Superirreducibility of Polynomials, Binomial Coefficient Asymptotics and Stories From My Classroom

by

Lara Du

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mathematics)
in The University of Michigan
2020

Doctoral Committee:

Professor Mattias Jonsson, Co-Chair
Professor Trevor Wooley, Co-Chair
Professor Ratindranath Akhoury
Professor Jeffrey Lagarias
Dr Elaine Lande

Lara Du

hjkl@umich.edu

ORCID id: 0000-0001-6749-4867

ACKNOWLEDGEMENTS

This work was partially supported by NSF grant DMS-1701577 and a Department of Mathematics Fellowship. I was extremely fortunate during my graduate school years to have had excellent mentors that helped me grow both as a mathematician and as a teacher.

Mathematically, I benefitted from regular meetings with Professors Trevor Wooley, Mattias Jonsson and Jeff Lagarias, who have all advised me in some form during my PhD programme. Working with Trevor, I really found my place in the mathematical community: learning how to develop as an academic and have meaningful mathematical conversations while staying true to my own beliefs and roots. From Mattias, I learned to pursue the math that I found interesting, seeking out opportunities that were best for me and finding the conviction I needed to see them through. From Jeff, I learned to problem-solve, to adjust the research question I was asking if appropriate, to take things as slowly as I needed, but to never give up on seeking beautiful mathematical truths.

Throughout graduate school, I've also benefitted enormously from collaboration with Gene Kopp, Jonathan Bober and Dan Fretwell and from the mathematical mentorship of Evangelia Gazaki.

I am also extremely grateful to Elaine Lande, Shylynn Lofton, Talia Thorson, Kevin Carde, Marisa Debowsky, McKenna Shaw and Elizabeth Collins-Wildman for all they've done to help me become a better teacher. Their expertise, belief and guidance helped shape me into the educator that I am today.

Navigating graduate school would not have been possible for me without the tireless work of the Michigan Mathematics Department office staff. Their kindness and support were really instrumental in helping me work through my degree requirements.

Finally, I want to thank my wonderful, hilarious, hardworking and dynamic students that have made up such an amazing part of my professional life for the past six years. I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

ABSTRACT

In the first main section of this thesis, I investigate superirreducible polynomials over fields of positive characteristic and also over $\mathbb{Q}$ and $\mathbb{Z}$. An n-superirreducible polynomial $f(x)$ is an irreducible polynomial that remains irreducible under substitutions $f(g(x))$ for g of degree at most n. I find asymptotics for the number of 2-superirreducible polynomials over finite fields. Over the integers, I give examples of both families of superirreducible polynomials and families of irreducible polynomials which have an obstruction to superirreducibility. The writing and results on finite fields in this section have come from a collaboration with Jonathan Bober, Dan Fretwell, Gene Kopp and Trevor Wooley. The results over $\mathbb{Z}$ and $\mathbb{Q}$ are my own independent work.

In the second section I determine the asymptotic growth of certain arithmetic functions $A(n)$, $B(n)$ and $C(n)$, related to digit sum expansions. I consider these functions as sums over primes p up to n. I obtain unconditional results as well as results with better error terms conditional on the Riemann Hypothesis. The results over primes have come from collaboration with Jeff Lagarias. I also independently solved the analogous problem of summing over all positive integers $b \leq n$.

Finally in the third section, I discuss mathematical education via the lens of interviews and interactions. I consider my role as a teacher through multiple real-life anecdotes and what those stories have taught me. My interviews were conducted with young mathematicians from Bronx, NY that I got the opportunity to talk to as a result of my employment with Bridge to Enter Advanced Mathematics during the summer of 2019. The anecdotes I give are from working with teenaged students from a variety of different cultural, socioeconomical and mathematical backgrounds.

CHAPTER I

Introduction

1.1 Polynomial Related Definitions

Definition 1.1.1. *A polynomial over a ring R is an expression built from constants and variables. A polynomial in the single variable x has form $f(x) = \sum_{k=0}^{n} a_k x^k$, where the* ***coefficients*** *a_k are constant elements in the ring R. The* ***degree*** *of the polynomial is n: the highest power of x occurring with a nonzero coefficient multiplying it.*

Polynomials are ever-present in the world around us. For example, if we throw an object, the arc made by its path can be modelled by a quadratic (degree 2) polynomial. Since polynomials are some of the nicest functions to visualise, approximate and work with, we often approximate other functions using polynomials. In fact the Stone-Weierstrass Theorem, proven in 1937 states that any continuous function on a closed interval can be approximated to any degree of accuracy by a polynomial function. Therefore interest rates, in economics, which are modelled by exponential functions can be approximated by polynomials. So can trigonometric functions, which come up frequently in physics and engineering, for example to describe water ripples in a lake, or a bridge reacting to a hurricane.

A polynomial $f(x)$ is **reducible** over R if it can be written as $f(x) = g(x)h(x)$ with g and h polynomials of degree at least one with coefficients in R. f is said to be **irreducible** otherwise. A value a in the ring R is a **root** of the polynomial f if $f(a) = 0$. This is equivalent to saying that $f(x)$ can be written as $f(x) = (x - a)h(x)$ for some other polynomial $h(x)$ with coefficients in R.

1.2 Finding Roots for Polynomials of Degree at Most 4

Finding roots of polynomials and determining irreducibility are problems that have been studied throughout history. The ancient Greeks discovered that a quadratic polynomial $ax^2 + bx + c$ is reducible over the rationals if and only if $b^2 - 4ac$ is a perfect square. Later on, in the 7th century, the Indian mathematician Brahmagupta discovered that the two roots for a quadratic polynomial are given by the formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

taking the positive square root for one root and the negative square root for the other. If $b^2 - 4ac = 0$ then the quadratic has a repeated root. We can arrive at this formula algebraically by setting $ax^2 + bx + c = 0$ and completing the square.

Just increasing the degree of the polynomial by one already makes finding roots much more difficult. Many generations of mathematicians pondered how to solve cubics, making slow partial progress in understanding cubics that had specific forms. It wasn't until the sixteenth century when a general method was discovered by Tartaglia. There now exist a few different methods of solving the cubic. Here I'll illustrate Tartaglia's method.

Tartaglia's method consists purely of algebraic manipulation. The first step is to complete the cube: by replacing $y + \frac{1}{3}a$ by x in $y^3 + ay^2 + by + c$, we can assume that any cubic equation can be reduced to one of the form $x^3 + px + q$. Tartaglia's method then splits into two cases: $p > 0$ and $p < 0$ (clearly if $p = 0$, we can just directly take cube roots). We'll show how to find the roots when $p > 0$. The solution with $p < 0$ is handled similarly.

Let $f(x) = x^3 + px + q$. Suppose that $p > 0$ and that r is a root of f. We write r in the form $r = \sqrt[3]{u} - \sqrt[3]{v}$ for some u and v. We see that

$$\begin{aligned} r^3 &= (\sqrt[3]{u} - \sqrt[3]{v})^3 \\ &= u - v - 3\sqrt[3]{v}(\sqrt[3]{u})^2 + 3\sqrt[3]{u}(\sqrt[3]{v})^2 \\ &= (u - v) - 3\sqrt[3]{u}\sqrt[3]{v}(\sqrt[3]{u} - \sqrt[3]{v}) \\ &= (u - v) - 3(\sqrt[3]{u}\sqrt[3]{v})r. \end{aligned}$$

Rearranging, we see that we get a cubic in r with no r^2 term:

$$r^3 + 3(\sqrt[3]{u}\sqrt[3]{v})r - (u - v) = 0.$$

Since we know that $f(r) = 0$, we can equate coefficients to get a pair of equations:

$$p = 3(\sqrt[3]{u}\sqrt[3]{v}), \qquad q = -(u - v).$$

By substituting, we can come up with quadratics satisfied by u and v:

$$\begin{aligned} u^2 + qu - \left(\frac{p}{3}\right)^3 &= 0, \\ v^2 - qv - \left(\frac{p}{3}\right)^3 &= 0. \end{aligned}$$

Now using the quadratic formula on each, we can solve for u and v. This allows us to find three (not necessarily distinct) solutions for r, recalling that $r = \sqrt[3]{u} - \sqrt[3]{v}$. We get that

$$r = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} - \sqrt[3]{\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

There's also an algebraic way to solve the quartic. Similar to what we did with the cubic, we can complete the fourth power and assume that our quartic has form

$$x^4 + ax^2 + bx + c$$

with $a, b, c \in \mathbb{Q}$. We look for a root of the form $u + iv \in \mathbb{C}$, where u, v are complex numbers with $v \neq 0$. Substituting in we obtain:

$$(u + iv)^4 + a(u + iv)^2 + b(u + iv) + c.$$

We can expand this out and set both the coefficient of 1 and the coefficient of i equal to 0. Although this is not a necessary condition, it is a sufficient one: both coefficients being zero will be enough for $u + iv$ to be a root. We get the following pair of equations:

$$u^4 - 6u^2v^2 + v^4 + au^2 - av^2 + bu + c = 0,$$
$$4u^3v - 4uv^3 + 2auv + bv = 0.$$

Since we assumed that $v \neq 0$, we can factor it out of the second equation to conclude that $4u^3 - 4uv^2 + 2au + b = 0$. And so we solve for v^2 to get that $v^2 = \frac{4u^3+2au+b}{4u}$. We now substitute this into the first equation and obtain

$$u^4 - 6u^2\frac{4u^3 + 2au + b}{4u} + \left(\frac{4u^3 + 2au + b}{4u}\right)^2 + au^2 - a\frac{4u^3 + 2au + b}{4u} + bu + c = 0. \qquad \{*\}$$

We see that amazingly, the odd degree terms of u in this equation cancel to 0, so that multiplying through $\{*\}$ by u^2 we get a rational degree 6 polynomial with only even degree terms, making it a cubic in u^2. Since we know how to solve cubics, we can obtain solutions for u^2. Taking the square root, we obtain solutions for u. This then gives us solutions for v^2 and so v. Putting all this together, we've solved our quartic.

1.3 Impossibility of the quintic and Galois Theory

We've outlined ways to find roots of all polynomials from degree up until 4 in the previous section. Unfortunately from degree 5 onwards, there aren't such methods. Here we'll discuss the theory behind why this is the case. To begin with, we'll need the following two definitions from group theory.

Definition 1.3.1. *A simple group is one that has no non-trivial normal subgroups.*

Definition 1.3.2. *For a group G, a composition series is a sequence of normal subgroups $\{e\} = G_0 \trianglelefteq G_1 \trianglelefteq \cdots \trianglelefteq G_n \trianglelefteq G$ such that each of the factor groups G_{i+1}/G_i is simple.*

A finite group being solvable means that all of its factor groups G_{i+1}/G_i are abelian. Solvability of a polynomial is determined by solvability of the Galois group of its splitting field. We state this more precisely via the following two theorems and in the subsequent corollary conclude unsolvability of degree n polynomials for $n \geq 5$.

Theorem 1.3.3. *(Jordan-Hölder) Every finite group has a composition series. Any two composition series have the same factor groups (up to rearrangement).*

Theorem 1.3.4. *(Abel-Ruffini) Let F be a field of characteristic 0. A polynomial $f(x) \in F[x]$ is solvable by radicals over F if and only if its splitting field K/F has a solvable Galois group.*

Corollary 1.3.5. *For any $n \geq 5$, there exist degree n polynomials which cannot be solved by radicals.*

To prove Corollary 1.3.5, we'll need to make use of the following three lemmas from Group and Field Theory:

Lemma 1.3.6. *For $n \geq 5$, the alternating group A_n is simple.*

Lemma 1.3.7. *(Dedekind's Theorem) Let $f(x)$ be a monic polynomial of degree n, with integer coefficients. Let $k \leq n$ be an integer. Suppose there exists a prime p such that $f \pmod{p}$ has a an irreducible irreducible factor of degree k. Then the Galois group G for the splitting field of f contains a permutation with a cycle of length k.*

Lemma 1.3.8. *If H is a subgroup of S_n acting transitively on $\{1, 2, \ldots, n\}$ and H contains both an $(n-1)$-cycle and a transposition, then H must be equal to S_n.*

We can now begin the proof of Corollary 1.3.5.

Proof. We use Lemma 1.3.7 to show that there are degree n polynomials with Galois group S_n. First, we take three polynomials $g(x)$ (of degree n, irreducible modulo 2), $h(x)$ (of degree $(n-1)$, irreducible modulo 3) and $k(x)$ (of degree 2, irreducible modulo 5). Next, we choose integers a and b such that a is congruent to $1 \pmod 2$ and 0 modulo 3 and 5 and b is congruent to $1 \pmod 3$ and 0 modulo 2 and 5. By Chinese remainder Theorem, this is possible. Given these choices, we finally build a degree n polynomial $f(x)$ by taking

$$f(x) = ag(x) + bxh(x) + (1-a-b)x^{n-2}k(x).$$

By reducing modulo 2, we see that $f(x) \equiv g(x)$ is irreducible $\pmod 2$ and therefore is irreducible over $\mathbb{Q}$. However by considering the factorisations modulo 3 and 5, we get that

$$\begin{aligned} f(x) &\equiv xh(x) \pmod 3, \\ f(x) &\equiv x^{n-2}k(x) \pmod 5. \end{aligned}$$

By properties of $h(x)$ and $k(x)$, we get that the Galois group for the splitting field of $f(x)$ must contain an $(n-1)$-cycle and a transposition. Since the Galois group acts transitively on the roots of f (because $f(x)$ is irreducible), by Lemma 1.3.8 it must be the entire symmetric group S_n.

The symmetric group S_n has normal subgroup A_n. However by lemma 1.3.6 A_n is simple for $n \geq 5$. This tells us that (up to rearrangement), the composition series for S_n is

$$S_n \trianglerighteq A_n \trianglerighteq \{e\}.$$

Since A_n is nonabelian, this tells us that S_n is not solvable, so that a polynomial with Galois group S_n can't be solvable by radicals. □

Despite this apparent snag in finding roots of polynomials, there is hope of at least identifying when a polynomial is reducible. One such technique is the rational root theorem:

Theorem 1.3.9. *A polynomial $f(x) = c_nx^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0 \in \mathbb{Z}[x]$ has a rational root $\frac{a}{b}$ with $gcd(a, b) = 1$ only if $a|c_0$ and $b|c_n$.*

We note that the rational root is far from sufficient for proving irreducibility: a polynomial that has no roots can still be reducible. For example the quintic polynomial $f(x) = x^5 + x^4 + 3x^3 + 3x^2 + 3x + 1$ has no rational roots: by rational root theorem, any rational root would have form $\frac{a}{b}$ with $gcd(a, b) = 1, a|1$ and $b|1$. Therefore we just need to verify that 1 and -1 are not roots and indeed they're not. However $f(x)$ is reducible over $\mathbb{Q}$: $f(x) = (x^2 + x + 1)(x^3 + 2x + 1)$.

1.4 Methods for showing irreducibility

Now instead of trying to find roots of polynomials, we consider the opposite problem: proving polynomials are irreducible over a given field. One method to show irreducibility for a polynomial $f(x)$ is to find a p prime and a finite field $\mathbb{F}_p$ over which $f(x)$ is irreducible. This then gives us that $f(x)$ is irreducible over $\mathbb{Z}$ (because if it wasn't, a factorisation in $\mathbb{Z}$ would give a factorisation over $\mathbb{F}_p$. This method is effective because it reduces the problem to checking whether $f(x)$ is divisible by finitely many lower degree polynomials over $\mathbb{F}_p$, instead of having to check infinitely many over $\mathbb{Z}$.

Another tool that is used to show irreducibility is Newton Diagrams, which we'll use as a first approach in Subsection 2.2.1 to tackle superirreducibility over $\mathbb{Z}$. The Newton diagram for a polynomial $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ with respect to the prime p is a subset of $\mathbb{R}^2$. It's constructed by plotting degrees along the x axis (from 0 to n) and the maximum power of p dividing a_i for each i, $0 \leq i \leq n$ along the y axis and then taking the lower convex hull of these points. This lower convex hull follows the perimeter of the convex hull from $(0, v_p(a_0))$ to $(n, v_p(a_n))$. We call the set of lines in this lower convex hull the system of vectors for the Newton diagram of $f(x)$ and note that if a line has interior integer points, then it can be expressed as the union of several shorter vectors.

Using Newton Diagrams, we can appeal to the Dumas criterion below:

Theorem 1.4.1. *(Dumas) Let $f = gh$, where f, g and h are polynomials with integer coefficients. Then the system of vectors of the segments for f is the union of the systems of vectors of the segments for g and h in their respective Newton Diagrams (provided p is the same for all the polynomials).*

The Dumas criterion is applicable to show us that multiple families of polynomials are irreducible. It is proved in [23], page 53, Theorem 2.2.1 . We give a few examples of its applications below.

Example 1.4.2. *(Eisenstein's criterion) Let $f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ be a polynomial with integer coefficients. Suppose that there exists a prime p such that $p|a_0, p|a_1, \ldots, p|a_{n-1}$ but $p \not| \, a_n$ and $p^2 \not| \, a_0$. Then $f(x)$ is irreducible over $\mathbb{Z}$.*

Proof. The Newton diagram of f consists simply of one vector with endpoints $(0, 1)$ and $(n, 0)$. This has no interior integer points, so can't be written as the union of multiple vectors. □

Example 1.4.3. *We can apply Eisenstein's criterion (with $p = 3$) on the polynomial $f(x) = x^5 + 18x^4 + 81x^3 + 21x^2 + 6x + 3$ to get the following Newton Diagram:*

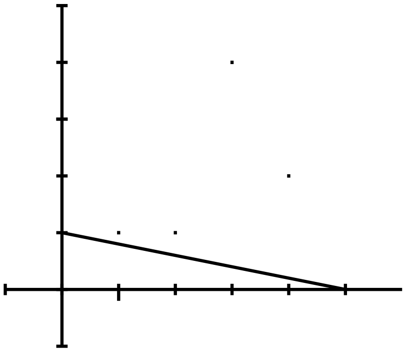

Figure 1.1: Newton Diagram for the polynomial $f(x) = x^5 + 18x^4 + 81x^3 + 21x^2 + 6x + 3$ with $p = 3$

We see that the Newton diagram consists of a single line segment with slope $-\frac{1}{5}$. This line segment contains no interior points, so $f(x)$ must be irreducible.

