real gradient, hence the name *lower* convex hull. Graphically, this line is the lower envelope of the points $(i, \nu(a_i))$.

The utility of these Newton polygons is that they give us information about the roots of the polynomial f over K. This is interesting, as the Newton polygon is a purely algebraic object, but their geometric shape is closely related to the factorisation of f.



Newton Polygon for $f(T) = p^5 + pT + p^3T^2 + p^2T^3 + T^5$.

Theorem 4.2. Let K be a complete valued field with valuation ν. Let $f(T) = a_nT^n + \cdots + a_0 \in K[T]$ be a polynomial with Newton polygon Γ, and L/K the splitting field of f.

Suppose Γ has a line segment $(r, \nu(a_r)) \to (s, \nu(a_s))$ with $s \geq r$, and slope $-m \in \mathbb{R}$. Then, f has precisely $s - r$ roots in L, whose valuation is m.

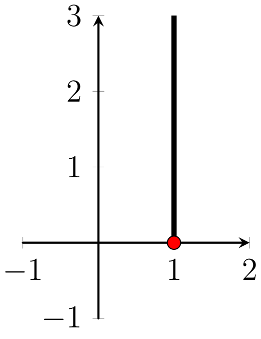
In other words, suppose Γ has successive vertices $(x_0, y_0), \ldots, (x_k, y_k)$ with $x_0 < \cdots < x_k$. Then, over K, the polynomial f factors as

$$f(x) = f_1(x) \ldots f_k(x)$$

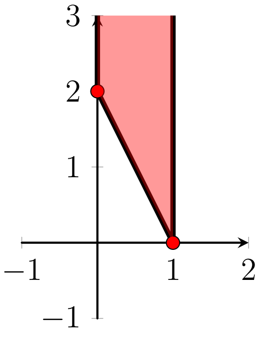
such that $\deg f_i = x_i - x_{i-1}$, and all the roots of f_i have valuation $-\frac{y_i - y_{i-1}}{x_i - x_{i-1}}$ over K. This is since the successive gradients of the line segments of Γ are strictly increasing, and thus all $s - r$ roots of f with valuation m as guaranteed in the theorem must belong to a single factor f_k.

Before we attempt to prove this theorem, it is useful to observe the conclusion of this theorem for polynomials of small degree.

- For $f(T) = T - \alpha$, the Newton polygon is a single line segment joining $(0, \nu(\alpha))$ with $(1, 0)$. Hence, f has a single root α with valuation $-\nu(\alpha)$, which is also its gradient.



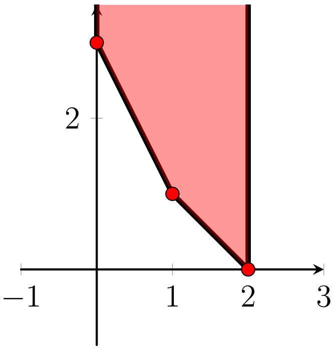
Newton Polygon for $f(T) = T$.



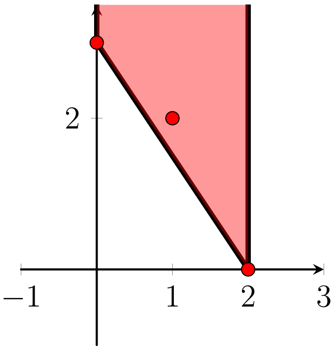
Newton Polygon for $f(T) = T - p^2$.

- For $f(T) = (T - \alpha)(T - \beta)$ with $\nu(\alpha) \geq \nu(\beta)$, the constant term is $\alpha\beta$, so one point on the polygon is $(0, \nu(\alpha) + \nu(\beta))$. Another point is $(2, 0)$, from the leading coefficient.
  - If $\nu(\alpha) > \nu(\beta)$, then $\nu(\alpha + \beta) = \nu(\beta)$ so we have two distinct slopes, which is consistent with the theorem.

- If $\nu(\alpha) = \nu(\beta)$, then $\nu(\alpha+\beta) \geq \nu(\alpha) = \nu(\beta)$, so we have a single line segment for the Newton polygon, again consistent with the theorem.



Newton Polygon for $f(T) = p^3 + pT + T^2$.



Newton Polygon for $f(T) = p^3 + p^2T + T^2$.

Proof of Theorem 4.2. We may assume $a_n = 1$ by dividing out by a_n. This only changes Γ up to vertical shifting, which does not affect the gradients of Γ. Enumerate the roots of f as α_i such that

$$\begin{aligned} \nu(\alpha_1) = \cdots = \nu(\alpha_{r_1}) = m_1 \\ \nu(\alpha_{r_1+1}) = \cdots = \nu(\alpha_{r_2}) = m_2 \\ \vdots \\ \nu(\alpha_{r_k+1}) = \cdots = \nu(\alpha_n) = m_{k+1}, \end{aligned}$$

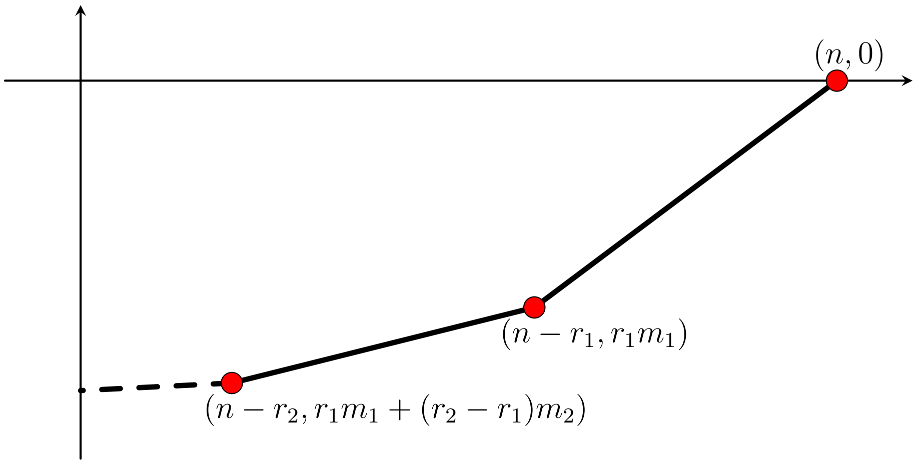
where we have $m_1 < \cdots < m_{k+1}$. Of course, we would like to show that $-m_i$ are the gradients of the Newton polygon, which would prove the result. We have the following inequalities:

$$\begin{aligned} \nu(a_n) &= \nu(1) = 0 \\ \nu(a_{n-1}) &= \nu\left(\sum_i \alpha_i\right) \geq \min_{i\leq n} \nu(\alpha_i) = m_1 \\ \nu(a_{n-2}) &= \nu\left(\sum_{i\neq j} \alpha_i\alpha_j\right) \geq \min_{1\leq i\neq j\leq n} \nu(\alpha_i\alpha_j) = 2m_1 \\ &\vdots \\ \nu(a_{n-r_1}) &= \nu\left(\sum_{i_1\neq\ldots\neq i_{r_1}} \alpha_{i_1}\ldots\alpha_{i_{r_1}}\right) = \min_{1\leq i_1\neq\ldots\neq i_{r_1}\leq n} \nu(\alpha_{i_1}\ldots\alpha_{i_{r_1}}) = r_1m_1 \end{aligned}$$

where equality holds at the r_1–th stage since the term $\alpha_1\ldots\alpha_{r_1}$ has valuation strictly less than all other terms in the sum, and we recall that $\nu(a+b) = \nu(a)$ if $\nu(a) < \nu(b)$. Continuing on, we obtain more inequalities:

$$\begin{aligned} \nu(a_{n-r_1+1}) &\geq r_1m_1 + m_2 \\ &\vdots \\ \nu(a_{n-r_2}) &= r_1m_1 + (r_2-r_1)m_2 \\ &\vdots \\ \nu(a_{n-r_3}) &= r_1m_1 + (r_2-r_1)m_2 + (r_3-r_2)m_3 \end{aligned}$$

and so on. This implies the Newton polygon Γ of f looks like the following:



By the inequalities above, we know that the points drawn in are vertices of the Newton polygon, since all the other points lie on or above these segments. In particular, for $j \geq 2$, the j–th segment has gradient

$$\frac{\left(r_1m_1 + \sum_{i=1}^{j-2}(r_{i+1} - r_i) \cdot m_{i+1}\right) - \left(r_1m_1 + \sum_{i=1}^{j-1}(r_{i+1} - r_i) \cdot m_{i+1}\right)}{r_ij - r_{j-1}} = \frac{-(r_j - r_{j-1}) \cdot m_j}{r_j - r_{j-1}} = -m_j,$$

and for $j = 1$, the gradient is simply

$$\frac{0 - r_1m_1}{n - (n - r_1)} = -m_1.$$

It follows that $-m_i$ are the gradients of the Newton polygon, as desired. □

An immediate corollary of this theorem concerns the degrees of irreducible factors of polynomials over $\mathbb{Q}_p$.

Corollary 4.3. Let f be a polynomial over $\mathbb{Q}_p$ with Newton polygon Γ. Suppose $d \geq 1$ is a positive integer that divides the denominator of all slopes (in lowest terms) of Γ.

Then, d must also divide the degree of each irreducible factor of f of degree $x_i - x_{i-1}$.

Proof. Take an irreducible factor f_i of f, and let $\alpha \in \overline{\mathbb{Q}}_p$ be some root of f_i. Denote $K = \mathbb{Q}_p(\alpha)$. Then, by Theorem 4.2, d divides the denominator of $\nu(\alpha)$. By Theorem 3.1, we obtain

$$|\alpha|_p^n = |N_{K/\mathbb{Q}_p}(\alpha)|_p \implies n \cdot \nu_p(\alpha) = \nu_p(N_{K/\mathbb{Q}_p}(\alpha)) \in \mathbb{Z}$$

since $N_{K/\mathbb{Q}_p}(\alpha)$ is a rational number. Hence, d must divide $n = [K : \mathbb{Q}_p] = \deg f_i$. □

4.1 Application to the Truncated Exponential

The importance of this corollary is that the Newton polygon of the truncated exponential f_n takes a form that is easy to analyse, as the p–adic valuation of factorials has a well–known formula.

Suppose n has the p–adic expansion

$$n = a_1p^{n_1} + a_2p^{n_2} + \cdots + a_tp^{n_t}$$

where $0 < a_i < p$ for each i, and $n_1 > n_2 > \cdots > n_t$. A well–known result regarding factorials is that

$$\nu_p(n!) = \sum_{k \geq 1} \left\lfloor \frac{n}{p^k} \right\rfloor = \frac{n - s_p(n)}{p - 1},$$

where $s_p(n) = a_1 + \cdots + a_t$ is the sum of digits in the p–adic expansion of n.

Lemma 4.4. Denote $x_i = a_1 p^{n_1} + \cdots + a_i p^{n_i}$. Then, the vertices of the Newton Polygon Γ of f_n are

$$\{(x_i, -\nu_p(x_i!)) : 1 \leq i \leq t\}.$$

Proof. Any number $x_i < y < x_{i+1}$ can be written as

$$y = a_1 p^{n_1} + \cdots + a_i p^{n_i} + b$$

where $0 < b < a_{i+1} p^{n_{i+1}}$. Hence, the gradient between $(x_i, -\nu_p(x_i!))$ and $(y, -\nu_p(y!))$ is

$$m = \frac{-\nu_p(y!) + \nu_p(x_i!)}{y - x_i} = \frac{(x_i - s_p(x_i)) - (y - s_p(y))}{(y - x_i)(p-1)} = -\frac{b - s_p(b)}{b(p-1)}.$$