Example 1.4.4. *According to the Dumas criterion, we can give a condition for irreducibility that's a little more general than Eisenstein's criterion. Namely, that if the Newton diagram for a polynomial f consists of exactly one segment with no interior integer points, then f must be irreducible. For example a polynomial having the following Newton diagram must be irreducible:*

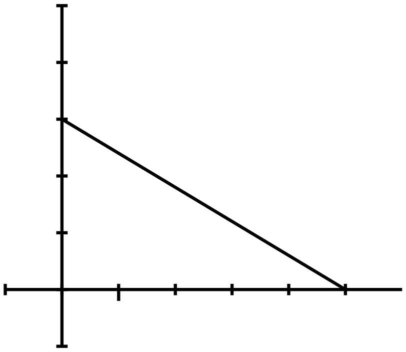

Figure 1.2: Newton Diagram consisting of exactly one segment

We see that this Newton diagram has a single line segment with endpoints at $(0, 3)$ and $(5, 0)$, with slope $-\frac{3}{5}$. Therefore, there are no interior integer points.

Example 1.4.5. *In contrast to Example 1.4.4, a polynomial having the following similar Newton diagram need not be irreducible:*

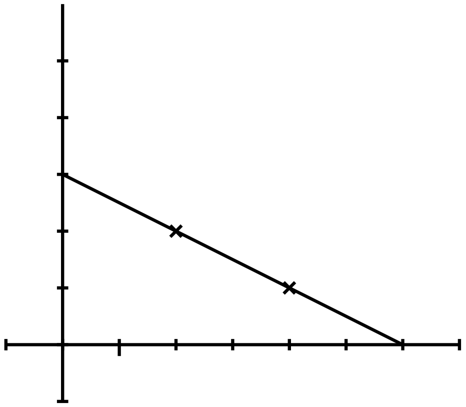

Figure 1.3: Newton Diagram with interior integer points

Although this appears to again to only consist of one line segment, we observe that the line segment has endpoints at $(0,3)$ *and* $(6,0)$*, with slope* $-\frac{1}{2}$*. Therefore this line segment has two interior integer points: at* $(2,2)$ *and at* $(4,1)$*. Therefore, the most that a polynomial with such a Newton diagram could factor is as a product of irreducible quadratics. Indeed if we take* $p = 2$*, we can find such a polynomial. Letting* $f(x) = x^6 + 2x^4 + 4x^2 + 8$*, we get the Newton diagram above and indeed,* $f(x)$ *is the product of three irreducible polynomials:* $f(x) = (x^2+2)(x^2-2x+2)(x^2+2x+2)$.

1.5 Substitutions that preserve irreducibility

Superirreducible polynomials are polynomials that resist factorization under polynomial substitutions. More precisely, for a positive integer k, a **k-superirreducible polynomial** $f(x)$ is not only irreducible, but remains irreducible after any substitution $x = g(t)$ where the degree of $g(t)$ is at most k. The concept of superirreducibility was introduced (although not by name) by Bober, Fretwell, Martin, and Wooley in [4] as a potential limitation on (poly)smoothness of polynomial compositions.

We note that a polynomial $f(x)$ is irreducible over R if and only if for every $a \in R$, $f(x+a)$ is irreducible over R. This is because $f(x) = g(x)h(x)$ if and only if $f(x+$

$a) = g(x+a)h(x+a)$. Therefore the condition of 1-superirreduciblility is equivalent to irreducibility. A 0-superirreducible polynomial will turn out to be precisely a (possibly reducible) polynomial with no roots in the base ring. 2-superirreducibility is when things start to get tricky: if $f(x)$ is an irreducible polynomial and $g(x)$ is a quadratic polynomial, it's not apparent whether $f(g(x))$ will be irreducible. As we increase the degree of g, answering such a question seems to get increasingly difficult.

1.6 Infinite Descent

The method of infinite descent was originally developed by Fermat in the 17th century to show that certain equations have no integer solutions. We appeal to this method in subsections 2.2.3 and 2.2.4 to prove that certain families of polynomials are 2-superirreducible over $\mathbb{Z}$. To apply this method, we show that if a given equation has an integer solution then we can always construct another solution out of it that's strictly smaller in magnitude. This is impossible if we're working within the integers and so the given equation must have no integer solutions at all. We'll illustrate how the method of infinite descent works via a couple of examples.

Example 1.6.1. *We can show via infinite descent that $\sqrt{2}$ is irrational.*

Proof. Suppose that $\sqrt{2}$ is rational, so for some $p, q \in \mathbb{Z}$ it can be written as $\sqrt{2} = \frac{p}{q}$. Squaring both sides yields $2 = \frac{p^2}{q^2}$ so that $p^2 = 2q^2$. Since the right hand side of this equation is an even integer, the left hand side must be too, so that p is divisible by 2. This tells us p^2 is divisible by 4, so $2q^2$ is divisible by 4 and q is even as well. But then we can replace p by the integer $p' := \frac{p}{2}$ and q by the integer $q' := \frac{q}{2}$ and write $\sqrt{2}$ as $\frac{p'}{q'}$. This process can continue indefinitely, which is impossible inside the integers. □

Example 1.6.2. *Using infinite descent, we can prove the case $n = 4$ of Fermat's Last Theorem: that the equation $X^4 + Y^4 = Z^4$ has no nontrivial integer solutions.*

Proof. It suffices to show that the equation $X^4 + Y^4 = Z^2$ has no nontrivial integer solutions. Assume it has solution (x, y, z). Then by dividing through by common factors, we can assume that $(x, y) = 1$, $(x, z) = 1$ and $(y, z) = 1$. We see that exactly one of x and y are odd (for if they're both odd, then $z^2 \equiv 2 \pmod 4$), which is impossible). Assume without loss of generality that x is odd and y is even. Then we can parametrise the Pythagorean triple (x^2, y^2, z) using coprime positive integers m and n by

$$x^2 = m^2 - n^2, \qquad y^2 = 2mn, \qquad z = m^2 + n^2.$$

From this, we receive $x^2 + n^2 = m^2$ giving rise to another Pythagorean triple where x, m, n are all pairwise coprime. We parametrise again, remembering that x is odd, via the coprime positive integers k, l:

$$x = k^2 - l^2, \qquad n = 2kl, \qquad m = k^2 + l^2.$$

Now taking in both parametrisations together, we can write

$$\begin{aligned} y^2 &= 2mn \\ &= 4kl(k^2 + l^2) \end{aligned}$$

so that $\left(\frac{y}{2}\right)^2 = kl(k^2 + l^2)$. Since k and l are coprime, they must each be coprime with $k^2 + l^2$. Since y is even $\left(\frac{y}{2}\right)^2$ is a square integer, so that k, l and $k^2 + l^2$ must all be perfect squares. We can write

$$k = a^2, \qquad l = b^2, \qquad k^2 + l^2 = c^2$$

for a, b, c positive integers and observe that this gives us $c^2 = a^4 + b^4$. We now note that $c^2 = m$ so $c \leq m < z$. Therefore given any positive integer solution (x, y, z) to $X^4 + Y^4 = Z^2$, we can construct another positive integer solution (a, b, c) with $c < z$. Over the positive integers it isn't possible that this process continue indefinitely.

□

1.7 The Theory of Characters

Our analysis of superirreducibles over finite fields in Section 2.1, relies on the technique of characters on unit groups of finite fields. We'll give a brief overview of characters here.

Definition 1.7.1. *For $m \geq 1$, a multiplicative character $\chi : (\mathbb{Z}/m\mathbb{Z})^* \to \mathbb{C}$ is a multiplicative group homomorphism satisfying $\chi(ab) = \chi(a)\chi(b)$ for $(ab, m) = 1$.*

We define the trivial character by $\chi(g) = 1$ for all $g \in (\mathbb{Z}/m\mathbb{Z})^*$. We can extend the notion of multiplicative characters to unit groups of all finite fields:

Definition 1.7.2. *Let G be any finite abelian group. A character of G is a group homomorphism $\chi : G \to \mathbb{C}^*$*

The set of characters forms a finite group with composition given by function multiplication. The identity element is the trivial character. The inverse for each character is given by $\chi^{-1}(x) = (\chi(x))^{-1}$. Since G is finite, we see that all characters must take values that are nth roots of unity in $\mathbb{C}$, for some n. Therefore, the value the inverse of a character takes is exactly its complex conjugate.

We have the following nice orthogonality relations for characters:

Proposition 1.7.3. *(i) Let χ_1, χ_2 be characters on a finite abelian group G. Then, summing over all group elements x, we have*

$$\sum_{x \in G} \chi_1(x)\overline{\chi_2(x)} = \begin{cases} |G| & \text{if } \chi_1 = \chi_2 \\ 0 & \text{otherwise} \end{cases}$$

(ii) If instead we sum over all characters, then we have

$$\sum_{\chi \text{character of } G} \chi(x)\overline{\chi(y)} = \begin{cases} |G| & \text{if } x = y \\ 0 & \text{otherwise} \end{cases}$$

Proof. The proofs of (i) and (ii) are very similar, so we'll just show how to prove (i). Since the set of characters form a group, we see that $\chi_1\overline{\chi_2}$ is a character. If $\chi_1 = \chi_2$, then $\chi_1\overline{\chi_1}$ is the trivial character, so that $\sum_{x\in G}\chi_1(x)\overline{\chi_1(x)} = \sum_{x\in G} 1 = |G|$.

We suppose that $\chi_1 \neq \chi_2$. Then we call $\chi_1\overline{\chi_2} = \psi$ a fixed character of G, so that

$$\sum_{x\in G}\chi_1(x)\overline{\chi_2(x)} = \sum_{x\in G}\psi(x).$$

We can choose an element $g \in G$ such that $\psi(g) \neq 1$ (this is possible because $\chi_1 \neq \chi_2$). Then we have that

$$\begin{aligned}\sum_{x\in G}\psi(x) &= \sum_{x\in G}\psi(g)\psi(g^{-1}x)\\ &= \psi(g)\sum_{x\in G}\psi(g^{-1}x)\\ &= \psi(g)\sum_{x\in G}\psi(x).\end{aligned}$$

Since $\psi(g) \neq 1$, we've established that $\sum_{x\in G}\psi(x) = 0$.

□

1.8 Superirreducibility Results over Finite Fields

With my collaborators Jonathan Bober, Dan Fretwell, Gene Kopp and Trevor Wooley I've counted the number of 2-superirreducible polynomials over finite fields and come up with the following asymptotic formulas. We began with classical Möbius inversion arguments and then used analytic methods from [29] and [27] to bound character sums.

Definition 1.8.1. *We let $s_2(q,d)$ denote the number of 2-superirreducible polynomials of degree d over $\mathbb{F}_q$.*

Theorem 1.8.2. *Let d be odd. Then $s_2(q,d) = 0$.*

Theorem 1.8.3. *Let d be even and q an odd prime power. If we let $d \to \infty$, then we have*

$$s_2(q,d) = \frac{q^d}{d2^q} + O\left(\frac{1}{d}q^{d/2}\right).$$

Theorem 1.8.4. *For $q \gg d^2$, we have that $s_2(q,d) = 0$.*

We also found that our problem of counting 2-superirreducibles over finite fields is equivalent to counting points on affine curves. Specifically, the number of 2-superirreducibles of degree n has main term equal to $\frac{1}{d2^q}$ multiplied by the number of points on

$$C_n := \{(x, y_1, \dots, y_n\} \in \mathbb{A}^{n+1} : \delta y_j^2 = x + u_j \text{ for } 1 \leq j \leq n\},$$

where δ denotes some quadratic non-residue and the u_j are some points in $\mathbb{F}_q$. In Subsection 2.1.5, we show that these curves are complete intersections and we're also able to compute the genus of each C_n to be $1 + 2^{n-1}(n-2)$. We would like to use Weil bounds in Theorem 8.1 from [11]:

Theorem 1.8.5. *Let X be a non-singular complete intersection over $\mathbb{F}_q$ of dimension n. Let b' be the nth Betti number of X. We take $b = b'$ for n odd and $b = b' - 1$ for n even. Then we have that*

$$|\#X(\mathbb{F}_q) - \#\mathbb{P}^n(\mathbb{F}_q)| \leq bq^{\frac{n}{2}}.$$

Knowing the genus allows us to compute the Betti numbers, which would ideally give us an explicit upper bound for the error in approximating the number of points on the complex intersection by the number of points in $\mathbb{P}^n$. Unfortunately these curves are not smooth so Weil bounds are not directly applicable. However we show in Subsection 2.1.5 that there's a workaround to this problem using methods of Ghorpade and Lachaud [14]. This gives us an alternative proof to Theorem 1.8.3 .

1.9 Superirreducibility Results over $\mathbb{Q}$ and $\mathbb{Z}$

As in the proofs of specific cases of Fermat's Last Theorem, I have used infinite descent arguments to show that certain families of Fermat type polynomials must be 2-superirreducible over $\mathbb{Z}$ and $\mathbb{Q}$. Specifically:

Theorem 1.9.1. *For any nonzero $a \in \mathbb{Z}$, all the polynomials in the following families are 2-superirreducible over $\mathbb{Z}$:*

$$\{x^4 + a^4\} \text{ and } \{x^4 + 2a^4\}.$$

The polynomials above all have even degrees and so we must ask ourselves whether there exist any odd degree 2-superirreducibles. Clearly there are none of degree 1, for if $f(x) = x + a$, we may simply set $g(x) = x^2 - a$ and get that $f(g(x)) = x^2$, which factors. Schinzel proved in [24] that there are no 2-superirreducible polynomials of degree 3. In fact, he went further than this and managed to prove that for $d \geq 3$, there are no $(d-1)$-superirreducibles of degree d. Therefore the first odd-degree case that's of interest to us is when $\deg(f) = 5$.

I don't know whether there are any 2-superirreducible polynomials of degree 5. I suspect that there are. For example, by testing over several finite fields, the irreducible polynomial $x^5 + 2x + 1$ remains irreducible under monic quadratic substitutions, making it a good candidate for 2-superirreducibility over $\mathbb{Z}$. In fact I've proven that all of the polynomials in the family$\{x^{2k+1} + 2x + 1\}$ are irreducible and remain irreducible under a large number of quadratic substitutions.

Theorem 1.9.2. *For $k \geq 2$, all the polynomials in the family $\{x^{2k+1} + 2x + 1\}$ are irreducible and stay irreducible under every polynomial substitution of the form $g(x) = ax^2 + b$ with $a, b \in \mathbb{Z}$.*

As well as infinite descent, this proof also required tools from Complex Analysis and some Algebraic Number Theory in analysing the set of discriminants.

Conjecture 1.9.3. *For $k \geq 2$, all the polynomials in the family $\{x^{2k+1} + 2x + 1\}$ are 2-superirreducible.*

In the opposite direction, there are many cases where we can find an obstruction to superirreducibility. In such cases, the polynomial f has a form which we can exploit to find an explicit substitution g that gives a reducible $f(g(x))$. Indeed, this is what Schinzel did to prove that over $\mathbb{Z}$ and for $d \geq 3$, there are no $(d-1)$ superirreducibles of degree d. Drawing inspiration from his arguments, I went on to prove the following list of theorems which give families of polynomials that have an obstruction to superirreducibility.

Theorem 1.9.4. *If $f(x)$ is a polynomial of degree n with rational coefficients, then f is not $(n+k)$-superirreducible over $\mathbb{Z}$, nor $\mathbb{Q}$ for any integer $k \geq 0$.*

Theorem 1.9.5. *For N odd, there are no 2-superirreducibles over $\mathbb{Q}$ of the form $ax^N - b$.*

Theorem 1.9.6. *If $f(x)$ is a polynomial over $\mathbb{Q}$ of degree $2N$, its linear term has nonzero coefficient and all of its other odd degree terms have zero coefficients, then $f(x)$ is not N-superirreducible.*

This last theorem is of particular interest, because it says that the only quartics which could be 2-superirreducible are biquadratic ones. This tells us that 'most' irreducible quartics are not 2-superirreducible. It also tells us that the Galois group of a 2-superirreducible quartic must have order at most 8, which is much smaller than $|S_4|$. This gives us the intuition that somehow Galois groups for superirreducible polynomials should be 'small.' However, the polynomial x^5+2x+1 has Galois group S_5, which is 'big' and it's irreducible under all quadratic substitutions of form ax^2+b.

1.10 Future work with Superirreducibility

1.10.1 Galois Groups

I'd like to gain a better understanding of 2-superirreducibility for quintics. One strategy for doing this involves delving deeper into Fermat type equations and their proofs of nonexistence of solutions and utilising infinite descent arguments combined with Theorem 6.1 from [4].

I'm interested in investigating the connection between the size of Galois groups of polynomials and the probability that the polynomials are superirreducible. Data so far seems to suggest that Galois groups need to be relatively small for 2-superirreducibility. In the quartic case, by Theorem 2.2.8 for $f(x)$ to be 2-superirreducible over the rationals, we need it to be of the form $ax^4 + bx^2 + c$. Such biquadratic quartics are known to have small Galois groups (size at most 8). Something I'm interested in doing for polynomials of degree d is counting the fraction of Galois groups of order at least d and at most $2d$. One approach is do this over number fields of bounded discriminant (studied in [30]) and look for asymptotics as the size of the discriminant goes to infinity. Another is to consider polynomials with coefficients in bounded intervals $[-H, H]$ and study the asymptotics as $H \to \infty$.

1.10.2 Analysis over function fields

Another direction I'd like to take is investigating superirreducibility over function fields. I.e. I'll take polynomials in x with coefficients in another variable t and investigate how substituting in other polynomials of this form affects irreducibility. There has been quite a bit of machinery developed for understanding irreducibility of polynomials over function fields, for example in [1] and [22], that I plan to use as a starting point. Similar to my work over finite fields, I'd like to come up with asymptotics for the number of 2-superirreducibles over the function fields $\mathbb{F}_q(t)$. It's my hope that work over function fields will shed light on higher degree superirreducibility over the fields $\mathbb{F}_q$ and $\mathbb{Q}$.

1.11 Classical Analytic Number Theory Techniques

In Chapter 3, I find asymptotics for certain arithmetic functions related to Farey fractions and to products of binomial coefficients. To do this, I mainly use classical Analytic Number Theory methods. The purpose of this subsection is to give a brief introduction to such methods.

One of the main techniques I'll use in chapter 3 is the Euler-Maclaurin summation formula. The formula, proven in [26], is as follows

Theorem 1.11.1. *For any integer $k \geq 0$ and for any function $f \in C^{k+1}[a,b]$, where $a, b \in \mathbb{Z}$, we have*

$$\sum_{a<n\leq b} f(n) = \int_a^b f(t)\,dt + \sum_{r=0}^{k} \frac{(-1)^{r+1}B_{r+1}}{(r+1)!}(f^{(r)}(b) - f^{(r)}(a)) + \frac{(-1)^k}{(k+1)!}\int_a^b B_{k+1}(t) f^{(k+1)}(t)\,dt.$$

In this formula, the $B_m(t)$ denote the Bernoulli functions, discovered by Jacob Bernoulli in the seventeenth century. To build these, we first take the Bernoulli polynomials $b_m(t)$ obtained via the generating function

$$\sum_{m=0}^{\infty} b_m(t)\frac{y^m}{m!} = \frac{ye^{ty}}{e^y - 1}.$$

The Bernoulli function $B_m(t)$ is the period 1 function which coincides with $b_m(t)$ on $[0, 1)$. Finally, the B_m denote the Bernoulli numbers obtained from the Bernoulli functions by taking $B_m := B_m(0)$. The Bernoulli numbers and functions have many beautiful applications. A nice exposition is given in [5], chapter 5.8. One such application is that they give us a relatively straightforward way to calculate values of the Riemann Zeta function at all positive even integers.

The Euler-Maclaurin summation formula lets us approximate sums with integrals to a high degree of accuracy, allowing us to make sense of the error terms. For example, in

Theorem $I.5$ of [26], the Euler Maclaurin summation formula is used to find an asymptotic formula for the truncated harmonic series:

Theorem 1.11.2. *For $n \geq 1$ we have*

$$\sum_{m \leq n} \frac{1}{m} = \log n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{\theta}{60n^4}$$

where γ is Euler's constant and $\theta = \theta(n) \in [0, 1]$.

Euler's constant $\gamma \approx 0.577$ is conjectured to be transcendental. Since its discovery, γ has been ubiquitous across Analytic Number Theory, showing up in relation to Bernoulli Numbers and the Riemann Zeta function, among other places. A thorough exposition on Euler's constant and its applications is given in [18].

Another technique I use extensively in chapter 3 is the prime number theorem, which says that the number of primes less than n is asymptotic to $\frac{n}{\log n}$. We define the indicator function on primes via

$$\mathbb{I}_P(n) = \begin{cases} 1 & \text{if } n \text{ is prime} \\ 0 & \text{otherwise} \end{cases}$$

and the prime counting function via $\pi(x) := \sum_{n \leq x} \mathbb{I}_P(n)$.

Theorem 1.11.3. *(Prime Number Theorem) There exists a constant $c > 0$ such that*

$$\pi(x) = \frac{x}{\log x} + O(xe^{-c\sqrt{\log x}}).$$

The prime number theorem has a long and rich history. That $\pi(x)$ was about $\frac{x}{\log x}$ was originally conjectured by Gauss in the late eighteenth century. In 1838, Legendre connected the prime counting function to the logarithmic integral $\int_2^x \frac{1}{\log t}\, dt$. The prime number theorem was eventually proven independently by Hadamard [15] and de la Vallée Poussin [8] in the late 20th century, using complex analysis methods.

Finally, in chapter 3, we'll get various expressions containing factorials arising from repeated integration by parts of the logarithmic integral function $\mathrm{li}(x) = \int_2^x \frac{1}{\log t}\, dt$. To bound such expressions, I'll use Stirling's formula which provides an accurate estimate of the size of $n!$:

Theorem 1.11.4. *For all $n \geq 1$, we have*

$$\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} \leq n! \leq e n^{n+\frac{1}{2}} e^{-n}.$$

1.12 Digit Sum Expansions

We define the base b expansion of a positive integer n to be $\sum_{i=0}^{k} a_i b^i$, where each $a_i = a_i(b,n) \leq b-1$ and $a_k \geq 1$ so that $b^k \leq n < b^{k+1}$.

Example 1.12.1. *1. If $n = 93$, the base 2 expansion would be $1 + 2^2 + 2^3 + 2^4 + 2^6$, so that $a_0 = 1, a_1 = 0, a_2 = 1, a_3 = 1, a_4 = 1, a_5 = 0, a_6 = 1$.*

2. If $n = 132$, the base 3 expansion would be $2 * 3 + 2 * 3^2 + 3^3 + 3^4$, so that $a_0 = 0, a_1 = 2, a_2 = 2, a_3 = 3, a_4 = 4$.

Base b expansions are unique for positive integers. We construct it by first taking the highest multiple of the highest power of b which is smaller than n. Say this is $a_k b^k$. We then do the same for $n - a_k b^k$, to get the highest multiple of the highest power of b smaller than $n - a_k b^k$, say $a_j b^j$. We continue this process until it terminates. We observe for all $k \geq 0$ that

$$\begin{aligned} \sum_{i=0}^{k-1} a_i b^i &\leq \sum_{i=0}^{k-1} (b-1) b^i \qquad \text{by definition of the } a_i \\ &= (b-1)\frac{b^k - 1}{b-1} \\ &< b^k. \end{aligned}$$

From the base b expansion, we can build the following functions:

Definition 1.12.2. *1. The sum of digits function $d_b(n)$ (to base b) is given by:*

$$d_b(n) := \sum_{i \geq 0} a_i(b, n).$$

2. The running digit sum function $S_b(n)$ (to base b) is given by:

$$S_b(n) := \sum_{j-0}^{n-1} d_b(j).$$

$S_b(n)$ gives an 'average' sum of digits for integers between 0 and n.

The functions in definition 1.12.2 have been extensively studied by number theorists in the past century. In 1940, Bush proved in [6] (with slightly different notation) that as $n \to \infty$, the quantity $S_b(n)$ is asymptotically equivalent to $\frac{n(b-1)\log n}{2 \log b}$. Over time, other people found better and better error terms for this asymptotic. In 1975, in [10], Delange found a closed form expression for $S_b(n)$:

Theorem 1.12.3. *(Delange, 1975) For every integer $b \geq 2$, we have*

$$S_b(n) = \left(\frac{b-1}{2}\right)\left(n \log_b n + \tilde{f}_b(\log_b n)\right)$$

where the $\tilde{f}_b(x)$ are non-positive continuous, nowhere differentiable functions that have Fourier series expansions

$$\tilde{f}_b(x) = \sum_{k \in \mathbb{Z}} \tilde{c}_b(k) e^{2\pi i k x}$$

whose Fourier coefficients are

$$\begin{aligned} \tilde{c}_b(k) &= -\frac{1}{k\pi i}\left(1 + \frac{2k\pi i}{\log b}\right)^{-1} \zeta\left(\frac{2k\pi i}{\log b}\right) \qquad \textit{for } k \neq 0 \\ \tilde{c}_b(0) &= \frac{1}{\log b}(\log(2\pi) - 1) - \left(\frac{b+1}{2(b-1)}\right). \end{aligned}$$

1.13 Binomial Coefficients and Farey Fractions–Past Work

Definition 1.13.1. *The Farey sequence of order n is a finite set of fractions*

$$\left\{\frac{k}{m} : \frac{k}{m} \in (0, 1], m \leq n\right\}.$$

The distribution of Farey fractions approaches the uniform distribution on $[0,1]$ as n tends to ∞.

The complete products of binomial coefficients of order n are the integers

$$\overline{G}_n := \prod_{k=0}^{n} \binom{n}{k}.$$

Lagarias and Mehta showed that $\overline{G}_n$ is equal to the reciprocal of the product of all the elements in the Farey sequence of order n. We can verify this explicitly for small values of n.

Example 1.13.2. *If $n = 5$ we have that*

$\overline{G}_n = \prod_{k=0}^{5} \binom{5}{k}$*. The reciprocal of the product of all the elements of the Farey sequence of order 5 is*

$$\frac{1}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)\left(\frac{1}{4}\right)\left(\frac{2}{4}\right)\left(\frac{3}{4}\right)\left(\frac{1}{5}\right)\left(\frac{2}{5}\right)\left(\frac{3}{5}\right)\left(\frac{4}{5}\right)}.$$

Both of these expressions simplify to 2500.

There's a connection between Farey fractions, products of binomial coefficients and the sum of digit and running digit sum functions $d_b(n)$ and $S_b(n)$. This connection was found by Lagarias and Mehta in their paper [19], Theorem 5.1:

Theorem 1.13.3. *(Lagarias and Mehta, 2014) Let the prime p be fixed. Then for all $n \geq 1$,*

$$\nu_p(\overline{G}_n) := ord_p(\overline{G}_n) = \frac{1}{p-1}\left(2S_p(n) - (n-1)d_p(n)\right).$$

1.14 Binomial Coefficients and Farey Fractions–Main Results

Using Theorem 1.13.3, my collaborator Jeff Lagarias and I found asymptotics for other arithmetic functions related to Farey fractions and products of binomial coefficients. Namely, we took logarithms in the expression for $\nu_p(\overline{G}_n)$, so that

$$\log \overline{G}_n = A(n) - B(n)$$

where

$$A(n) = \sum_{p \leq n} \frac{2}{p-1} S_p(n) \log p$$
$$B(n) = \sum_{p \leq n} \frac{n-1}{p-1} d_p(n) \log p$$

and analysed both pieces to understand the magnitude of the contribution from each of $A(n)$ and $B(n)$ to $G(n)$. We found asymptotics for these functions as $n \to \infty$, considering what the error term would be unconditionally and assuming the Riemann Hypothesis. An interesting note is that Euler's constant γ appears in the main terms of our asymptotics for both functions.

The theorems we proved are:

Theorem 1.14.1. *Let* $A(n) = \sum_{p \leq n} \frac{2}{p-1} S_p(n) \log p$

1. *There is a constant* $c > 0$*, such that for* $n \geq 4$

$$A(n) = \left(\frac{3}{2} - \gamma\right) n^2 + O\left(n^2 \exp(-c\sqrt{\log n})\right)$$

where γ *denotes Euler's constant.*

2. *Assuming the Riemann Hypothesis, for* $n \geq 4$ *and any* $\epsilon > 0$

$$A(n) = \left(\frac{3}{2} - \gamma\right) n^2 + O\left(n^{7/4} (\log n)^2\right).$$

Theorem 1.14.2. *Let* $B(n) = \sum_{p \leq n} \frac{n-1}{p-1} d_p(n) \log p$

1. *There is a constant* $c > 0$*, such that for* $n \geq 4$

$$B(n) = (1 - \gamma) n^2 + O\left(n^2 \exp(-c\sqrt{\log n})\right)$$

where γ *denotes Euler's constant.*

2. *Assuming the Riemann Hypothesis, for* $n \geq 4$ *and any* $\epsilon > 0$

$$B(n) = (1 - \gamma) n^2 + O\left(n^{7/4} (\log n)^2\right).$$

Theorem 1.14.3. *Let* $C(n) = \sum_{p \le n} \tilde{f}_p(\log_p n) \log p$

1. *For all* $n \geq 4$,

$$C(n) = \left(\frac{1}{2} - \gamma\right) n - \sum_{k=1}^{m} k! \frac{n}{(\log n)^k} + O\left(2^{m+1}(m+1)! \frac{n}{(\log n)^{m+1}}\right).$$

2. *Assuming the Riemann Hypothesis, for* $n \geq 4$ *and any* $\epsilon > 0$

$$C(n) = \left(\frac{1}{2} - \gamma\right) n - \sum_{k=1}^{\lfloor \frac{1}{2} \log n \rfloor} k! \frac{n}{(\log n)^k} + O\left(n^{3/4} (\log n)^2\right).$$

I've also independently extended the ideas in the proofs of these theorems and proven asymptotics for analogous functions $A'(n)$, $B'(n)$ and $C'(n)$. For these new functions, instead of considering base p radix expansions over the primes, I considered expansions over all integers b with $1 < b \leq n$ and came up with the following theorems:

Theorem 1.14.4. *Let* $A'(n) = \sum_{2 \le b \le n} \frac{2 \log b}{b-1} S_b(n)$ *For all* $n \geq 4$ *we have*

$$A'(n) = \left(\frac{3}{2} - \gamma\right) n^2 \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right) n^2 + O(n^{3/2} \log n)$$

where γ *denotes Euler's constant and* α *denotes the first Stieltjes constant given by*

$$\alpha := \lim_{n \to \infty} \left(\sum_{k=1}^{n} \frac{\log k}{k} - \int_1^n \frac{\log t}{t} \, dt \right).$$

Theorem 1.14.5. *Let* $B'(n) = \sum_{2 \le b \le n} \frac{n-1}{b-1} d_b(n) \log b$. *For all* $n \geq 4$ *we have*

$$B'(n) = (1 - \gamma) n^2 \log n + (\gamma + \alpha - 1) n^2 + O(n^{3/2} \log n).$$

Theorem 1.14.6. *Let* $C'(n) = \sum_{2 \le b \le n} \frac{2 \log b}{b-1} f_b(\log_b(n))$ *For all* $n \geq 4$ *we have*

$$C'(n) = \left(\frac{1}{2} - \gamma\right) n \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right) n + O(\sqrt{n} \log n).$$

Summing over all integers, the unconditional results for $A'(n)$, $B'(n)$ and $C'(n)$ have error terms which save a power of n. This is a big improvement from the unconditional results for $A(n)$, $B(n)$ and $C(n)$, whose error terms were only able to save a logarithm.

In these Theorems, Euler's constant γ has value approximately equal to 0.5772 and the first Stieltjes constant α has numerical value approximately equal to -0.0728.

1.15 Stories from my classroom

When I've thought of my thesis, I've always wanted it to be a representative collection of my graduate school work. As well as being a researcher, being a teacher has been an integral part of my graduate school experience, both at the University of Michigan where I've been blessed to have been the instructor of record for multiple calculus classes and also through the various non-university math outreach opportunities I've had. Accordingly, it's impossible for me to catalogue the past six years without reflecting on the challenges, unique learning opportunities and indescribable happiness my students have given me. And so, I devote one chapter of my thesis to reflect, to tell their stories and to let their voices be heard through direct interviews and through my indirect recollections.

I've written this chapter using the social science qualitative technique of autoethnography [13], which first appeared in the 1970's works of American anthropologists. I don't make generalisations to education as a whole, cite figures to try to prove hypotheses or give techniques for other educators to implement. I just reflect on and analyse my experiences and interactions, in an effort to extract more meaningful depth from my memories.

One of my literary inspirations is Mel Levine's 'A Mind at a Time' [20], where a lot of emphasis is put on case studies, describing each student as an individual learner and techniques the author used to work with that student. In his book, Levine writes at length about different types of learning disabilities. However, instead of giving instructions for how a generic student with a given disability should be handled, he focuses on understanding each student as their own person and techniques for helping them succeed in their classes. In my stories, I try to take a similar approach, being mindful to not identify 'trends', or give guidelines for how to work with 'similar types' of students. From the educator's perspective, I also avoid saying that any given approach I used with one specific student should be used by all teachers, or even all teachers ‘similar to me.’

My teaching chapter is loosely divided into two main sections. In the first, I share some of the challenges faced in math by three college students (all three are people of colour, from Bronx NY) that I interviewed, linking them to the challenges I’ve faced. I also talk about the role that positive mathematical mentorship has played in our lives. In the second section, I share a series of stories about my interactions with students, to the best of my ability to recall them. Within this section, I showcase some of my most precious interactions with students and reflect on what they’ve taught me.

CHAPTER II

Superirreducibility of Polynomials

2.1 Superirreducibility over finite fields

2.1.1 Preliminary definitions and observations

Superirreducible polynomials are polynomials that resist factorization under polynomial substitutions. More precisely, for a positive integer k, a k-superirreducible polynomial $f(x)$ is not only irreducible, but remains irreducible after any substitution $x = g(t)$ where the degree of $g(t)$ is at most k.

The condition of 1-superirreduciblility is equivalent to irreducibility. A 0-superirreducible polynomial will turn out to be precisely a (possibly reducible) polynomial with no roots in the base ring.

The concept of superirreducibility was introduced (although not by name) by Bober, Fretwell, Martin, and Wooley [4] as a potential limitation on (poly)smoothness of polynomial compositions.

Definition 2.1.1. *Let R be a commutative domain with unity, and let F be its field of fractions. A polynomial $f(x) \in R[x]$ is k-superirreducible over R if, for all $g(t) \in R[t]$ of degree $\deg g \leq k$, the composition $f(g(t))$ is irreducible in $F[t]$.*

We are interested in counting superirreducible polynomials over the finite field $\mathbb{F}_q$.