Letting $b_0 = a_{i+1} p^{n_{i+1}}$, it remains to show

$$-\frac{b - s_p(b)}{b(p-1)} \geq -\frac{b_0 - s_p(b_0)}{b_0(p-1)} \iff \frac{s_p(b)}{b} \geq \frac{s_p(b_0)}{b_0} = p^{-n_{i+1}}.$$

Denote $n \overset{\text{def}}{=} n_{i+1}$ for simplicity, and take the p–adic expansion $b = b_0 + b_1 p + \cdots + b_n p^n$. The inequality is equivalent to

$$p^n(b_0 + b_1 + \cdots + b_n) - (b_0 + b_1 p + \cdots + b_n p^n) \geq 0$$

which is true since $p^i < p^n$ for all $i < n$. □

A corollary of the above proof is that the Newton Polygon of f_n has slopes

$$m_i = -\frac{p^{n_i} - 1}{p^{n_i}(p-1)}.$$

Using our previous result about Newton polygons, we can obtain the following results:

Corollary 4.5. For prime p, suppose $p^m \mid n$. Then p^m divides the degree of all factors of f_n over $\mathbb{Q}_p$.

Proof. Since $p^m \mid n$, notice that $m \leq n_t < n_{t-1} < \cdots$. Hence, $m < n_i$ for each i, and so p^m divides the denominator of each m_i by the formula above. The result now follows from Theorem 4.2.

Corollary 4.6. Suppose $p^m \leq n$. Then p^m divides the degree of the splitting field of f_n over $\mathbb{Q}_p$.

Proof. Clearly, $m \leq n_1$. Hence, p^m divides the denominator of m_1 by the formula above. By a similar argument to Corollary 4.3, if $L/\mathbb{Q}_p$ is the splitting field of f_n, then since

$$[L : \mathbb{Q}_p] \cdot \nu_p(\alpha) = \nu_p(N_{L/\mathbb{Q}_p}(\alpha)) \in \mathbb{Z},$$

we obtain p^m divides $[L : \mathbb{Q}_p]$, which is what we wanted. □

4.2 Irreducibility of f_n

We are now ready to prove the irreducibility of f_n over $\mathbb{Q}$.

Suppose f_n has an irreducible factor g. Then, by Corollary 4.5, by taking the reduction of f_n over $\mathbb{Q}_p$, we know that $p^m \mid n$ implies p^m divides the degree of all factors of g. In particular, p^m divides the degree of g, and hence $n \mid \deg g$. This shows that f_n is irreducible over $\mathbb{Q}$, as desired. □

Later, we will use Corollary 4.6 to obtain the existence of a p–cycle in the Galois group of f_n, for many prime values p. The next section will then show that this implies $\text{Gal}(f_n/\mathbb{Q})$ contains A_n.

5 Jordan's Symmetric Group Theorem

It remains to show that the Galois group contains A_n, as this would leave us to compute the discriminant to fully determine the Galois group.

In general, there isn't a single nice algorithm or method to show such facts. For smaller n, there are some tricks involving computing the Lagrange resolvent (for cubics) or the resolvent cubic (for quartics). Unfortunately, these rely on the heuristic that there are a relatively small number of subgroups of S_n for small n, and such methods don't generalise well to general n.

However, for special classes of polynomials, there are group theoretic approaches that can be used to show such facts. One such approach is used here.

Definition 5.1. Let Ω be a finite, non–empty set. A finite permutation group G acting on Ω is *primitive* if it is transitive, and the only G–invariant equivalence relations on Ω are the trivial ones.

Equivalently, the action is primitive if for any partition $\Omega = \Omega_1 \cup \cdots \cup \Omega_k$ into non–empty subsets, the action of G on Ω reduces to a permutation of the sets Ω_k if and only if either $k = 1$ or $k = |\Omega|$.

The point of the above definition is the following theorem about primitive group actions, which looks surprisingly close to the missing piece of our proof.

Theorem 5.2. Let Ω, G be as before, such that $|\Omega| = n$ and G is primitive on Ω. Suppose G contains a p–cycle for some prime $p < n - 2$. Then, G contains A_n.

Clearly, this theorem is not true if G is just transitive on Ω. For example, the cyclic group C_9 acts on itself by left multiplication, is transitive, and contains a 3–cycle, yet does not contain A_9.

The idea is to apply this theorem to $\mathrm{Gal}(f_n/\mathbb{Q})$ acting on its roots. We know from Part II Galois Theory that, as long as f is irreducible over $\mathbb{Q}$, the Galois group acts transitively on its roots. If we can also show the action is primitive, the above theorem will show that $\mathrm{Gal}(f/\mathbb{Q})$ contains A_n.

Corollary 5.3. Let $f(T) \in \mathbb{Q}[T]$ be an irreducible polynomial of degree n. Suppose $\mathrm{Gal}(f/\mathbb{Q})$ contains a p–cycle for some prime $\frac{n}{2} < p < n - 2$. Then, $\mathrm{Gal}(f/\mathbb{Q})$ contains A_n.

In other words, if we have a cycle of prime size p where p is large enough, then we can force the transitive group action to also be primitive. This isn't surprising, since the existence of such p–cycles mean that any nontrivial partition fixed by G must have its components of size at least p.

Proof of Corollary. As above, denote $G = \mathrm{Gal}(f/\mathbb{Q})$, acting on Ω, the set of roots of f. Since f is irreducible, G acts transitively on Ω.

Choose the prime p as guaranteed in the corollary. Take any partition $\Omega = \Omega_1 \cup \cdots \cup \Omega_k$ that is fixed by the action of G. For each Ω_i, there are two cases:

- There is an element $g \in G$ of order p, such that for some distinct $x, y \in G$, $g(x) = y$. Then, the orbit of x under G has size p, and is fully contained in Ω_i.

  It follows $|\Omega_i| \geq p > n/2$.