Definition 2.1.2. *Define $S_k(q, d)$ to be the set of monic k-superirredible polynomials of* *degree d over $\mathbb{F}_q$. Denote the size of $S_k(q,d)$ by $s_k(q,d) = \#S_k(q,d)$.*

2.1.2 Elementary cases

Proposition 2.1.3. *Let p be a prime number. Then for all $\ell \geq 1$, we have $s_p(p^\ell, d) = 0$.*

Proof. We let $f(x) \in \mathbb{F}_{p^\ell}[x]$ be a monic irreducible polynomial of degree d, and write

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

We note that $a_j = a_j^{p^\ell}$. By the linearity of the Frobenius automorphism $\alpha \mapsto \alpha^p$,

$$f(t^p) = \sum_{j=0}^{d} a_j^{p^\ell} t^{pj} = \left(\sum_{j=0}^{d} a_j^{p^{\ell-1}} t^j \right)^p$$

Therefore, no such $f(x)$ is p-superirreducible. Hence $s_p(p^\ell, d) = 0$. □

Choosing $p = 2$ in the above proposition gives us that there are no 2-superirreducibles over finite fields of characteristic 2. So in terms of understanding 2-superirreducibility over finite fields, we can move to consider the case when q is odd.

Proposition 2.1.4. *Let $f(x) \in \mathbb{F}_q[x]$ be a monic irreducible polynomial of degree d, and let α be a root of $f(x)$ in $\mathbb{F}_{q^d}$. If $g(t)$ is a quadratic polynomial, the composition $f(g(t))$ factors in $\mathbb{F}_q[t]$ if and only if $g(t) = \alpha$ has a solution in $\mathbb{F}_{q^d}[t]$.*

Proof. ($\implies$) Suppose that $f(g(x))$ splits in $\mathbb{F}_q[x]$. We note that for this to happen, it must split as the product of two irreducible polynomials, each of degree d. If this were not the case then we'd have a factor $h(x) \in \mathbb{F}_q[x]$ of $f(g(x))$ with $deg(h) = e < d$. Let β be a root of h. We know that $\mathbb{F}_q(g(\beta)) \subset \mathbb{F}_q(\beta)$ and $[\mathbb{F}_q(\beta) : \mathbb{F}_q] < d$. Therefore the minimal polynomial of $g(\beta)$ has degree strictly smaller than d. However since f is irreducible, it is the minimal polynomial of $g(\beta)$, giving a contradiction.

From now we may suppose that $f(g(x)) = h_1(x)h_2(x)$ with each h_j irreducible of degree d. We look at the set of roots of both of the h_j's: $\{\beta_1, \beta_2, \ldots \beta_{2d}\}$. Then for each i,

$[\mathbb{F}_q(\beta_i) : \mathbb{F}_q] = d$. Since $g(\beta_i)$ is a root of f for all i, we have that $[\mathbb{F}_q(g(\beta_i)) : \mathbb{F}_q] = d$, so that $\mathbb{F}_q(g(\beta_i)) = \mathbb{F}_q(\beta_i)$ and β_i must be inside the splitting field for f.

We show that all the β_i's are distinct. If h_j has a root β_i, then since h_j is irreducible, it's the minimal polynomial for β_i. If the root is repeated then β_i would also be a root of h_j'. This tells us that h_j divides h_j', which is impossible. Next we consider a β_i which is a singular root of both h_1 and h_2. The Galois group of the splitting field for $f(g(x))$ acts transitively on β_i. For the action to be well defined we need the set of roots of h_1 and h_2 to be exactly the same, so the only case left to consider is $f(g(x)) = c(h(x))^2$. In this case, every root of $f(g(x))$ is a repeated root and so must be a root of $g'(x)f'(g(x))$. For each such β_i we have that $g(\beta_i)$ is a root of f. We know that $g'(\beta_i) \neq 0$, because otherwise h (the minimal polynomial of β_i) would divide g'. Therefore $f'(g(\beta_i)) = 0$ and so $g(\beta_i)$ is a repeated root for f, contradicting the fact that f is irreducible.

Next we see that for all elements β_i, we have that $g(\beta_i)$ is a root of f. Since g is quadratic, each value taken by g can only be obtained from at most two distinct inputs, i.e. from plugging in at most two distinct β_i's. Since there are $2d$ of them and f has only d roots, by pigeonhole principle, α must be equal to $g(\beta_i)$ for some i.

$(\Longleftarrow)$ Suppose $g(t) = \alpha$ has a solution in $\mathbb{F}_{q^d}[t]$. This means $f(g(t))$ has a zero in $\mathbb{F}_{q^d}$. Let's call this zero β, so that $\mathbb{F}_q(\beta) \subset \mathbb{F}_q(\alpha) = \mathbb{F}_{q^d}$. However, $\mathbb{F}_q(\alpha) = \mathbb{F}_q(g(\beta))$ so the reverse inclusion $\mathbb{F}_q(\alpha) \subset \mathbb{F}_q(\beta)$ holds. This tells us that $\mathbb{F}_q(\alpha) = \mathbb{F}_q(\beta)$, so that $\mathbb{F}_q(\beta)$ is a degree d extension over $\mathbb{F}_q$, generated by $\{1, \beta, \beta^2, \ldots, \beta^{d-1}\}$. We can then write β^d as an $\mathbb{F}_q$-linear combination of elements in this set, telling us the minimal polynomial of β over $\mathbb{F}_q$ has degree at most d.

Since β is a root of $f(g(t))$, the minimal polynomial of β must divide $f(g(t))$. Since $f(g(t))$ is of degree $2d$ and the minimal polynomial of degree at most d we get a proper factorisation of $f(g(t))$.

□

Remark 2.1.5. *An analogous line of reasoning establishes Proposition 2.1.4 in the case that $\mathbb{F}_q$ is replaced by $\mathbb{Q}$*

Proposition 2.1.6. *If q is odd and d is odd, then $s_2(q,d) = 0$.*

Proof. Let $f(x) \in \mathbb{F}_q[x]$ be a monic irreducible polynomial of degree d, and let α be a root of $f(x)$ in $\mathbb{F}_{q^d}$. By Proposition 2.1.4, if $g(t)$ is a quadratic polynomial, the composition $f(g(t))$ factors in $\mathbb{F}_q[t]$ if and only if $g(t) = \alpha$ has a solution in $\mathbb{F}_{q^d}$.

If α is a square in $\mathbb{F}_{q^d}$, then taking $g(t) = t^2$ gives a solution to $g(t) = \alpha$ lying inside $\mathbb{F}_{q^d}$ and shows that $f(x)$ is not 2-superirreducible.

If α is not a square in $\mathbb{F}_{q^d}$, choose any nonsquare $b \in \mathbb{F}_q$. Because d is odd, b is also a nonsquare in $\mathbb{F}_{q^d}$. This is because if $z^2 = b$ was solvable over $\mathbb{F}_{q^d}$ then $\mathbb{F}_q(\sqrt{b})$ would be a subfield of $\mathbb{F}_{q^d}$. This is impossible by multiplicativity of field extension degrees. Thus, $b^{-1}\alpha$ is a square in $\mathbb{F}_{q^d}$; that is, $bt^2 = \alpha$ has a solution over $\mathbb{F}_{q^d}$. Taking $g(t) = bt^2$ shows that $f(x)$ is not 2-superirreducible.

Thus, there are no 2-superirreducible polynomials of degree d in $\mathbb{F}_q[x]$ and so

$$s_2(q,d) = 0$$

□

2.1.3 Heuristic count for the number of 2-superirreducibles

By a classic argument using Möbius inversion, the number of monic irreducible polynomials of degree d in $\mathbb{F}_q[x]$ is about q^d/d. For completeness, we present the argument.

We know that $x^{q^d} - x$ is equal to the product of all the irreducibles dividing it. Therefore, by looking at degrees, we have

$$q^d = \sum_{e|d} e \cdot (\#\text{monic irreducibles of degree e})$$

where the sum is taken over all monic irreducibles dividing $x^{q^d} - x$. Applying Möbius inversion, we arrive at

$$d \cdot (\#\text{monic irreducibles of degree d}) = \sum_{e|d} \mu\left(\frac{d}{e}\right) q^e$$

where we note that the term on the right hand side with the largest power of q is q^d. Dividing both sides by d, we get the desired approximation.

Next, we'll use this estimate for the number of irreducible polynomials to approximate the number of 2-superirreducibles of degree d over $\mathbb{F}_q$.

Let $f(x) \in \mathbb{F}_q[x]$ be a monic irreducible polynomial of degree d, and let $\alpha \in \mathbb{F}_{q^d}$ be a root of $f(x)$. For a quadratic polynomial $g(t) \in \mathbb{F}_q[t]$, by Proposition 2.1.4 the composition $f(g(t))$ factors if and only if $g(t) = \alpha$ has a solution.

The conditions '$g(t) = \alpha$ has a solution' are not all independent for different $g(t)$. In particular, $g(t) = \alpha$ has a solution in $\mathbb{F}_{q^d}$ if and only if $g(t+v) = \alpha$ has a solution, for any $v \in \mathbb{F}_q$; after such a substitution, we may assume $g(t)$ is of the form $g(t) = at^2 + c$ for $a, c \in \mathbb{F}_q$. Because d is even, a is a square in $\mathbb{F}_{q^d}$, so it actually suffices to consider only the polynomials $g(t) = t^2 + c$.

To recap, $f(x)$ is 2-superirreducible if and only if $f(t^2+c)$ is irreducible for all $c \in \mathbb{F}_q$, that is, if and only if for any c, the polynomial $t^2 + c = \alpha$ has no solution. To rephrase, we are asking that all the additive shifts $\alpha - c$ of α by elements in the base field are not squares in $\mathbb{F}_{q^d}$. Half the elements of $\mathbb{F}_{q^d}$ are squares, so, heuristically, this happens for each c 'with probability $\frac{1}{2}$'. There are q choices for c and so by treating '$\alpha - c$ is a square' as 'independent events', we predict that $f(x)$ is 2-superirreducible with probability $1/2^q$.

After multiplying by the number of choices of irreducible $f(x)$, our heuristic predicts

$$s_2(q, d) \approx \frac{q^d}{d2^q}.$$

We'll show in the next section that this heuristic gives the correct answer as $d \to \infty$.

2.1.4 The large d limit

The asymptotic formula predicted by the heuristic will follow in the large d limit from Weil's theorem resolving the Riemann hypothesis for curves—specifically, the Weil bound for higher autocorrelations of the quadratic character.

Theorem 2.1.7. *Fix q an odd prime power. As $d \to \infty$,*

$$s_2(q,d) = \frac{q^d}{d2^q} + O\left(\frac{1}{d}q^{d/2}\right).$$

The proof of this asymptotic formula that we give in this section is a more rigorous version of the heuristic argument given in Section 2.1.3.

Definition 2.1.8. *Let χ_q be the nontrivial quadratic character $\chi_q : \mathbb{F}_q^\times \to \{1,-1\}$, where*

$$\chi_q(n) = \begin{cases} 1 & \text{if } n \text{ is a nonzero square in } \mathbb{F}_q \\ -1 & \text{if } n \text{ is not a square in } \mathbb{F}_q \end{cases}$$

χ_q is extended to all of $\mathbb{F}_q$ by setting $\chi_q(0) = 0$.

Definition 2.1.9. *We define the order n autocorrelations of χ_q with offsets $u_1, \ldots, u_n \in \mathbb{F}_q$ to be the numbers*

$$a_q(u_1, \ldots, u_n) := \sum_{\alpha \in \mathbb{F}_q} \chi_q(\alpha + u_1) \cdots \chi_q(\alpha + u_n).$$

Note that $a_q(u_1, \ldots, u_n) \in \mathbb{Z}$. As the order of the u_j doesn't matter, we will also use the notation $a_q(U) = a_q(u_1, \ldots, u_n)$ where $U = \{u_1, \ldots, u_n\}$.

Proposition 2.1.10. *Let q be an odd prime power, and let d be even. The number $s_2(q,d)$ of monic 2-superirreducible polynomials of degree d over $\mathbb{F}_q$ has the following expression in terms of autocorrelations of the quadratic character χ_{q^d}.*

$$s_2(q,d) = \frac{1}{d2^q} \sum_{\substack{d' \mid d \\ 2 \nmid \frac{d}{d'}}} \mu\left(\frac{d}{d'}\right) \left(q^{d'} + \sum_{\{\} \neq U \subseteq \mathbb{F}_q} (-1)^{|U|} a_{q^{d'}}(U)\right).$$

Proof. Consider a monic irreducible polynomial $f(x) \in \mathbb{F}_q[x]$ of degree d. Let α be a root of $f(x)$ in $\mathbb{F}_{q^d}[x]$. The polynomial $f(x)$ is 2-superirreducible if and only if $f(t^2-u)$ is irreducible in $\mathbb{F}_q[t]$ for every $u \in \mathbb{F}_q$. This happens when $t^2 - u = \alpha$ has no solutions in $\mathbb{F}_{q^d}$; in other words, when $\chi_{q^d}(\alpha+u) = -1$ for all $u \in \mathbb{F}_q$.

The indicator function for 2-superirreducibility can thus be expressed algebraically:

$$\prod_{u\in\mathbb{F}_q} \frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right) = \begin{cases} 1, & \text{if } f \text{ is 2-superirreducible,} \\ 0, & \text{otherwise.} \end{cases}$$

Each irreducible polynomial of degree d has d roots $\alpha_1, \ldots, \alpha_d \in \mathbb{F}_{q^d}$, and the α_i that arise are exactly the elements of $\mathbb{F}_{q^d}$ not living in any proper subfield. This is because any proper subfield of $\mathbb{F}_{q^d}$ is a normal extension of $\mathbb{F}_q$, so if a polynomial had a root in such a subfield, it would split completely in that subfield. Counting over all irreducible polynomials of degree d, we get all the α's in $\mathbb{F}_{q^d}$ not living in any proper subfield. Thus, we divide by d to avoid overcounting and obtain the following formula for $s_2(q,d)$:

$$s_2(q,d) = \frac{1}{d} \sum_{\substack{\alpha\in\mathbb{F}_{q^d} \\ \alpha\notin\mathbb{F}_{q^{d'}} \\ \text{for } d'|d,\, d'\neq d}} \prod_{u\in\mathbb{F}_q} \frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right).$$

Using Möbius inversion, we can remove the condition that α does not live in a proper subfield of $\mathbb{F}_{q^d}$ in the following way.

$$\begin{aligned} s_2(q,d) &= \frac{1}{d} \sum_{\substack{\alpha\in\mathbb{F}_{q^d} \\ \alpha\notin\mathbb{F}_{q^{d'}} \\ \text{for } d'|d,\, d'\neq d}} \prod_{u\in\mathbb{F}_q} \frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right) \\ &= \frac{1}{d}\sum_{\alpha\in\mathbb{F}_{q^d}} \prod_{u\in\mathbb{F}_q} \frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right) - \frac{1}{d}\sum_{\substack{d'|d \\ d'\neq d}} \sum_{\substack{\alpha\in\mathbb{F}_{q^{d'}} \\ \alpha\notin\mathbb{F}_{q^e} \\ \text{for } e|d',\, e\neq d'}} \prod_{u\in\mathbb{F}_q} \frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right) \\ &= \frac{1}{d}\sum_{\alpha\in\mathbb{F}_{q^d}} \prod_{u\in\mathbb{F}_q} \frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right) - \frac{1}{d}\sum_{\substack{d'|d \\ d'\neq d}} d' s_2(q,d'). \end{aligned}$$

Therefore we have

$$\frac{1}{d}\sum_{\alpha\in\mathbb{F}_{q^d}}\prod_{u\in\mathbb{F}_q}\frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right)=\frac{1}{d}\sum_{d'|d}d's_2(q,d').$$

Applying Möbius inversion now we get

$$\begin{aligned}s_2(q,d)&=\frac{1}{d}\sum_{d'|d}\mu\left(\frac{d}{d'}\right)\sum_{\alpha\in\mathbb{F}_{q^{d'}}}\prod_{u\in\mathbb{F}_q}\frac{1}{2}\left(1-\chi_{q^d}(\alpha+u)\right)\\&=\frac{1}{d2^q}\sum_{d'|d}\mu\left(\frac{d}{d'}\right)\sum_{\alpha\in\mathbb{F}_{q^{d'}}}\prod_{u\in\mathbb{F}_q}\left(1-\chi_{q^d}(\alpha+u)\right).\end{aligned}$$

The quadratic character χ_{q^d} on $\mathbb{F}_{q^d}$ restricts to the trivial character on $\mathbb{F}_{q^{d'}}$ when $\frac{d}{d'}$ is even and to $\chi_{q^{d'}}$ when $\frac{d}{d'}$ is odd. The product over u vanishes in the former case. Thus,

$$\begin{aligned}s_2(q,d)&=\frac{1}{d2^q}\sum_{\substack{d'|d\\2\nmid\frac{d}{d'}}}\mu\left(\frac{d}{d'}\right)\sum_{\alpha\in\mathbb{F}_{q^{d'}}}\prod_{u\in\mathbb{F}_q}\left(1-\chi_{q^{d'}}(\alpha+u)\right)\\&=\frac{1}{d2^q}\sum_{\substack{d'|d\\2\nmid\frac{d}{d'}}}\mu\left(\frac{d}{d'}\right)\left(q^{d'}+\sum_{\{\}\neq U\subseteq\mathbb{F}_q}(-1)^{|U|}a_{q^{d'}}(U)\right).\end{aligned}$$

This is the desired formula for $s_2(q,d)$. □

When $n=1$, we have $a_q(u)=0$. When $n=2$ and $u_1\neq u_2$, instead $a_q(u_1,u_2)$ is a quadratic Jacobi sum, so $a_q(u_1,u_2)=\pm1$. (See Chapter 8 of [16] for the elementary theory of Jacobi sums.) The higher-order correlations become more complicated, but they can be bounded.

[There is actually no need to fix the base field in the following argument. The u_j could be taken to live in a field extension of $\mathbb{F}_q$, and we could send $q\to\infty$ rather than $d\to\infty$.]

Lemma 2.1.11. *Fix q, and fix $u_1,\ldots,u_n\in\mathbb{F}_q$. The following asymptotic holds as $d\to\infty$:*

$$a_{q^d}(u_1,\ldots,u_n)\leq(n-1)q^{d/2}.$$

Proof. We consider the hyperelliptic curve C with affine equation

$$y^2 = (x+u_1)\cdots(x+u_n).$$

The number of points on the affine locus of C is

$$\sum_{\alpha\in\mathbb{F}_{q^d}} \big(1+\chi_{q^d}\left((\alpha+u_1)\cdots(\alpha+u_n)\right)\big) = q^d + a_{q^d}(u_1,\dots,u_n)$$

where on each occasion $\chi_{q^d}\left((\alpha+u_1)\cdots(\alpha+u_n)\right)=1$, we add 2 to the sum, accounting for the two possible square roots of each y^2 giving two distinct points.

The curve C has genus at most $g=\lfloor\frac{n-1}{2}\rfloor$. This is because $y^2=(x+u_1)\cdots(x+u_n)$, where the right hand side is a degree n polynomial in x and the genus of a hyperelliptic curve is completely determined by the degree ([3], Proposition 12.9). Thus, by the Weil bound [11],

$$\big|a_{q^d}(u_1,\dots,u_n)\big| \le 2gq^{d/2} \le (n-1)q^{d/2}.$$

□

Now we complete the proof of Theorem 2.1.7.

Proof. By Proposition 2.1.10,

$$s_2(q,d) = \frac{1}{d2^q}\sum_{\substack{d'|d\\ 2\nmid\frac{d}{d'}}} \mu\left(\frac{d}{d'}\right)\left(q^{d'} + \sum_{\{\}\neq U\subseteq\mathbb{F}_q}(-1)^{|U|}a_{q^{d'}}(U)\right).$$

By Lemma 2.1.11, there exists a constant C, dependent on q but independent of d and d', such that

$$\big|a_{q^{d'}}(U)\big| \le Cq^{d'/2}$$

for all nonempty subsets U of $\mathbb{F}_q$. Thus,

$$\begin{aligned}
\left| s_2(q,d) - \frac{q^d}{d2^q} \right| &\leq \frac{1}{d2^q} \sum_{\substack{d'|d \\ 2\nmid \frac{d}{d'} \\ d' \neq d}} q^{d'} + \frac{1}{d2^q} \sum_{\substack{d'|d \\ 2\nmid \frac{d}{d'}}} (2^q - 1) C q^{d'/2} \\
&\leq \frac{1}{d2^q} \sum_{d' \leq d/2} q^{d'} + \frac{1}{d2^q} \sum_{d' \leq d} 2^q C q^{d'/2} \\
&= \frac{1}{d2^q} \left(\frac{q}{q-1} q^{d/2} \right) + \frac{C}{d} \left(\frac{q^{1/2}}{q^{1/2}-1} q^{d/2} \right) \\
&= \frac{C'}{d} q^{d/2}
\end{aligned}$$

for some constant C' dependent on q but independent of d. We have now proven Theorem 2.1.7. □

2.1.5 Genus Calculation and An Alternative Way to Compute the Large d Limit

The number of 2-superirreducibles can also be counted using the point counts for the affine curves

$$C_n(\{u_1, \ldots, u_n\}; \delta) = \{(x, y_1, \ldots, y_n) \in \mathbb{A}^{n+1} : \delta y_j^2 = x + u_j \text{ for } 1 \leq j \leq n\}.$$

More specifically, let the u_j range over all of $\mathbb{F}_q$, and let $\delta \in \mathbb{F}_{q^d}$ be a quadratic nonresidue. We note that we don't have to divide by 2 in the product since for each $\delta^{-1}(\alpha + u)$ that is a square in $\mathbb{F}_{q^d}$, there are two possible square roots. Therefore this gives rise to two points. Then,

$$\begin{aligned}
\#C_n(\mathbb{F}_q; \delta)(\mathbb{F}_{q^d}) &= \sum_{\alpha \in \mathbb{F}_{q^d}} \prod_{u \in \mathbb{F}_q} \left(1 + \chi_{q^d}\left(\delta^{-1}(\alpha + u)\right)\right) \\
&= \sum_{\alpha \in \mathbb{F}_{q^d}} \prod_{u \in \mathbb{F}_q} \left(1 - \chi_{q^d}(\alpha + u)\right) \\
&= q^d + \sum_{\{\} \neq U \subseteq \mathbb{F}_q} (-1)^{|U|} a_{q^d}(U).
\end{aligned}$$

Thus, Proposition 2.1.10 may be rephrased as

$$s_2(q,d) = \frac{1}{d2^q} \sum_{\substack{d'|d \\ 2\nmid \frac{d}{d'}}} \mu\left(\frac{d}{d'}\right) \#C_n(\mathbb{F}_q;\delta_{d'})(\mathbb{F}_{q^{d'}}) \tag{2.1.1}$$

where $\delta_{d'}$ is any quadratic nonresidue in $\mathbb{F}_{q^{d'}}$.

We next show that the projectivisation of C_n is a complete intersection in $\mathbb{P}^{n+1}$. We can then use this fact to calculate its arithmetic genus.

Lemma 2.1.12. *For any value of $n \geq 1$, the projectivisation of C_n is a complete intersection.*

Proof. Let I be the ideal generated by C_n. For each n, we define the rings R_n by

$$R_n := \frac{\mathbb{F}_{q^d}[X,Z,Y_1,\ldots,Y_n]}{I}.$$

To show that C_n is a complete intersection, we need to show that R_n has Krull dimension 2. In order to do this, we use the fact that R_n is an $\mathbb{F}_{q^d}[X,Z]$-module. More precisely we have

$$R_n \cong \bigoplus_{1\leq i_1<i_2<\cdots<i_l\leq n} \mathbb{F}_{q^d}[X,Z]Y_{i_1}Y_{i_2}\ldots Y_{i_l}.$$

I.e. the R_n are built out of polynomials where each monomial term can contain any power of X or Z, but can only contain each Y_i to at most the first power.

By definition of Krull dimension of a module, we have

$$\begin{aligned} \dim_{\mathbb{F}_{q^d}[X,Z]}(R_n) &:= \dim\left(\frac{\mathbb{F}_{q^d}[X,Z]}{Ann(R)}\right) \\ &= \dim\left(\mathbb{F}_{q^d}[X,Z]\right) \\ &= 2. \end{aligned}$$

The above is because $\mathbb{F}_{q^d}[X,Z]$ appears as a summand of R_n, so $Ann(R_n) = \{0\}$. The Krull dimension of $\mathbb{F}_{q^d}[X,Z]$ is 2 since it's just a polynomial ring in 2 variables.

Since R_n is a direct sum, we know that its Krull dimension as a ring is just the Krull dimension of $\mathbb{F}_{q^d}[X, Z]$. Therefore we conclude that R_n has the same Krull dimension as a ring and as an $\mathbb{F}_{q^d}[X, Z]$-module. This tells us R_n is a complete intersection, as desired. □

The next formula we will write down is due to Pieter Belmans and given in [2]. We'll use it to find the genera of our curves $\{C_n\}$.

Theorem 2.1.13. *A complete intersection of degrees $(d_1, d_2, \ldots, d_{k-1})$ inside $\mathbb{P}^k$ has genus*

$$g = 1 + \frac{1}{2}\left(\prod_{i=1}^{k-1} d_i\right)\left(\sum_{i=1}^{k-1} d_i - k - 1\right).$$

This formula gives us the following theorem:

Theorem 2.1.14. *For each n, the curve C_n has genus $1 + 2^{n-1}(n-2)$.*

Lemma 2.1.15. *For each fixed set $(\delta, u_1, u_2, \ldots, u_n)$ in $\mathbb{C}$ we can write down explicitly the singular points for the corresponding curve C_n. These are isolated over $\mathbb{C}$.*

Proof. We begin by finding the points not at infinity. We find the partial derivatives for the polynomials defining C_n and consider the Jacobian matrix. The set of polynomials are

$$\{f_i(x, y_1, y_2, \ldots, y_n) := \delta y_i^2 - x - u_i\}_{i=1}^n.$$

Each f_i has the set of partial derivatives

$$\frac{\partial f_i}{\partial x} = -1, \qquad \frac{\partial f_i}{\partial y_i} = 2\delta y_i, \qquad \frac{\partial f_i}{\partial y_j} = 0 \quad \text{for } j \neq i.$$

Therefore the Jacobian matrix corresponding to C_n is

$$\begin{pmatrix}
-1 & 2\delta y_1 & 0 & 0 & \ldots & 0 & 0 \\
-1 & 0 & 2\delta y_2 & 0 & \ldots & 0 & 0 \\
-1 & 0 & 0 & 2\delta y_3 & \ldots & 0 & 0 \\
-1 & \cdot & \cdot & \cdot & \ldots & \cdot & 0 \\
-1 & \cdot & \cdot & \cdot & \ldots & \cdot & 0 \\
-1 & \cdot & \cdot & \cdot & \ldots & \cdot & 0 \\
-1 & 0 & 0 & 0 & \cdots & 0 & 2\delta y_n
\end{pmatrix}$$

This matrix is $n \times (n+1)$ and so C_n has singularities at exactly the points where this matrix has rank strictly less than n. We know that δ is a quadratic non-residue, so it has to be non-zero. Therefore looking at the $n \times n$ minors, we see that rank strictly smaller than n is achieved if and only if at least two of the y_j's are zero at the same time. We can explicitly compute these points. For example if $y_1 = y_2 = 0$ the set of points are

$$\left(-u_1, 0, 0, \pm\sqrt{\frac{u_3 - u_1}{\delta}}, \pm\sqrt{\frac{u_4 - u_1}{\delta}}, \ldots, \pm\sqrt{\frac{u_n - u_1}{\delta}}\right)$$

where the choice of ± in each case gives 2^{n-2} possible points.This also covers the case with more than two of the $u_i = 0$, for example if $u_1 = u_2 = u_3 = 0$. We get similar sets when restricting $y_l = y_m = 0$: points that have 0's in the lth and mth coordinates, $-u_l$ in the first coordinate and $\pm\sqrt{\frac{u_k - u_l}{\delta}}$ in the kth coordinate for all k not equal to $1, l$ or m. Finally we realise that there's one more point in the projectivisation of C_n: the point at infinity. This is also a singular point. Combining all the sets we've come up with gives the collection of singular points, which we can see are isolated over $\mathbb{C}$. □

Remark 2.1.16. *Since the singularities for the C_n are isolated over $\mathbb{C}$ it means that they must be isolated over $\mathbb{F}_{q^N}$ for some N large enough.*

Corollary 7.2 in [14] gives us a way to bound the number of points on the curve C_n, which is a complete intersection with only isolated singular points over $\mathbb{F}_{q^N}$. The corollary tells us that

$$|\#C_n(\mathbb{F}_{q^N}) - q^N)| \leq b'_{n-2}(n, \mathbf{d})q^{N/2} + (b'_{n-1}(n+1, \mathbf{d}) + 1)q^{(N-1)/2} \tag{2.1.2}$$

where **d**$=$ $(2, 2, \ldots, 2)$ is the multidegree of C_n and b'_j denotes the jth primitive Betti number computed using the formula

$$b'_j(n, \mathbf{d}) = (-1)^{j+1}(j+1) + (-1)^{n+1}\sum_{c=r}^{n+1}(-1)^c\binom{n+1}{c+1}\sum_{\nu\in M(c)}\mathbf{d}^\nu$$

where $M(c)$ denotes the set of r-tuples $(\nu_1, \ldots, \nu_r)$ of positive integers such that $\nu_1 + \cdots + \nu_r = c$ and $d^\nu = d_1^{\nu_1}\ldots d_r^{\nu_r}$

We can now complete an alternative proof to Theorem 2.1.7

Proof. We have from (2.1.1) that

$$s_2(q, d) = \frac{1}{d2^q}\sum_{\substack{d'|d \\ 2\nmid\frac{d}{d'}}}\mu\left(\frac{d}{d'}\right)\#C_n(\mathbb{F}_q; \delta_{d'})(\mathbb{F}_{q^{d'}}).$$

For d large enough we have that the main term in this sum arises when $d' = d$ and is equal to $\frac{1}{d2^q}\#C_n(\mathbb{F}_{q^d})$, which we see from (2.1.2) is $\frac{q^d}{d2^q} + O(q^{d/2})$. Since $2 \nmid \frac{d}{d'}$, the next largest possible term in the sum would come from when $d' = \frac{d}{3}$, in which case $\frac{1}{d2^q}\#C_n(\mathbb{F}_{q^{d/3}})$ is $O(q^{d/3})$. Therefore we're able to conclude that $s_2(q, d) = \frac{q^d}{d2^q} + O(q^{d/2})$.

□

Remark 2.1.17. *In fact given our fixed* d*, we know for all* $m \leq d$ *that*

$$|\#C_n(\mathbb{F}_{q^m}) - q^m)| \leq b'_{m-2}(m, \boldsymbol{d})q^{m/2} + (b'_{m-1}(m+1, \boldsymbol{d}) + 1)q^{(m-1)/2}$$

will hold. This is because the bounds arise from the primitive Betti numbers which are computed from the roots of the Hasse-Weil functions in $\mathbb{F}_{q^d}$. If the set of roots are $\{\alpha_1, \ldots \alpha_\ell\}$ in $\mathbb{F}_{q^d}$, then they're $\{\alpha_1^{1/d}, \ldots \alpha_\ell^{1/d}\}$ in $\mathbb{F}_q$, giving us corresponding Betti numbers and bounds in $\mathbb{F}_q$. A more complete explanation of this process is given in [28].

2.1.6 Vanishing in the large q limit

If we fix the degree d, the heuristic argument suggests that $s_2(q,d) = 0$ for all q sufficiently large, because

$$\sum_{q \geq N} \frac{q^d}{d2^q} \to 0 \text{ as } N \to \infty.$$

Based on the heuristic as well as ample numerical evidence, we conjecture that in fact $s_2(q,d) = 0$ for q large enough. In particular, we can prove that $s_2(q,d) = 0$ for $q \gg d^2$.

The idea for proving $s_2(q,d) = 0$ for $q \gg d^2$ is as follows. We saw in Subsection 2.1.3 for $f \in \mathbb{F}_q[t]$ irreducible of degree d and α a root of f in $\mathbb{F}_{q^d}$, that f is 2-superirreducible over $\mathbb{F}_q$ iff $\alpha + c$ is a nonsquare in $\mathbb{F}_{q^d}$ for every $c \in \mathbb{F}_q$. This means the nontrivial quadratic character χ on $\mathbb{F}_{q^d}$ is identically equal to -1 on $\{\alpha + c : c \in \mathbb{F}_q\}$. So for there to be no 2-superirreducibles of degree d, it's enough to show for every $\alpha \in \mathbb{F}_{q^d}$ that there exists some $c \in \mathbb{F}_q$ such that $\chi_{q^d}(\alpha + c) = 1$. This will be true as long as we have $|\sum_{k=0}^{q-1} \chi_{q^d}(\alpha + u_k)| < q$, where the sum is over all $u_k \in \mathbb{F}_q$.

Before proving $s_2(q,d) = 0$ for $q \gg d^2$, we introduce a few definitions which we'll use in the proof.

Definition 2.1.18. *Suppose $q = p^n$. Let $\{\beta_1, \ldots, \beta_n\}$ be a basis of $\mathbb{F}_q$ over $\mathbb{F}_p$. For each integer k with $0 \leq k < q$, we define ζ_k by $\zeta_k := k_1\beta_1 + \cdots + k_n\beta_n$ where the k_i are given by the p-adic expansion of k: $k = k_1 + k_2p + \cdots + k_np^{n-1}$.*

Definition 2.1.19. *Let k and j be integers with $k \geq 0$ and $j < q$. We define $k \oplus j$ by*

$$k \oplus j = i \iff \zeta_k + \zeta_j = \zeta_i \text{ for } 0 \leq i < q.$$

Theorem 2.1.20. *Let χ be the nontrivial quadratic character on $\mathbb{F}_{q^d}$. Let $\alpha \in \mathbb{F}_{q^d}$ be such that $\mathbb{F}_{q^d} = \mathbb{F}_q(\alpha)$. Then we have that $|\sum_{k=0}^{q-1} \chi(\alpha + u_k)| < q$, where the sum is over all nonzero $u_k \in \mathbb{F}_q$.*

Proof. We use the idea from the proof of Theorem 3.1 in Winterhof [29] and Corollary 2.4 in Wan [27]. Denote $\sum_{k=0}^{q-1}\chi(\alpha+u_k)$ by $S(\alpha)$. We fix j with $0 \leq j \leq q-1$. We can see that each u_k can be written $\zeta_{j\oplus l}$ for some unique l with $0 \leq l \leq q-1$. This is because the β's form a basis for $\mathbb{F}_{q^d}/\mathbb{F}_q$ and if $l_1 \neq l_2$ then $\zeta_{l_1\oplus j} \neq \zeta_{l_2\oplus j}$. This tells us that:

$$S(\alpha) = \sum_{l=0}^{q-1}\chi(\alpha+\zeta_{l\oplus j}).$$

Next we restrict the j's such that $0 \leq j \leq J-1$. We sum over all such j's in this interval to get:

$$J|S(\alpha)| \leq \left|\sum_{l=0}^{q-1}\sum_{j=0}^{J-1}\chi(\alpha+\zeta_{l\oplus j})\right|.$$

We define

$$W := \sum_{l=0}^{q-1}\sum_{j=0}^{J-1}\chi(\alpha+\zeta_{l\oplus j}).$$

The triangle inequality tells us that

$$\left|\sum_{l=0}^{q-1}\sum_{j=0}^{J-1}\chi(\alpha+\zeta_{l\oplus j})\right| \leq \sum_{l=0}^{q-1}\left|\sum_{j=0}^{J-1}\chi(\alpha+\zeta_{l\oplus j})\right|.$$

Next by Cauchy-Schwarz, we obtain

$$\begin{aligned} W^2 &\leq (q-1)\sum_{l=0}^{q-1}\left|\sum_{j=0}^{J-1}\chi(\alpha+\zeta_{l\oplus j})\right|^2 \\ &\leq (q-1)\sum_{\zeta\in\mathbb{F}_q}\left|\sum_{j=0}^{J-1}\chi(\zeta+\alpha+\zeta_j)\right|^2 \\ &= (q-1)\sum_{j_1,j_2=0}^{J-1}\sum_{\zeta\in\mathbb{F}_q}\chi\left((\zeta+\alpha+\zeta_{j_1})\left(\zeta+\alpha+\zeta_{j_2}\right)\right) \end{aligned}$$

where we got the last line using the fact that χ takes values in $\{-1,1\}$ and has order 2 (so that $\overline{\chi}=\chi$).