- There is no such g. Assume $|\Omega_i| > 1$, and find distinct $x, y \in \Omega_i$. Since there exists a p–cycle in G, and G acts transitively on Ω, there is a p–cycle containing x, and a p–cycle containing y.

  Since $2p > n$, there is an element common in both p–cycles. It follows that x and y are in the same p–cycle, and hence in the same orbit, which contradicts our earlier claim that there is no such g.

  It follows $|\Omega_i| = 1$.

In particular, if $\Omega_1 > 1$, then $|\Omega_1| > n/2$, and hence there cannot be any other nontrivial partition. It follows G preserves Ω_1. Since G acts transitive, $\Omega_1 = \Omega$.

Hence, it follows that any partition of Ω fixed by G is trivial, and hence the action is primitive. By the theorem, G contains A_n. □

Theorem 5.2 is perhaps the most difficult and technically involved result to prove in this essay and will require a lot of work in group theory. We will first prove several results about group actions, and eventually prove the theorem using these results.

5.1 Translates and Blocks

In a transitive group action, any partition fixed by G must have the property that each component of that partition has the same size. Indeed, any two such components must be bijectively mapped to each other by a group element taking an element in one component to another.

This motivates the following definition of a *block*, which is essentially one component of a partition fixed by G. The other components of the partition are obtained by *translating* this block by a group element.

Definition 5.4. Suppose G is a group acting transitively on a finite set Ω, and $\emptyset \neq X \subseteq \Omega$.

- The set $gX \overset{\text{def}}{=} \{gx \mid x \in X\}$ is called a *translate* of X.
- The set X is called a *block* if for any $g \in G$, either $gX = X$ or $gX \cap X = \emptyset$.

Translates are like cosets of a group G. Indeed, if G acts on a subgroup H, then H itself is a block, since the cosets partition H. This allows us to prove theorems such as Lagrange's theorem, which are based on these partitions. This intuition is important and motivate the following properties.

Lemma 5.5. If G acts transitivity on Ω, and $|\Omega|$ is prime, then G acts primitively on Ω.

Proof. By the above discussion, any partition of Ω corresponds to a block Δ whose translates cover Ω. Hence $|\Delta|$ divides $|\Omega|$. Since $|\Omega|$ is prime, it follows that $|\Delta| = 1$ or $|\Delta| = |\Omega|$. □

Lemma 5.6. The intersection of two blocks with nonempty intersection is a block.

Proof. Take any two blocks Δ, Δ', and assume they have non–empty intersection $\Delta \cap \Delta' = \Theta$. Suppose $\Theta \cap \Theta^g \neq \emptyset$ for some $g \in G$. By definition, $\Delta \cap g\Delta \neq \emptyset$ and $\Delta' \cap g\Delta' \neq \emptyset$. Since Δ, Δ' are blocks, it follows

$$\Theta = \Delta \cap \Delta' = g\Delta \cap g\Delta' = g\Theta.$$

It follows that Θ is also a block of Ω. □

Lemma 5.7. Suppose G acts transitively on Ω, and $X \subseteq \Omega$. Pick any $\omega \in \Omega$. Then, the set

$$\Delta \overset{\text{def}}{=} \bigcap_{g \in G,\, \omega \in gX} gX$$

is a block of G, containing ω.

In particular, if G is primitive, then Δ is a block containing ω, and so G fixes the partition of Ω into the blocks $\{g\Delta \mid g \in G\}$. By the definition of primitivity, this partition must be trivial. Hence, for $X \subsetneq \Omega$, clearly $|\Delta| \leq |X| < |\Omega|$, and so $\Delta = \{\omega\}$.

Proof of Lemma. Take any $h \in G$ with $\Delta \cap h\Delta \neq \emptyset$; it suffices to show $\Delta = h\Delta$.

- If $\omega \in h\Delta$, then $h^{-1}(\omega) \in gX$ whenever $\omega \in gX$. This means $\omega \in g\Delta \implies \omega \in gh\Delta$, so $\Delta \subseteq h\Delta$. Since translates have the same size, it follows $\Delta = h\Delta$.
- Otherwise, choose $\alpha \in \Delta \cap h\Delta$. Since G acts transitively on Ω, there is $k \in G$ with $k(\omega) = \alpha$. By definition of α, this means $\omega \in k^{-1}\Delta, hk^{-1}\Delta$ and by definition, in Δ.

  Hence, we have $\omega \in \Delta \cap k^{-1}\Delta \neq \emptyset$, and by the previous argument, $\Delta = k^{-1}\Delta$. Similarly, $\Delta = hk^{-1}\Delta$. Applying k both sides yield $\Delta = h\Delta$, as desired. □

The result of the previous set bash is that we have the following trivial, yet useful corollary, which is sort of a separability property of blocks.

Corollary 5.8. Suppose G acts primitively on Ω, and $\emptyset \neq X \subsetneq \Omega$. Then, for any $\alpha \neq \beta \in \Omega$, there exists $g \in G$ with $\alpha \in gX$ but $\beta \notin gX$.

The proof of course follows from the lemma, where we deduced

$$\{\omega\} = \bigcap_{g \in G,\, \omega \in gX} gX.$$

This is a strengthening of the deduction when G is just transitive on Ω, which can be seen by taking $X = G \setminus \{\alpha\}$ which yields $g \in G$ with $g(\beta) = \alpha$. This strengthening will be important in proving a key lemma necessary for Jordan's Theorem.

5.2 Pointwise and Setwise Stabilisers

We wish to consider the notion of multiple transitivity and multiple primitivity, and its relation to the property of being a primitive action.

Primitivity does *not* imply 2–transitivity; indeed, $D_{10} = \langle r, s \mid r^5 = s^2 = e,\, srs^{-1} = r^{-1} \rangle$ acts on the cosets Ω of a subgroup $\langle s \rangle$ of order 2 by right multiplication. Since Ω has size 5, the action is primitive; yet, it is not 2–transitive, since otherwise D_5 will act transitively on the set of ordered pairs of distinct points on Ω, which has size 20, which is greater than the size of D_5.