For each $\alpha+\zeta_{j_1} \in \mathbb{F}_{q^d}$, there are at most d indices j_2 such that $\alpha+\zeta_{j_1}$ and $\alpha+\zeta_{j_2}$ are Galois conjugates over $\mathbb{F}_q$. For these indices, the inner sum can simply be estimated by q.

If $\alpha + \zeta_{j_1}$ and $\alpha + \zeta_{j_2}$ are not Galois conjugates, then their minimal polynomials (call them f_1 and f_2 respectively) are coprime and squarefree over $\mathbb{F}_q$. Let $f(t) := f_1(t)f_2(t)$. We can define a multiplicative character χ_f on $\left(\mathbb{F}_{q^d}[t]/(f)\right)^*$ via the formula

$$\chi_f(g(t)) = \chi(N(g(\alpha + \zeta_{j_1}))N(g(\alpha + \zeta_{j_2})))$$

where χ is again the quadratic character on $\mathbb{F}_{q^d}$. In this formula, we're letting $N(\beta)$ denote the norm of the β (i.e. the product of all of its Galois conjugates) over $\mathbb{F}_q$. We observe for $a \in \mathbb{F}_q$ that

$$\begin{aligned} \chi_f(a - t) &= \chi(N(a - \alpha - \zeta_{j_1})N(a - \alpha - \zeta_{j_2})) \\ &= \chi(f_1(a)f_2(a)). \end{aligned}$$

We show that χ_f can't be trivial. The f_i's have degrees dividing d (odd) and each $a - \alpha - \zeta_{j_i}$ is a square in $\mathbb{F}_{q^d}$ iff its Galois conjugates are squares. Therefore, since the f_i's have odd degrees, we can rewrite $\chi_f(a-t)$ as $\chi(a-\alpha-\zeta_{j_1})\chi(a-\alpha-\zeta_{j_2})$. By assumption, χ is non-trivial on each of the sets $\{a - \alpha - \zeta_{j_1} : a \in \mathbb{F}_q\}$ and $\{a - \alpha - \zeta_{j_2} : a \in \mathbb{F}_q\}$ (otherwise every element in one of these sets is a square and there's nothing to prove). By Chinese remainder theorem on the coprime f_i, we see that χ_f is non-trivial and so Weil's theorem on characters is applicable to χ_f.

We end up with the bound

$$W^2 < (q-1)Jdq + (q-1)J^2(2d-1)\sqrt{q}.$$

To make this bound as strong as possible, we choose $J = \lceil \sqrt{q} \rceil$ This tells us

$$\frac{W^2}{J^2} < (q-1)(3d-1)\sqrt{q}.$$

Finally we substitute this into $J|S(\alpha)| \leq W$ to get $|S(\alpha)| < 2.2q^{1/4}\sqrt{qd}$, which gives us $|S(\alpha)| < q$ as long as $q \gg d^2$. □

2.2 2-Superirreducibility over $\mathbb{Z}$ and $\mathbb{Q}$

2.2.1 Using Newton Diagrams to Understand 2-Superirreducibility

The technique of Newton Diagrams and the Dumas Criterion from Section 1.4 allows us to make some progress towards superirreducibility, however they don't suffice to completely prove superirreducibility. We give an example of how we can use this technique to show that a given irreducible polynomial remains irreducible under a non-trivial family of polynomial substitutions.

Example 2.2.1. *Using Dumas criterion and Newton Diagrams, we can prove that the irreducible polynomial $x^5 + 2x + 1$ remains irreducible over $\mathbb{Z}$ after substitution by*

$$g(x) = 5x^2 + a$$

for all $a \in \mathbb{Z}$.

Proof. We let $f(x) = x^5 + 2x + 1$ and $g(x) = 5x^2 + a$. Then we have

$$f(g(x)) = 5^5x^{10} + 5^5ax^8 + 2 \cdot 5^4a^2x^6 + 2 \cdot 5^3a^3x^4 + 5(5a^4 + 2)x^2 + (a^5 + 2a + 1).$$

We first note that if $f(g(x))$ splits over $\mathbb{Z}$, then it must do so as the product of two quintics. We can see this by assuming otherwise. Suppose that $h(x)$ is a factor of $f(g(x))$ of degree strictly less than 5. Then $h(x)$ has a root β and the extension $\mathbb{Q}(\beta)$ over $\mathbb{Q}$ is of degree strictly less than 5. But then $g(\beta) \in \mathbb{Q}(\beta)$ would be a root of f, which is impossible as f is an irreducible quintic.

Next, we consider the Newton diagram for $p = 5$. We're not sure what power of 5 will divide the constant term $a^5 + 2a + 1$ but we are able to draw the Newton diagram for the x^2 term onwards:

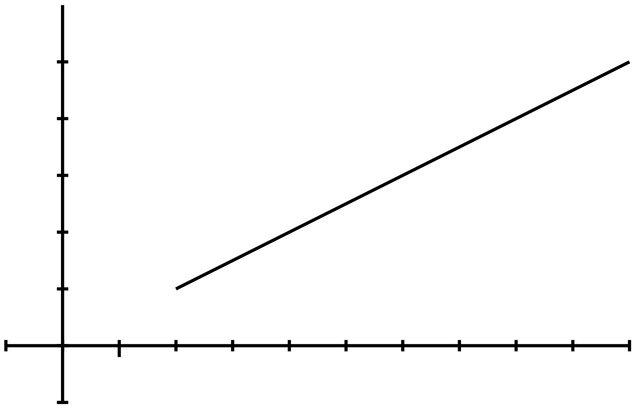

Figure 2.1: Partial Newton Diagram for $f(g(x))$ where $f(x) = x^5 + 2x + 1, g(x) = 5x^2 + a$ and $p = 5$

We don't know exactly where the points are on the Newton diagram for the x^4, x^6 and x^8 terms, but we note from the expansion of $f(g(x))$ that they must be strictly above the line. We suppose that $f(g(x))$ splits as $h_1(x)h_2(x)$ for irreducible quintics h_1 and h_2. Then by Dumas criterion, from the x-term onwards, their Newton diagrams must look like:

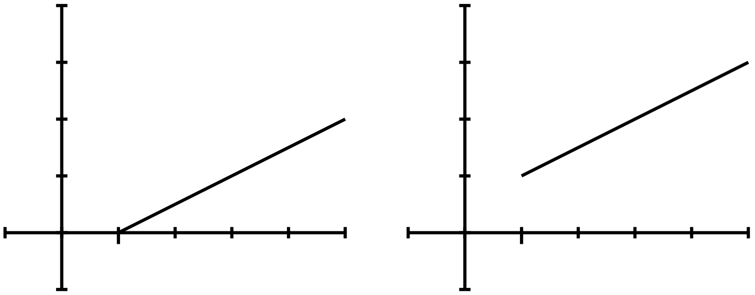

Figure 2.2: Partial Newton Diagrams for factors of $f(g(x))$, if $f(g(x))$ factors into two quintics, where $f(x) = x^5 + 2x + 1, g(x) = 5x^2 + a$ and $p = 5$

From these diagrams, we receive the expressions:

$$h_1(x) = a_0 + k_1x + 5a_2x^2 + 5a_3x^3 + 5^2a_4x^4 + 5^2x^5,$$

$$h_2(x) = b_0 + 5j_1x + 5^2b_2x^2 + 5^2b_3x^3 + 5^3b_4x^4 + 5^3x^5$$

where j_1 and k_1 are some integers not divisible by 5 and the a_i and b_i can be any integers. However we observe that if 5 doesn't divide b_0, then $f(g(x))$ has for its x-term a coefficient of $5j_1a_0 + k_1b_0$ which can't be equal to 0 (since $5j_1a_0$ is divisible by 5 but k_1b_0 is not.)

Therefore 5 must divide b_0.

From using $f(g(x)) = h_1(x)h_2(x)$, equating coefficients and remembering that b_0 must be divisible by 5, we receive:

$$x^2 : j_1 k_1 \equiv 2 \ \ (5)$$

$$x^4 : 5 | b_3 k_1 + a_3 j_1$$

$$x^6 : 5 | a_3 b_3 + k_1 + j_1$$

$$x^8 : 5 | a_3 + b_3.$$

From the x^8 term, we get that $b_3 = -a_3 + 5r_3$ for some integer r_3. If we substitute this into the expression for x^4, we receive $(-a_3 + 5r_3)k_1 + a_3 j_1 \equiv 0$, so that $a_3(j_1 - k_1) \equiv 0$. This tells us that either $a_3 \equiv 0 \pmod 5$ or that $j_1 - k_1 \equiv 0 \pmod 5$.

Considering the first possibility, if $a_3 \equiv 0 \pmod 5$, then from the x^8 term, we'd also have $b_3 \equiv 0 \pmod 5$. Next from the x^6 term, we'd deduce that $k_1 \equiv -j_1 \pmod 5$. Substituting this into the x^2 term, we receive $-j_1^2 \equiv 2 \pmod 5$ so that $j_1^2 \equiv 3 \pmod 5$, which is impossible.

Now considering the second possibility, $j_1 \equiv k_1 \pmod 5$. But this time, substituting into the x^2 term, we receive $j_1^2 \equiv 2 \pmod 5$, which is also impossible. Therefore, such h_1 and h_2 can't exist and $f(g(x))$ must be irreducible.

□

Remark 2.2.2. *As a generalisation of this example, we'll prove later in Subsection 2.2.4 that all polynomials in the family $\{x^{2k+1} + 2x + 1\}$ (for $k \geq 2$) are irreducible and remain irreducible under every substitution of form $g(x) = ax^2 + b$.*

2.2.2 Cases where there are no superirreducibles

We can identify families of polynomials over $\mathbb{Q}$ and $\mathbb{Z}$ that have obstructions to being superirreducible. In the next two theorems, we demonstrate that there are no degree n

polynomials which are k-superirreducible for any $k \geq n-1$.

Theorem 2.2.3. *If $f(x)$ is a polynomial of degree n with rational coefficients, then for any integer $k \geq 0$, f is not $(n+k)$-superirreducible over $\mathbb{Z}$ or over $\mathbb{Q}$.*

Proof. Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $g(x) = x^k f(x) + x$. Then we have

$$\begin{aligned} f(g(x)) &= f(x^k f(x) + x) \\ &= a_n(x^k f(x) + x)^n + a_{n-1}(x^k f(x) + x)^{n-1} + \cdots + a_1(x^k f(x) + x) + a_0 \\ &\equiv a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \pmod{f(x)} \\ &\equiv 0 \pmod{f(x)} \end{aligned}$$

so that $f(g(x))$ is divisible by $f(x)$ and so reducible. □

Corollary 2.2.4. *There are no 2-superirreducible quadratic polynomials over $\mathbb{Z}$ or over $\mathbb{Q}$.*

Theorem 2.2.5. *For every $d \geq 3$, there are no $(d-1)$-superirreducibles over $\mathbb{Q}$ or over $\mathbb{Z}$ of degree d.*

A slightly more general version of this is a Theorem of Schinzel's [24]. For completeness, we give the proof specific to our case.

Proof. Let $f(x)$ be an irreducible polynomial of degree d with integer coefficients. Suppose first that f is monic, so that $f(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$. Now we let α be a root of f in its splitting field over $\mathbb{Q}$ and $\beta = \frac{1}{\alpha}$. Then $\alpha^d = -(a_{d-1}\alpha^{d-1} + \cdots + a_1\alpha + a_0)$. Dividing both sides by α^{d-1} we find that $\alpha = -(a_{d-1} + a_{d-2}\beta + \cdots + a_1\beta^{d-2} + a_0\beta^{d-1})$. Therefore $\alpha = g(\beta)$ for a polynomial $g(x) \in \mathbb{Z}[x]$ of degree $d-1$. So by Proposition 2.1.4, $f(g(\beta))$ splits over $\mathbb{Z}$. This tells us there are no monic $(d-1)$-superirreducibles of degree d over the integers.

Now suppose that f is irreducible over the integers but not necessarily monic. Let $f(x) = a_d x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0$. Drawing inspiration from the monic case, we let $h(x) = x^d f\left(\frac{1}{x} - \frac{a_{d-1}}{(d-1)a_0}\right)$. Next we define $g(x)$ (our polynomial of degree $d-1$ to be substituted into f) as $g(x) = -\frac{h(x)}{a_d x} + \frac{1}{x} - \frac{a_{d-1}}{(d-1)a_d}$. We note that by adding the $\frac{1}{x}$ term, $g(x)$ is a polynomial. We also notice

$$\begin{aligned} f(g(x)) &= f\left(-\frac{h(x)}{a_d x} + \frac{1}{x} - \frac{a_{d-1}}{(d-1)a_d}\right) \\ &\equiv 0 \pmod{h(x)}. \end{aligned}$$

I.e. we've found a factorisation of $f(g(x))$ that has a degree d factor. This is a factorisation over the rationals. To get a factorisation over the integers, we simply need to rescale things. We see that if we rescale by $(d-1)a_d^2$, then we get a factorisation over $\mathbb{Z}$. I.e. if $\tilde{h}(x) = h((d-1)^2 a_d^2 x)$ and $\tilde{g}(x) = g((d-1)^2 a_d^2 x)$, then $\tilde{g}(x)$ and $\tilde{h}(x)$ are polynomials in $\mathbb{Z}[x]$ such that $f(\tilde{g}(x))$ is divisible by $\tilde{h}(x)$.

□

Corollary 2.2.6. *There are no* 2*-superirreducible polynomials of degree* 3*.*

Theorem 2.2.7. *For* $a, b \in \mathbb{Q}$ *and* N *odd, there are no* 2*-superirreducibles over* $\mathbb{Q}$ *of the form* $ax^N - b$*.*

Proof. We first deal with the case $f(x) = x^N - b$ with $b > 0$. We substitute in $g(t) = bt^2$ to get the following factorisation for $f(g(t))$:

$$\begin{aligned} f(g(t)) &= b^N t^{2N} - b \\ &= b(b^{N-1}t^{2N} - 1) \\ &= b(b^{\frac{1}{2}(N-1)}t^N + 1)(b^{\frac{1}{2}(N-1)}t^N - 1). \end{aligned}$$

Next we consider the case $f(x) = x^N + b$ with $b > 0$. This time, we let $g(t) = -bt^2$

and get the following factorisation:

$$\begin{aligned} f(g(t)) &= (-b)^N t^{2N} + b \\ &= -b^N t^{2N} + b \\ &= -b(b^{N-1} t^{2N} - 1) \\ &= -b(b^{\frac{1}{2}(N-1)} t^N + 1)(b^{\frac{1}{2}(N-1)} t^N - 1). \end{aligned}$$

Finally we look at the case $ax^N - b$ with a and b rationals. We make the substitution $g(t) = \frac{b}{a} t^2$ to get:

$$\begin{aligned} f(g(t)) &= a\left(\frac{b}{a}t^2\right)^N - b \\ &= \frac{b^N}{a^{N-1}} t^{2N} - b \\ &= b\left(\left(\frac{b}{a}\right)^{N-1} t^{2N} - 1\right) \\ &= b\left(\frac{b}{a}\right)^{N-1}\left(t^{2N} - \left(\frac{a}{b}\right)^{N-1}\right). \end{aligned}$$

The expression we get in this final line can be handled by the first two cases, completing the proof. □

Next we discuss how a family of polynomials fails to be superirreducible.

Theorem 2.2.8. *If $f(x)$ is a polynomial over $\mathbb{Q}$ of degree $2N$, its linear term has nonzero coefficient and all of its other odd degree terms have zero coefficients, then $f(x)$ is not N-superirreducible.*

Proof. Let $f(x) = a_{2N}x^{2N} + a_{2N-2}x^{2N-2} + \cdots + a_2x^2 + a_1x + a_0$, where in particular $a_1 \neq 0$. Let α be a root of f in its splitting field. Then we see that

$$\alpha = -\frac{1}{a_1}\left(a_{2N}\alpha^{2N} + a_{2N-2}\alpha^{2N-2} + \cdots + a_2\alpha^2 + a_0\right).$$

We now set $\beta = \alpha^2$. Then $\alpha = -\frac{1}{a_1}\left(a_{2N}\beta^N + a_{2N-2}\beta^{N-1} + \cdots + a_2\beta + a_0\right)$. So if we now define a degree N polynomial by $g(x) = -\frac{1}{a_1}\left(a_{2N}x^N + a_{2N-2}x^{N-1} + \cdots + a_2x + a_0\right)$, we see that β is a solution to $g(x) = \alpha$ in $\mathbb{Q}(\alpha)$ and so by Lemma 2.1.4, $f(g(x))$ splits over $\mathbb{Q}$. □

Corollary 2.2.9. *Let $f(x) = x^4 + ax^3 + bx^2 + cx + d$ be a monic quartic polynomial with rational coefficients and nonzero linear term. If $-a^3 \neq 16c$, then $f(x)$ cannot be 2-superirreducible over the rationals.*

Proof. By completing the fourth power, we can assume f has a zero coefficient for its third degree term. The condition $-a^3 \neq 16c$ ensures that the linear term has nonzero coefficient. We conclude by Theorem 2.2.8. □

We note that in the last corollary, since we're considering superirreducibility over $\mathbb{Q}$, if f was not monic, we could still divide out by the leading coefficient and apply the corollary.