Hence, in a definition of multiple transitivity and primitivity, we wish to strengthen the conditions to not just requiring the action to be primitive, but some of its stabilisers of certain sets too.

Definition 5.9. Let G be a group acting on a set Ω, and $X \subseteq \Omega$.

- The *pointwise stabiliser* of X is the subgroup $G_X \overset{\text{def}}{=} \{g \in G \mid \forall x \in X,\, g(x) = x\}$.
- The *setwise stabiliser* of X is the subgroup $G_{(X)} \overset{\text{def}}{=} \{g \in G \mid g(X) = X\}$.

In the case where $X = \{x\}$, we write $G_x \overset{\text{def}}{=} G_{\{x\}}$.

These definitions appear similar to the definitions of a centraliser and normalizer of a subgroup. In fact, they are analogous, since we recover those definitions by having G act on itself by conjugation and taking X to be a subgroup.

An obvious observation is that $G_{(X)}$ acts on X naturally, and whose kernel is G_X. Hence, we know $G_X \trianglelefteq G_{(X)}$. Though, for most of our purposes, we're more interested in the group G_X.

Definition 5.10. Let G be a group acting transitively on a set Ω, and $X \subsetneq \Omega$. We call X a *Jordan set* if G_X is transitive on $\Omega \setminus X$, and a *strongly Jordan set* if this action is primitive.

Definition 5.11. An action of G on Ω is *doubly primitive* if the action of G on Ω is 2–transitive, and the action of G_ω on $\Omega \setminus \{\omega\}$ is primitive for all $\omega \in \Omega$.

The following lemma shows that the two notions above are very closely tied to each other.

Lemma 5.12. Let G be a group acting primitively on a finite set Ω. Suppose there exists a non–empty Jordan set $X \subset \Omega$ with $|X| \leq |\Omega| - 2$. Then, G is 2–transitive on Ω.

Furthermore, if X can be chosen to be strongly Jordan, then G is doubly primitive on Ω.

In some sense, the content of this lemma is that the *minimal strongly Jordan set is a singleton* if one exists. Since translates of strongly Jordan sets are also strongly Jordan, this means that all singleton sets are strongly Jordan.

Proof. By definition, it suffices to show that $\{\omega\}$ is a strongly Jordan set. It hence suffices to show that G is doubly primitive on Ω. We prove this by induction on $|X|$; for $|X| = 1$, the result is trivial, so we may assume $|X| > 1$. For convenience, denote $Y = \Omega \setminus X$.

For $g \in G$, denote $H_g = \langle G_X, gG_Xg^{-1} \rangle$ the direct sum of the groups. We have the following two cases:

- **Case 1:** $|X| < \frac{1}{2}|\Omega|$. Then $|Y| > \frac{1}{2}|\Omega|$, and so for any $g \in G$, we have $Y \cap gY \neq \emptyset$.

  Notice that G_X is transitive on Y by definition, and gG_Xg^{-1} is transitive on gY. Since $Y \cap gY \neq \emptyset$, it follows H_g is transitive on $Y \cup gY$, and $\overline{X} \overset{\text{def}}{=} X \cap gX$ is the set of points fixed by H_g.

  Pick $\alpha, \beta \in X$. By Corollary 5.8 there exists $g \in G$ such that $\alpha \in gX$ but $\beta \notin gX$. Clearly $\alpha \in \overline{X}$, however $\beta \notin \overline{X}$, so $1 \leq |\overline{X}| < |X|$, and hence by induction, G is doubly transitive.

  For the case when X is strongly Jordan, again by induction, it suffices to show that H_g is primitive.

  Claim. Suppose $G = \langle H, K \rangle$ acts transitivity on Ω, such that:

  - H acts primitivity on $\Gamma \subset \Omega$, and $H \leq G_{\Omega \setminus \Gamma}$;
  - K acts primitively on $\Delta \subset \Omega$ and $K \leq G_{\Omega \setminus \Delta}$.

  Then, G acts primitively on Ω.

  Proof. Since G is transitive, it follows $\Gamma \cup \Delta = \Omega$, and $\Gamma \cap \Delta \neq \emptyset$. WLOG assume $|\Delta| > \frac{1}{2}|\Omega|$. Notice that Δ is an orbit of K by definition.

  Suppose Θ is a block of G, and by taking an appropriate translate, assume $\Theta \cap \Delta \supseteq \{\alpha\} \neq \emptyset$. By definition, Θ is a block of K, and also $K\alpha \overset{\text{def}}{=} \{k\alpha : k \in K\} = \Delta$ is a block of K. By Lemma 5.6, $\Theta \cap \Delta$ is a block of K. By primitivity of K on Δ, it follows $\Theta \cap \Delta$ is either $\{\alpha\}$ or Δ.

  - If $\Theta \cap \Delta = \Delta$, then $\Delta \subseteq \Theta$. We know $|\Theta|$ divides $|\Omega|$ and $|\Theta| \geq \Delta > \frac{1}{2}|\Omega|$, and hence $\Theta = \Omega$.
  - If $\Theta \cap \Delta = \{\alpha\}$, then for any $g \in G$, $\Theta^g \cap \Delta$ is also a block of K, and hence equal to $\{\alpha\}$ or Δ. If it's equal to Δ for some $g \in G$, then we conclude as before that $g\Theta = \Omega$, so $\Theta = \Omega$.

    Otherwise, for any $g \in G$, the intersection $\Theta^g \cap \Delta$ has at most one element. Since Θ is a block, there are at least $|\Delta| > \frac{1}{2}|\Omega|$ such disjoint sets. Hence, the number of distinct $g\Theta$'s, which is a divisor of $|\Omega|$, is greater than $\frac{1}{2}|\Omega|$ and hence is equal to $|\Omega|$. It follows $|\Theta| = 1$. ■

- **Case 2:** $|X| \geq \frac{1}{2}|\Omega|$. Then, $|Y| \leq \frac{1}{2}|\Omega|$. Again, by Corollary 5.8, there is $g \in G$ such that $\alpha \in gX$ but $\beta \notin gX$. Similar to Case 1, we have $\alpha \in \overline{X} \neq \emptyset$ and H_g is transitive on $Y \cup gY$.

  Hence, we obtain $Y \subsetneqq Y \cup gY \subsetneqq \Omega$. Again by induction, we obtain that G is doubly transitive, or primitive, depending on whether X is Jordan or strongly Jordan. □

Interestingly enough, the same proof as above shows that the normal closure of G_X in G,

$$N \overset{\text{def}}{=} \langle gG_Xg^{-1} : g \in G \rangle,$$

is doubly transitive if there is a Jordan set X, and doubly primitive if X can be chosen to be strongly Jordan.

Earlier, we showed that primitivity does not imply 2–transitivity. The above lemma gives one condition that is enough to deduce 2–transitivity, namely the existence of a strongly Jordan set.

5.3 An alternative form of Jordan's Theorem

So far, the previous theorem discussed the case where an extra condition *upgrades* us from primitivity to 2–transitivity. Could we generalise this to a condition that upgrades us from primitivity to k–transitivity for any reasonable k? The answer is yes – we just need the set X to be large enough. The following theorem is the generalisation, which is a first major step towards proving Jordan's Theorem.

Theorem 5.13 [Jordan's Theorem, Alternative version]. Let G be a group acting primitively on a finite set Ω. Suppose $X \subseteq \Omega$ is a strongly Jordan set such that $|X| \leq |\Omega| - 2$.

Then, the action of G on Ω is $(|X| + 1)$–transitive.

This tells us that the previous theorem is selling us short – we can get higher order transitivity if we know that the size of X is large enough.

Of course, this doesn't mean we can be lazy and not prove the last result – we will be using it with induction to lift the transitivity, one step at a time.

Proof. We will prove the result by induction on $|\Omega|$. Of course, the result is vacuously true for $|\Omega| = 1, 2$.

For $|\Omega| = 3$, clearly $|X| = 1$. Let $\Omega = \{x, y, z\}$ and $X = \{z\}$, and $H \leq \mathrm{Sym}(\Omega)$ the set of permutations of Ω fixed by G. Since G_X is primitive, and hence transitive on $\{x, y\}$, H has a transposition. Furthermore, since G acts primitively on Ω, it follows H is a transitive subgroup of S_3 containing a transposition.

Hence, $H = S_3$, and so G acts 2–transitively on Ω, as required.

Henceforth, assume the result holds for all $|\Omega| < k$. Let $k = |\Omega| > 3$, and pick $\omega \in \Omega$. By Lemma 5.12, we know that G_ω is primitive on $\Omega \setminus \{\omega\}$ for any $\omega \in \Omega$. In particular, for $\omega \in X$, we have the following:

$$\begin{aligned}
(G_\omega)_{X \setminus \{\omega\}} &= \{g \in G_\omega \mid g \cdot x = x \ \forall x \in X \setminus \{\omega\}\} \\
&= \{g \in G \mid (g \cdot \omega = \omega) \wedge (g \cdot x = x \ \forall x \in X \setminus \{\omega\})\} \\
&= G_X.
\end{aligned}$$

By the hypothesis, G_X is primitive on $\Omega \setminus X = (\Omega \setminus \{\omega\}) \setminus (X \setminus \{\omega\})$. Furthermore, by the condition, $|X \setminus \{\omega\}| \leq |\Omega \setminus \{\omega\}| - 2$.

Hence, by induction, the action of G_ω on $\Omega \setminus \{\omega\}$ is $|X|$–transitive. It follows that G is $(|X| + 1)$–transitive on Ω, as required. □

5.4 Proving the Jordan Symmetric Group Theorem

We will first need to prove the theorem in a special case for $p = 3$ to get a 3–cycle.

Theorem 5.14 [Jordan, $p = 3$]. Let G be a primitive permutation group on a finite set Ω of size n. Suppose G contains a 3–cycle. Then, G contains A_n.

Proof. Denote $\sigma = (a, b, c) \in G$ be a 3–cycle, and let $X = \Omega \setminus \{a, b, c\}$. By definition, G_X contains σ, which implies G_X acts transitively on $\{a, b, c\}$.

Since 3 is prime, combined with the above fact, G_X is primitive on $\{a, b, c\}$, and furthermore $|X| = |\Omega| - 3 < |\Omega| - 2$. Hence, Theorem 5.13 shows that G is $(|\Omega| - 2)$–transitive on Ω. However, it is well known that this implies $G = \mathrm{Sym}(\Omega)$ or $\mathrm{Alt}(\Omega)$, and in particular contains $\mathrm{Alt}(\Omega)$. □

Of course, the proof of this case relies on this final fact that a $n - 2$–transitive subgroup $G \leq S_n$ must contain A_n, which is not true if we replace $n - 2$ with anything lower.

However, in the general case, we can reduce the problem to the case where $p = 3$. This will be done by using the fact that there is a p–cycle to force the existence of a 3–cycle. This will be where we use the fact that $p \leq n - 3$.

Theorem 5.2. Let Ω, G be as before, such that $|\Omega| = n$ and G is primitive on Ω. Suppose G contains a p–cycle for some prime $p < n - 2$. Then, G contains A_n.

Proof of Jordan's Theorem. We have already proven the result for $p = 3$. Hence, assume $p \neq 3$.

Let G, Ω be as in the theorem, and $\sigma \in G$ a p–cycle of G. Define the set

$$X \overset{\text{def}}{=} \{\omega \in \Omega \mid \sigma(\omega) = \omega\}.$$

Clearly, $|\Omega \setminus X| = p$, since the only elements not fixed by σ are part of a p–cycle by definition. Furthermore, it is obvious that $\sigma \in G_X$ and G_X acts transitively on $\Omega \setminus X$. Since p is prime, the action is primitive.