2.2.3 Discussion of 2-superirreducibility for degree 4 polynomials

Theorem 2.2.10. *The polynomial $f(x) = x^4 + 1$ is 2-superirreducible over the integers.*

Proof. Let α be a root of f in its splitting field. Let $g(x) = ax^2 + bx + c$ with $a, b, c \in \mathbb{Z}$. Suppose that $f(g(x))$ factors over the integers so that it's the product of two polynomials, $h_1(x)$ and $h_2(x)$, say.

We show that the h_i must each be of degree 4. Suppose that $\deg(h_1(x)) < 4$ and that it has a root γ. Then $g(\gamma)$ is a root of f and since f is irreducible, it's the minimal polynomial for $g(\gamma)$. However, we have that $\mathbb{Q}(g(\gamma)) \subset \mathbb{Q}(\gamma)$. Since $\deg(h_1) < 4$ we know $[\mathbb{Q}(\gamma) : \mathbb{Q}] < 4$, so that $[\mathbb{Q}(g(\gamma)) : \mathbb{Q}] < 4$. This is impossible since f has degree 4 and is the minimal polynomial of $g(\gamma)$.

Let β be a root of $g(x)-\alpha$ in the splitting field for f. We observe that $\mathbb{Q}(\alpha)=\mathbb{Q}(g(\beta))$ so that $\mathbb{Q}(\alpha)$ is a subfield of $\mathbb{Q}(\beta)$. However $g(\beta)=\alpha$ implies that β must be a root of either h_1 or h_2, so that $[\mathbb{Q}(\beta):\mathbb{Q}]=d$. This tells us that $\mathbb{Q}(\alpha)=\mathbb{Q}(\beta)$.

Using the fact that $g(\beta)=\alpha$, we have that $(2a\beta+b)^2=4a\alpha+(b^2-4ac)$, which we can rewrite as $m\alpha+n$ with m,n integers. Taking norms we find:

$$\begin{aligned}(N(2a\beta+b))^2 &= N(m\alpha+n)\\ &= m^4N(\alpha)+N(n).\end{aligned}$$

Since $2a\beta+b$ is an algebraic integer, its norm must be an integer. This gives us the equation $m^4+n^4=z^2$ where m,n and z are all integers. However it's known by infinite descent (for example, we showed it in Example 1.6.2)) that this equation has no non-trivial integer solutions.

Since z is a norm of a nonzero element, it can't be zero. For g to be a quadratic , we know that $a\neq 0$ so that $m=4a\neq 0$. The only possibility left is that $n=0$, but then this would mean that g has a repeated root. I.e. $g(x)=r(sx-t)^2$ for some $r,s\in\mathbb{Z}$. This would then mean $f(g(x))=r^4(sx-t)^8+1$ which is irreducible over the rationals. We see this is the case because we can rewrite $f(g(x))$ as r^4y^8+1. This polynomial has roots $\left\{\sqrt{\frac{1}{r}}exp\left(\frac{(2k+1)\pi i}{8}\right)\right\}_{k=0}^{7}$. The set of permutations on these roots contains the Galois group for y^8+1, which acts transitively on the set $\left\{exp\left(\frac{(2k+1)\pi i}{8}\right)\right\}_{k=0}^{7}$. Therefore the set of permutations for the roots of $f(g(x))$ acts transitively, so we see that it's really a Galois group and $f(g(x))$ is irreducible.

So we conclude that in all cases $f(g(x))$ can't have a factorisation over $\mathbb{Z}$ and so $f(x)$ is 2-superirreducible over the integers.

□

Corollary 2.2.11. *For any $a\in\mathbb{Z}$, the polynomials $f(x)=x^4+a^4$ are 2-superirreducible* *over $\mathbb{Z}$.*

Proof. We run the same argument as for Theorem 2.2.10. This time we end up with

$$z^2 = m^4 + (an)^4$$

which has no non-trivial solutions over the integers. □

Theorem 2.2.12. *The polynomial $x^4 + 2$ is 2-superirreducible over the integers.*

Proof. We run the argument from Theorem 2.2.10 to get the equation $t^4 + 2y^4 = z^2$. We see that for some relatively prime integers p and q, we can write $z = t^2 + \frac{2p}{q}y^2$. Substituting this into $t^4 + 2y^4 = z^2$, we get that $\frac{t^2}{y^2} = \frac{q^2-2p^2}{2pq}$. Now assuming t and y are coprime, we get the parametrisation

$$t^2 = q^2 - 2p^2, \qquad y^2 = 2pq.$$

Putting this parametrisation into our original equation gives us $z^2 = q^4 + 4p^4$, which we'll show has no non-trivial integer solutions.

We see that $(q^2, 2p^2, z)$ is a Pythagorean triple, which gives us the parametrisation $q^2 = m^2 - n^2$, $2p^2 = 2mn, z = m^2 + n^2$. Dividing through by common factors, we can assume that $q^2, 2p^2$ and z are pairwise coprime. For example if k divides both q^4 and $4p^4$, then k divides q^2 so k divides q. We can then assume that m and n are coprime, since if k divides both m and n, we'd have that k divides $q, 2p$ and z. But since $p^2 = mn$, this tells us m and n must both be perfect squares, allowing us to write q^2 as a difference of fourth powers. By a known infinite descent (1.6) argument (such as the one given in [7]), the equation $q^2 = a^4 - b^4$ has no non-trivial solutions over the integers. The case of trivial solutions can be disregarded by exactly the same reasoning as in the proof of Theorem 2.2.10

□

We observe that our technique used in the last proof doesn't work to show that $f(x) = x^4 + 3$ is 2-superirreducible because the equation $m^4 + 3n^4 = z^2$ has nontrivial integer solutions (e.g. $m = 1, n = 1, z = 2$). Similarly $f(x) = x^4 + 5$ gives rise to the equation $m^4 + 5n^4 = z^2$ which has $m = 1, n = 2, z = 9$ as a solution. However, what we just showed does give the following corollary.

Corollary 2.2.13. *For $a \in \mathbb{Z}$, all the polynomials in the set $\{x^4 + 2a^4\}$ are 2-superirreducible over the integers.*

Remark 2.2.14. *If an integer is a $4k$-th power, then it is automatically a 4th power. Therefore all the polynomials in the families $\{x^{4k} + a^4\}$ and $\{x^{4k} + 2a^4\}$ are 2-superirreducible over the integers.*

We can also say something about 2-superirreducibility of a large family of non-monic quartic polynomials.

Theorem 2.2.15. *The polynomial $f(x) = ax^4 + b$ is 2-superirreducible if it's irreducible and a is a non-square mod b, or b is a non-square mod a.*

Proof. We first note that $ax^4 + b$ is almost always irreducible over the integers. In fact it's only reducible when it can be written as $c(x^4 - k^2)$ for some $c, k \in \mathbb{Z}$. We're then done by Legendre's Theorem (presented in [16]), which says that the equation $ax^4 + by^4 - z^2 = 0$ has nontrivial integer solutions if and only if a is a square mod b and b is a square mod a. □

2.2.4 A family of weakly 2-superirreducible polynomials

Definition 2.2.16. *Let R be a commutative domain and F its field of fractions. The polynomial $f(x) \in R[x]$ is weakly k-superirreducible over R if for all $g(t) \in R[t]$ of form $ax^j + b$ with $a, b \in \mathbb{Z}$ and $j \leq k$, the composition $f(g(t))$ is irreducible over R.*

We now state the main theorem of this section, which we'll prove at the end of the section.

Theorem 2.2.17. *For $k \geq 2$, all the polynomials in the family $\{x^{2k+1}+2x+1\}$ are weakly 2-superirreducible over the integers.*

Remark 2.2.18. *Concretely, Theorem 2.2.17 is saying that all the polynomials in the family are irreducible (and so remain irreducible after linear substitutions) and that under a substitution of the form $g(x) = ax^2 + b$, they remain irreducible.*

Lemma 2.2.19. *For $k \geq 1$, any polynomial of the form $x^{2k+1} + 2x + 1$ is irreducible over $\mathbb{Z}$.*

Proof. Let $f(x) = x^{2k+1} + 2x + 1$ and $F(x) = x^{2k+1} f\left(\frac{1}{x}\right) = x^{2k+1} + 2x^{2k} + 1$. We see that $f(x)$ is reducible over the integers if and only if $F(x)$ is.

Next we define $h(x)$ by $h(x) = F(x) - 1$. We see that $h(x)$ can be factored as $x^{2k}(x+2)$. Considering $h(x)$ as a polynomial over $\mathbb{C}$, we see that it has all but one of its roots inside the unit disk. Our goal is to use the symmetric Rouché Theorem to conclude that F has all but one of its roots inside the unit disk.

According to the symmetric Rouché Theorem (presented in [23]), the condition we need to verify for the above to be true is that the strict inequality $|F - h| < |F| + |h|$ holds on the unit circle. We see that $|F - h|$ is identically 1. However we have

$$\begin{aligned} |F| + |h| &\geq |h| \\ &= |x^{2k}||x+2| \\ &= |x+2|. \end{aligned}$$

We see that when $x \neq -1$, we have that $|x + 2| > 1$. But direct computation at $x = -1$ shows that $F(-1) = 2$ and $h(-1) = 1$, so here we also have $|F - h| < |F| + |h|$.

Therefore the symmetric Rouché is applicable and we conclude that $F(x)$ has at most one root outside the unit disk.

If $F(x)$ was reducible over the integers, then there would exist two polynomials $j_1(x), j_2(x)$ in $\mathbb{Z}[x]$ such that $F(x) = j_1(x)j_2(x)$. By looking at their zero order terms, j_1 and j_2 must both have a product of roots that is a nonzero integer. Therefore they must both have at least one root outside the unit disk, resulting in $F(x)$ having at least two roots outside the unit disk. This gives a contradiction, so we can conclude that $F(x)$ must be irreducible over the integers. Now we see that our original polynomial $x^{2k+1} + 2x + 1$ must also be irreducible over $\mathbb{Z}$. □

Next, we aim to understand the ring of algebraic integers in the splitting field of $x^{2k+1} + 2x + 1$. We'll do this by computing its discriminant.

Lemma 2.2.20. *The discriminant of the polynomial* $f(x) = x^{2k+1} + 2x + 1$ *is*

$$(-1)^k \left((2k+1)^{2k+1} + 2^{2k+1}(2k)^{2k}\right).$$

Proof. Let θ be a root of $f(x) = x^{2k+1} + 2x + 1$. We'll use $\{1, \theta, \theta^2, \ldots, \theta^{2k}\}$ as a basis for $\mathbb{Q}(\theta)$ over $\mathbb{Q}$. By definition, the discriminant of $f(x)$ is the determinant of the matrix whose (ij)-entry is given by $tr\ (\theta^{i-1}\theta^{j-1})$.

We have that $tr(1) = 2k+1$ (for it's just the trace of the $(2k+1)$-dimensional identity matrix). Then we use the fact that $\theta^{2k+1} = -2\theta - 1$ to see that $tr(\theta) = 0$ since it's the trace of the matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \dots & -1 \\ 1 & 0 & 0 & 0 & \dots & -2 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Next, the matrix for multiplication by θ^2 in our basis is

$$\begin{pmatrix} 0 & 0 & 0 & \dots & -1 & 0 \\ 0 & 0 & 0 & \dots & -2 & -1 \\ 1 & 0 & 0 & \dots & 0 & -2 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

This pattern continues until θ^{2k-1} so that the elements $\theta, \theta^2, \theta^3, \dots, \theta^{2k-1}$ all have trace zero.

Next we see that the matrix for multiplication by θ^{2k} in our basis is

$$\begin{pmatrix}
0 & -1 & 0 & \dots & 0 & 0 \\
0 & -2 & -1 & \dots & 0 & 0 \\
0 & 0 & -2 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 \\
\cdot & \cdot & \cdot & \dots & \cdot & 0 \\
\cdot & \cdot & \cdot & \dots & \cdot & 0 \\
\cdot & \cdot & \cdot & \dots & \cdot & -1 \\
1 & 0 & 0 & \cdots & 0 & -2
\end{pmatrix}$$

which has trace $-4k$. The matrix for multiplication by θ^{2k+1} in our basis is

$$\begin{pmatrix}
-1 & 0 & 0 & \dots & 0 & 2 \\
-2 & -1 & 0 & \dots & 0 & 4 \\
0 & -2 & -1 & \dots & 0 & 0 \\
0 & 0 & -2 & \dots & 0 & 0 \\
0 & 0 & 0 & \dots & 0 & 0 \\
\cdot & \cdot & \cdot & \dots & \cdot & 0 \\
\cdot & \cdot & \cdot & \dots & \cdot & 0 \\
\cdot & \cdot & \cdot & \dots & -1 & 0 \\
1 & 0 & 0 & \cdots & -2 & -1
\end{pmatrix}$$

which has trace $-(2k+1)$. Continuing onwards, the matrix for multiplication by θ^{2k+2} in our basis is

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 2 & 1 \\ -1 & 0 & 0 & \dots & 4 & 4 \\ -2 & -1 & 0 & \dots & 0 & 4 \\ 0 & -2 & -1 & \dots & 0 & 0 \\ 0 & 0 & -2 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & 0 \\ \cdot & \cdot & \cdot & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & -1 & 0 \end{pmatrix}$$

which has trace 0. This pattern continues, similarly to what happened above with $\theta, \theta^2, \dots, \theta^{2k-1}$, with columns shifting to the left (and more and more columns of mostly zeros except the elements $(1, 4, 4)$ appearing on the right). We see that this will result in zeros along the diagonal until the $(1, 4, 4)$ has shifted so that the second four is the bottom-rightmost entry of our matrix. This happens exactly at the element θ^{4k}, so this element has trace $8k$. This is because its matrix will be:

$$\begin{pmatrix} 0 & 2 & 1 & 0 & \dots & 0 \\ 0 & 4 & 4 & 1 & \dots & 0 \\ 0 & 0 & 4 & 4 & \dots & 0 \\ 0 & 0 & 0 & 4 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & 0 & \dots & 1 \\ -1 & 0 & 0 & 0 & \dots & 4 \\ -2 & -1 & 0 & 0 & \dots & 4 \end{pmatrix}$$

which has $(1, 1)$-entry zero and has 4 as all of its other diagonal entries.

We see that to get the discriminant of $f(x)$, we need to compute the determinant of the following $(2k+1)$-dimensional square matrix:

$$\begin{pmatrix}
2k+1 & 0 & 0 & \dots & 0 & -4k \\
0 & 0 & 0 & \dots & -4k & -(2k+1) \\
0 & 0 & \dots & -4k & -(2k+1) & 0 \\
\cdot & \cdot & \cdot & \dots & \cdot & \cdot \\
\cdot & \cdot & \cdot & \dots & \cdot & \cdot \\
\cdot & \cdot & \cdot & \dots & \cdot & 0 \\
\cdot & \cdot & \cdot & \dots & 0 & 0 \\
0 & -4k & -(2k+1) & \dots & 0 & 0 \\
-4k & -(2k+1) & 0 & \dots & 0 & 8k
\end{pmatrix}$$

We expand along the first row to see that the discriminant is exactly

$$(-1)^k\left((2k+1)^{2k+1} + 2^{2k+1}(2k)^{2k}\right)$$

as desired. We note that this value is odd for all k.

□

Corollary 2.2.21. *Let θ be a root of the polynomial $x^{2k+1}+2x+1$. Since the discriminant of $x^{2k+1}+2x+1$ is odd, elements of the ring of algebraic integers for the splitting field can all be expressed in the form*

$$\left\{\frac{a_0 + a_1\theta + \cdots + a_{2k}\theta^{2k}}{m}, a_i \in \mathbb{Z}, m \textit{ odd}\right\}.$$

We're now ready to prove Theorem 2.2.17.

Proof. Let $f(x) = x^{2k+1} + 2x + 1$ and $g(x) = ax^2 + c$. Then

$$f(g(x)) = (ax^2+c)^{2k+1} + 2(ax^2+c) + 1.$$

For this composition to be reducible, we need $ax^2 = \theta - c$ to be solvable in the splitting field for f. However we note that

$$ax^2 = \theta - c \implies a^2x^2 = a\theta - ac.$$

By replacing ax with $\tilde{x}$, we see that $\tilde{x}^2 = M\theta + N$ is an algebraic integer with M dividing N.

By Corollary 2.2.21, we can write $\tilde{x}^2 = \left(A_{2k}\theta^{2k} + A_{2k-1}\theta^{2k-1} + \cdots + A_1\theta + A_0\right)^2$, with the A_i's rational numbers having odd denominators. Therefore we can consider them all as 2-adic integers. By the argument in the previous paragraph, we can set

$$\left(A_{2k}\theta^{2k} + A_{2k-1}\theta^{2k-1} + \cdots + A_1\theta + A_0\right)^2 = M\theta + N.$$

Examining this expression, we see that after simplification, we need all of $\theta^2, \theta^3, \ldots, \theta^{2k-1}, \theta^{2k}$ to end up with zero coefficients.