From the alternative version of Jordan's Theorem, G is $(|X| + 1)$–transitive on Ω. Any element $g \in G$ naturally induces a permutation of $\mathrm{Sym}(X)$, and hence by the multiple transitivity, there is a natural homomorphism $\varphi : G_{(X)} \to \mathrm{Sym}(X)$. It follows that

$$\mathrm{Sym}(X) \cong G_{(X)}/\ker\varphi = G_{(X)}/G_X.$$

Furthermore, since G is a permutation group, no non–identity element $g \in G$ fixes all elements of Ω, and hence no non–identity element $g \in G_X$ fixes all elements of $\Omega \setminus X$. It follows G_X acts faithfully on $\Omega \setminus X$. Hence, $G_X \leq \mathrm{Sym}(\Omega \setminus X)$, a group of size $p!$.

Next, let $P \overset{\text{def}}{=} \langle\sigma\rangle$. Clearly, $|P| = p$. Also, $P \subseteq G_X$ since σ fixes all members of X, and thus so do all its powers. Since $p \mid p!$ but $p^2 \nmid p!$, it follows P is a Sylow p–subgroup of G_X.

An argument of Frattini states that for groups $P \leq H \leq G$ such that P is a Sylow p–subgroup of H, then G can be written as the product of its subgroups

$$G = N_G(P)H$$

where $N_G(P)$ is the normalizer. Hence, $G_{(X)} = NG_X$. By the second isomorphism theorem,

$$G_{(X)}/G_X = NG_X/G_X \cong N/(N \cap G_X) = N/N_X \implies N/N_X \cong \mathrm{Sym}(X).$$

Observe that since $|X| = |\Omega| - p \geq 3$, the groups $\mathrm{Sym}(X)$, and hence $\mathrm{Alt}(X)$, contains a 3–cycle.

For a group $H \leq G$, denote $H' = [H, H] \leq G$ the commutator subgroup of H. We have

$$\mathrm{Alt}(X) = \mathrm{Sym}(X)' = (N/N_X)' = N'N_X/N_X \cong N'/(N' \cap N_X),$$

and hence there exists $\tau \in N'$ that induces a 3–cycle on X. Next, consider a homomorphism

$$\rho : N \to \mathrm{Aut}(P), \quad \rho(n)(\sigma) \overset{\text{def}}{=} n^{-1}\sigma n,$$

which has kernel $\ker(\rho) = C_N(P)$, the centralizer of P in N. By the first isomorphism theorem,

$$N/C_N(P) \cong \mathrm{Im}(\rho) \leq \mathrm{Aut}(P).$$

However, since $P = \langle \sigma \rangle$, $\mathrm{Aut}(P) \cong (\mathbb{Z}/p\mathbb{Z})^\times$ is an abelian group, and hence so is $N/C_N(P)$.

We claim that $N' \leq C_N(P)$. It suffices to show that $aba^{-1}b^{-1} \in \ker(\rho)$ for all $a, b \in N$. Since $N = N_{G_{(X)}}(P)$, we know that $a^{-1}Pa = b^{-1}Pb = P$, and hence there is $k, \ell \in \mathbb{Z}$ such that $a^{-1}\sigma a = \sigma^k$ and $b^{-1}\sigma b = \sigma^\ell$. Hence,

$$(aba^{-1}b^{-1})^{-1}\sigma(aba^{-1}b^{-1}) = \sigma^{k\ell k^{-1}\ell^{-1}} = \sigma,$$

and so $N' \leq C_N(P)$ as required. This means, $\tau \in N' \leq C_N(P)$ and hence τ commutes with σ.

Now, we analyse the permutation of Ω induced by τ. On $\Omega \setminus X$, τ commutes with a p–cycle σ. It is well known that an element τ commutes with a p–cycle σ in $\mathrm{Sym}(\Omega \setminus X) \cong S_p$ if and only if $\tau \in \langle \sigma \rangle$. On the other hand, τ induces a 3–cycle on X.

Hence, the element τ^p fixes $\Omega \setminus X$ and still induces a 3–cycle on X since $\gcd(3, p) = 1$ by assumption. Hence, G contains a 3–cycle. The theorem follows from our earlier discussion for the $p = 3$ case. □

6 Discriminant of f_n

A very useful tool in computing the Galois groups of a polynomial is to analyse the discriminant of the polynomial, which will tell us whether the group is contained in A_n or not.

Theorem 6.1. The discriminant of f_n is $(-1)^{n(n-1)/2}(n!)^n$.

Proof. Recall from Part II Galois Theory that the discriminant of a polynomial is given by

$$\mathrm{disc}(f) = (-1)^{\binom{n}{2}} \mathrm{Res}(f, f')$$

where $\mathrm{Res}(f, g)$ is the resultant of the two polynomials f, g. Furthermore, if $f(T) = a_nT^n + \cdots + a_0$ and $g(T) = b_mT^m + \cdots + b_0$, their resultant is given by the formula:

$$\mathrm{Res}(f, g) = \det \begin{bmatrix} a_n & a_{n-1} & \cdots & a_1 & a_0 & 0 & \cdots & 0 \\ 0 & a_n & a_{n-1} & \dots & a_1 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_1 & a_0 \\ b_m & b_{m-1} & \dots & b_1 & b_0 & 0 & \dots & 0 \\ 0 & b_m & b_{m-1} & \dots & b_1 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & b_m & b_{m-1} & \cdots & b_1 & b_0 \end{bmatrix}.$$

Hence, we obtain

$$\operatorname{Res}(f_n, f_n') = \det \begin{bmatrix} 1 & n & \cdots & \cdots & n! & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & 1 & n & \cdots & \cdots & n! \\ n & \cdots & \cdots & n! & & & \\ & n & \cdots & \cdots & n! & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & n & \cdots & \cdots & n! \end{bmatrix} = \det \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ n & \cdots & \cdots & n! & & & \\ & n & \cdots & \cdots & n! & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & n & \cdots & \cdots & n! \end{bmatrix} = (n!)^n.$$

It follows that $\operatorname{disc}(f_n) = (-1)^{n(n-1)/2}(n!)^n$. □

In particular, if $4 \mid n$, then the discriminant is a square number, and by a result in Part II Galois Theory, the Galois group is contained in A_n.

Otherwise, if $4 \nmid n$, then one of the following occur:

- n is odd. By Bertrand's Postulate, there is always a prime number p between $n/2$ and n. Hence, $p \mid n!$ but $p^2 \nmid n!$, and so $n!$ is never a square of an integer;
- n is even, but is not divisible by 4. In this case, it is easy to see that $\frac{n(n-1)}{2}$ is odd, and so the discriminant is negative, and hence not a square of an integer.

It follows that $\operatorname{Gal}(f_n/\mathbb{Q}) \subseteq A_n$ if and only if $4 \mid n$.

7 Conclusion and Further Results

Finally, we can prove the main result of this essay by combining the results we have proven.

Theorem 1.1. *The truncated exponential polynomial f_n has Galois group A_n over $\mathbb{Q}$ if and only if $4 \mid n$, and S_n otherwise.*

Proof. Denote $G = \operatorname{Gal}(f_n/\mathbb{Q})$.

- From Section 4.2, f_n is irreducible.
- From Corollary 4.6, G contains a p–cycle for any $p \leq n$.
- By Bertrand's Postulate[1], we can find a prime number p with $n/2 < p < n-2$.
- By Corollary 5.3, G contains A_n.
- By Theorem 6.4, we know that G is contained in A_n if and only if $4 \mid n$. □

7.1 Other results

Using just the results we have proved so far, we can immediately generalise the result to a larger class of polynomials, of the form

$$f(T) = c_n \frac{T^n}{n!} + c_{n-1}\frac{T^{n-1}}{(n-1)!} + \cdots + c_1 T + c_0.$$

[1] Bertrand of course states that there is a prime between $n/2$ and n for $n \geq 4$. However, we can modify the proof slightly to show that there is one strictly between $n/2$ and $n-2$ for $n \geq 8$.

where each c_i are nonzero and have all prime factors strictly greater than n. This ensures all results about the p–adic factorisations we have proved earlier remains true, and hence the Galois group contains A_n, to be determined by its discriminant.

However, this is quite a limited class of polynomials and we hope to do better. Indeed, we can at least show that polynomials of such form are irreducible in a much more general layout.

Theorem 7.1. Let

$$f(T) = \frac{T^n}{n!} + c_{n-1}\frac{T^{n-1}}{(n-1)!} + \cdots + c_1 T \pm 1$$

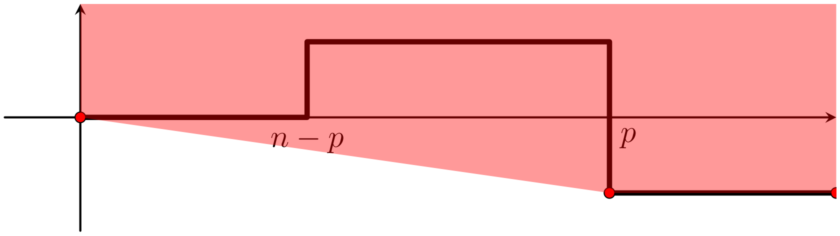
be a polynomial with $c_i \in \mathbb{Z}$. Then, f is irreducible over $\mathbb{Z}$.

Of course, the problem is then to find other conditions that guarantee such a polynomial's Galois group contains A_n, as we did using Jordan's theorem. In general, this is a very difficult problem, and we won't discuss it here; instead, we will present examples of such polynomials where we happen to be able to reuse Jordan's theorem in an alternative setting.

Example. Consider the Laguerre Polynomial

$$L_n(T) = \sum_{k=0}^{n} \binom{n}{k} \frac{(-T)^k}{k!}.$$

By Theorem 7.1, L_n is an irreducible polynomial. Notice that for large n, there is a prime p strictly between $n/2$ and n by Bertrand's Postulate. For this p, the Newton Polygon looks like this:



Hence, we find that there is a slope whose gradient is $-1/p$, and thus we reach the same conclusion as Corollary 4.6 for just this prime p. By Jordan's Symmetric Group theorem, we obtain $\mathrm{Gal}(L_n/\mathbb{Q}) \supseteq A_n$.

It remains to compute the discriminant of L_n to determine whether the Galois group is A_n or S_n. It turns out a similar trick used in Theorem 6.1 yields the discriminant to be

$$\mathrm{disc}(L_n) = \prod_{1\le k\le n} k^{2k-1} = 1^1 \cdot 2^3 \cdot 3^5 \cdot \cdots \cdot n^{2n-1}.$$

Again by Bertrand's Postulate, since there is a prime p strictly between $n/2$ and n, we know that this p must divide the discriminant an odd number of times. It follows that $\mathrm{disc}(L_n)$ is never the square of an integer, and hence $\mathrm{Gal}(L_n/\mathbb{Q}) = S_n$.

References

[1] Wielandt, H. and Bercov, R. (1964) ‘6-13’, in Finite permutation groups. New York: Acad. Press.

[2] P. Erdos, A theorem of Sylvester and Schur, J. London Math. Soc. 9 (1934), 282-288.

[3] Coleman, R. (1987) ‘On the Galois Groups of the Exponential Taylor Polynomials’, L’Enseignement Mathématique, 33, pp. 183-189. doi:http://dx.doi.org/10.5169/seals-87891.

Acknowledgements

I would like to thank Professor Jack Thorne for providing me with the materials to write this mathematical essay.