We now consider the expansion of $\left(A_{2k}\theta^{2k} + A_{2k-1}\theta^{2k-1} + \cdots + A_1\theta + A_0\right)^2$. We will make use of the following relations (that are valid because θ is a root of f):

$$\begin{aligned}
&\theta^{2k+1} = -2\theta - 1\\
&\theta^{2k+2} = -2\theta^2 - \theta\\
&\cdots\\
&\cdots\\
&\theta^{4k-1} = -2\theta^{2k-1} - \theta^{2k-2}\\
&\theta^{4k} = -2\theta^{2k} - \theta^{2k-1}.
\end{aligned}$$

These relations tell us that when we expand out the parentheses, eliminate any terms of θ with powers higher than $2k$ and reduce modulo 2, we get two possible terms involving θ^j for each j. One term is $\left(A_{\frac{j}{2}}\theta^{\frac{j}{2}}\right)^2$, which arises whenever j is even. The other term arises because of the above relations.

The expressions $\theta^2, \theta^3, \ldots, \theta^{2k}$ all need to end up with zero coefficients. So then, we get from looking at even powers of θ that $A_1, A_2, \ldots, A_k, A_{2k-1}$ and A_{2k} are all zero modulo 2 and from looking at odd powers of θ that $A_{k+2}, A_{k+3}, \ldots, A_{2k-2}$ are all zero modulo 2. For example, let's take $j = 2$. Upon expansion, the first term involving θ^2 is $A_1^2\theta^2$. The second term is from $\theta^{2k+3} = -2\theta^3 - \theta^2$. However since $2k+3$ is odd, θ^{2k+3} can only come out of multiplying together θ^r and θ^s where $r + s = 2k + 3$ and $r \neq s$. Therefore all terms involving θ^{2k+3} must come with even coefficients, so disappear modulo 2. This tells us that $A_1 \equiv 0 \pmod 2$

The only coefficients we have left to deal with are A_0 and A_{k+1}. Using the relations we have, the coefficient of θ^2 is

$$2A_0A_2 + A_1^2 - 2A_{k+1}^2 - 4 \sum_{\substack{i+j=2k+2 \\ i\neq j}} A_iA_j - 2 \sum_{\substack{i+j=2k+3 \\ i\neq j}} A_iA_j.$$

We note that A_0 doesn't appear in either of the sums and that this whole expression should be even for the coefficient of θ^2 to vanish. Since 4 divides every other term in the expression, 4 must also divide $2A_{k+1}^2$, so that A_{k+1} is even. Finally, since M divides N and we're now able to conclude that $N \equiv A_0^2 \pmod 2$, we have that A_0 is also even.

We've managed to conclude that, considered as 2-adic integers, all the A_i's are divisible by 2. Therefore we can divide through the whole system of homogeneous Diophantine equations to get a new system with $\frac{M}{2}$ dividing $\frac{N}{2}$. By infinite descent, we are done.

□

2.2.5 Galois Groups

It's somehow intuitive to guess that for 2-superirreducibility, we should have large Galois groups and for non 2-superirreducibility, we should have small Galois groups. We have examples to the contrary going both ways.

Example 2.2.22. *By Theorem 2.2.5, there are no 2-superirreducible cubics over $\mathbb{Z}$. Therefore $x^3 + 2x + 1$ is definitely not 2-superirreducible. However it has Galois group S_3.*

Example 2.2.23. *$x^4 + 16$ has Galois group $C_2 \times C_2$, which has the smallest possible order for the Galois group of a degree 4 irreducible. However by Corollary 2.2.11, $x^4 + 16$ is 2 superirreducible over $\mathbb{Z}$.*

CHAPTER III

Binomial Coefficient Asymptotics

3.1 Introduction

3.1.1 Definitions and Preliminaries

Definition 3.1.1. *We write a positive integer n in base $b \geq 2$ as*

$$n := \sum_{i=0}^{k} a_i b^i \qquad \textit{for } b^k \leq n < b^{k+1}$$

with digits $0 \leq a_i \leq b-1$. Here $k = \lfloor \log_b(n) \rfloor$ and $a_i := a_i(n)$ with $a_k(n) \geq 1$

(1)The sum of digits function $d_b(n)$ (to base b) is

$$d_b(n) := \sum_{i=0}^{k} a_i(n)$$

with $k = \lfloor \log_b(n) \rfloor$

(2) The running digit sum function $S_b(n)$ (to base b) is

$$\sum_{j=0}^{n} d_b(j).$$

Delange's Theorem [10] then gives an expression for the running digit sum function:

Theorem 3.1.2. *(Delange) We have*

$$S_b(n) = \left(\frac{b-1}{2}\right) n \log_b(n) + f_b(\log_b(n))n$$

where $f_b(x)$ are non-positive continuous but not differentiable functions that have Fourier series expansions

$$f_b(x) = \sum_{k\in\mathbb{Z}} c_p(k)e^{2\pi ikx}$$

whose Fourier coefficients are,

$$c_b(k) = -\frac{b-1}{2k\pi i}\left(1+\frac{2k\pi i}{\log b}\right)^{-1}\zeta\left(\frac{2k\pi i}{\log b}\right) \qquad \text{for } k \neq 0$$
$$c_b(0) = \frac{b-1}{2\log b}(\log(2\pi)-1) - \left(\frac{b+1}{4}\right).$$

Definition 3.1.3. *The Farey sequence of order n is a finite set of fractions*

$$\left\{\frac{k}{m} : \frac{k}{m} \in (0,1], m \leq n\right\}.$$

Definition 3.1.4. *We define $\overline{G}(n) := \prod_{k=0}^{n}\binom{n}{k}$.*

The quantities $\overline{G}_n$ are interesting to study because not only is $\overline{G}_n$ the product of the binomial coefficients in the nth row of Pascal's triangle, but it's also the product of all the elements in the Farey sequence of order n.

The next theorem combines Theorems 3.2 and 5.1 from Lagarias and Mehta's paper [19]

Theorem 3.1.5. *We have two expressions for $\log\overline{G}_n$:*

$$\log\overline{G}(n) = \frac{1}{2}n^2 + O\left(n\log n\right) \tag{3.1.1}$$

and

$$\log\overline{G}(n) = \sum_{p\leq n}\frac{2\log p}{p-1}S_p(n) - \sum_{p\leq n}\frac{n-1}{p-1}d_p(n)\log p. \tag{3.1.2}$$

3.1.2 Main Results over Primes

Definition 3.1.6. *Let*

$$A(n) = \sum_{p\leq n}\frac{2\log p}{p-1}S_p(n)$$

and

$$B(n) = \sum_{p\leq n} \frac{n-1}{p-1} d_p(n) \log p.$$

According to (3.1.2) in Theorem 3.1.5, we have that

$$\log \overline{G}(n) = A(n) - B(n).$$

$B(n)$ can be thought of as a 'scaled average' function. Indeed $\frac{B(n)}{n(n-1)}$ gives exactly the average behaviour of $\frac{\log p}{p-1} d_p(n)$ over all the primes up to n. Although $d_p(n)$ itself is difficult to average, this allows us to say something about its average value when we scale it by a fairly elementary function $\frac{\log p}{p-1}$. Similarly $A(n)$ can be thought of as 'scaled double average function': first averaging over the $d_p(n)$ to get $S_p(n)$ and then taking another scaled average over the $S_p(n)$. $A(n)$ and $B(n)$ both have asymptotics that make a significant contribution to $\log \overline{G}_n$. We can state this precisely via the following main theorems:

Theorem 3.1.7. *Let* $A(n) = \sum_{p\leq n} \frac{2}{p-1} S_p(n) \log p$. *There is a constant* $c > 0$, *such that for* $n \geq 4$

$$A(n) = \left(\frac{3}{2} - \gamma\right) n^2 + O\left(n^2 \exp(-c\sqrt{\log n})\right)$$

where γ *denotes Euler's constant*

Theorem 3.1.8. *Let* $B(n) = \sum_{p\leq n} \frac{n-1}{p-1} d_p(n) \log p$. *There is a constant* $c > 0$, *such that for* $n \geq 4$

$$B(n) = (1-\gamma)n^2 + O\left(n^2 \exp(-c\sqrt{\log n})\right)$$

where γ *denotes Euler's constant.*

Theorem 3.1.9. *Let* $C(n) = \sum_{p\leq n} \frac{2\log p}{p-1} f_p(\log_p(n))$. *For a fixed integer* m, *for all* $n \geq 4$

$$C(n) = \left(\frac{1}{2} - \gamma\right) n - \sum_{k=1}^{m} k! \frac{n}{(\log n)^k} + O\left(2^{m+1}(m+1)! \frac{n}{(\log n)^{m+1}}\right)$$

where γ *denotes Euler's constant.*

We can also prove analogous theorems with better error terms when assuming the Riemann Hypothesis.

3.1.3 Outline of Proofs

We give a sketch of the series of steps used to prove the Theorems 3.1.7, 3.1.8 and 3.1.9.

First we find asymptotics for B_n. We break the proof of Theorem 3.1.8: that

$$B(n) = (1-\gamma)n^2 + O\left(n^2 e^{-\frac{c}{2}\sqrt{\log n}}\right)$$

into the following series of steps:

1. We show that the contribution of primes that are small relative to n (i.e. primes that are no more than $\sqrt{n}$) to $B(n)$ is negligible.
2. We then simplify the expression for $B(n)$, expressing $d_p(n)$ purely in terms of elementary functions. This is done by splitting the interval $(\sqrt{n}, n]$ into smaller subintervals, where on each subinterval, $d_p(n)$ has a simple expression. We then sum over all of the subintervals to get $B(n)$.
3. Next we write the expression for $B(n)$ as the difference of two sums of elementary functions. By invoking the Prime Number Theorem [9], we analyse each of these sums in turn, getting main terms $\frac{1}{2}n^2 \log n$ and $\frac{1}{2}n^2 \log n + (\gamma - 1)n^2$ respectively and an error term that is in each case $o(n^2)$. The theorem is proved by taking the difference of the sums.

To prove the asymptotics for $A(n)$ and $C(n)$: 3.1.7: that

$$A(n) = \left(\frac{3}{2} - \gamma\right) n^2 + O\left(n^2 \exp(-c\sqrt{\log n})\right)$$

and

$$C(n) = \left(\frac{1}{2} - \gamma\right) n - \sum_{k=1}^{m} k! \frac{n}{(\log n)^k} + O\left(2^{m+1}(m+1)! \frac{n}{(\log n)^{m+1}}\right)$$

we use the following series of steps:

1. We can deduce the asymptotics for $A(n)$ from the asymptotics for $B(n)$ because $\log \overline{G}_n = A(n) - B(n)$. At this stage, we'll have asymptotics for both $\log \overline{G}_n$ and B_n from Theorems 3.1.5 and 3.1.8.
2. We break down $A(n)$ into $A(n) = \pi(n) n \log n - nC(n)$.
3. We invoke the prime number theorem to write $\pi(n) n \log n$ in terms of the logarithmic integral $\int_2^n \frac{1}{\log t}\, dt$, which we can analyse via integration by parts, bounding the error terms. This will give us asymptotics for $C(n)$.

3.1.4 Main results over general non-prime base b

As an extension problem, we consider the analogous functions $A'(n)$, $B'(n)$ and $C'(n)$, when instead of restricting to primes p, we sum over all integers $b \leq n$. Precisely, they're defined as

Definition 3.1.10.

$$A'(n) = \sum_{2 \leq b \leq n} \frac{2 \log b}{b-1} S_b(n) \qquad B'(n) = \sum_{2 \leq b \leq n} \frac{n-1}{b-1} d_b(n) \log b$$

and

$$C'(n) = \sum_{2 \leq b \leq n} \frac{2 \log b}{b-1} f_b(\log_b(n)).$$

We can prove the following theorems for the asymptotic expansions of $A'(n)$, $B'(n)$ and $C'(n)$:

Theorem 3.1.11. *For all $n \geq 4$*

$$A'(n) = \left(\frac{3}{2} - \gamma\right) n^2 \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right) n^2 + O(n^{7/4}).$$

Theorem 3.1.12. *For all $n \geq 4$*

$$B'(n) = (1 - \gamma)\, n^2 \log n + (\gamma + \alpha - 1)\, n^2 + O(n^{7/4}).$$

Theorem 3.1.13. *For all* $n \geq 4$

$$C'(n) = \left(\frac{1}{2} - \gamma\right) n \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right) n + O(n^{3/4}).$$

Here, as before, γ denotes Euler's constant. α is the first Stieltjes constant, analysed in [21] among other places, with numerical value approximately equal to -0.0728. It's defined by the following limit:

$$\alpha := \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{\log k}{k} - \int_1^n \frac{\log t}{t}\, dt \right). \tag{3.1.3}$$

It's also of note that if we look at the Laurent Series expansion of the Riemann Zeta function around $s = 1$ we receive

$$\zeta(s) = \frac{1}{s-1} + \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} c_n (s-1)^n$$

where $c_0 = \gamma$ (Euler's constant) and $c_1 = \alpha$ (the first Stieltjes constant). For $n \geq 2$ each c_n has value equal to the nth Stieltjes constant defined by

$$c_n := \lim_{m\to\infty} \left(\sum_{k=1}^{m} \frac{(\log k)^n}{k} - \frac{(\log m)^{n+1}}{n+1} \right).$$

3.2 A Preliminary Reduction

Let $C(n) = \sum_{p\leq n} \frac{2\log p}{p-1} f_p(\log_p(n))$.

Lemma 3.2.1. $A(n) = \pi(n) n \log n + nC(n)$.

Proof. We invoke Delange's Theorem (Theorem 3.1.2).

$$\begin{aligned}
A(n) &= \sum_{p\le n} \frac{2\log p}{p-1} S_p(n) \\
&= \sum_{p\le n} \frac{2\log p}{p-1}\left(\frac{p-1}{2}\right) n\frac{\log n}{\log p} + n\sum_{p\le n}\frac{2\log p}{p-1} f_p(\log_p(n)) \\
&= n\log n \sum_{p\le n} 1 + n\sum_{p\le n}\frac{2\log p}{p-1} f_p(\log_p(n)) \\
&= \pi(n) n\log n + n\sum_{p\le n}\frac{2\log p}{p-1} f_p(\log_p(n)) \\
&= \pi(n) n\log n + nC(n).
\end{aligned}$$

□

3.3 Asymptotics for $B(n)$

3.3.1 Asymptotics for the sum over small primes

To find asymptotics for the function $B(n) = \sum_{p\le n}\frac{n-1}{p-1} d_p(n)\log p$, our first step is to notice that small values of p (relative to n), give small contributions that do not affect the main term in the asymptotics.

Lemma 3.3.1.

$$\sum_{p\le \sqrt{n}} \frac{n-1}{p-1} d_p(n)\log p = O(n^{3/2}).$$

Proof. We first note for each prime p that if $n = \sum_{i=0}^{k} a_i(n)p^i$ with each $a_i \le p-1$ and

$$k \le \left\lfloor \frac{\log(n)}{\log p} \right\rfloor + 1 \le \frac{\log(n)}{\log p} + 1$$

then

$$\sum_{p\leq\sqrt{n}} \frac{n-1}{p-1} d_p(n)\log p \leq \sum_{p\leq\sqrt{n}} \frac{n-1}{p-1}(p-1)\left(\frac{\log n}{\log p}+1\right)\log p$$
$$= \sum_{p\leq\sqrt{n}} (n-1)\left(\frac{\log n}{\log p}+1\right)\log p$$
$$= \sum_{p\leq\sqrt{n}} (n-1)\log(n) + \sum_{p\leq\sqrt{n}} (n-1)\log(p).$$

We can deal with the first sum using the prime number theorem (Theorem 1.11.3):

$$\sum_{p\leq\sqrt{n}} (n-1)\log(n) = (n-1)\log(n)(\pi(\sqrt{n}))$$
$$= (n-1)\log(n)\left(\frac{\sqrt{n}}{\log\sqrt{n}} + O\left(\exp\left(-a\sqrt{\log\sqrt{n}}\right)\right)\right)$$
$$= 2\sqrt{n}(n-1) + (n-1)(\log n)O(n^{\frac{1}{4}})$$
$$= O(n^{3/2}).$$

And since we have that

$$\sum_{p\leq\sqrt{n}} (n-1)\log(n) \geq \sum_{p\leq\sqrt{n}} (n-1)\log(p)$$

we know that $\sum_{p\leq\sqrt{n}} \frac{n-1}{p-1} d_p(n)\log p = O(n^{3/2})$. □

So we have the result that $B(n) = \sum_{\sqrt{n}\leq p\leq n} \frac{n-1}{p-1} d_p(n)\log p + O(n^{3/2})$.

3.3.2 Writing the sum in terms of elementary functions

$d_p(n)$ is much simpler for primes in the interval $(\sqrt{n}, n]$ than over the whole interval $[1, n]$ because it is the sum of at most 2 digits (since if $p > \sqrt{n}$ then $p^2 > n$), so $n = a + bp$ for some $0 \leq a, b \leq p-1$.)

We will split the interval $(\sqrt{n}, n]$ into blocks $(\frac{n}{j+1}, \frac{n}{j}]$ for each $j \leq \sqrt{n}$. This will allow us to express $d_p(n)$ in terms of elementary functions on each block. We can then analyse the sum of these functions more easily than we could $d_p(n)$ itself.

Lemma 3.3.2. *We have that*

$$\sum_{\sqrt{n}\leq p\leq n} \frac{n-1}{p-1} d_p(n) \log p = \sum_{\sqrt{n}\leq p\leq n} \frac{n(n-1)}{p-1} \log p - (n-1) \sum_{j=1}^{\sqrt{n}-1} j \left[\sum_{\frac{n}{j+1}<p\leq\frac{n}{j}} \log p \right].$$

Proof. First we suppose $\frac{n}{2} < p \leq n$. We notice at first that we must have $b > 0$ since if $b = 0$ then $n = a \leq p - 1 \leq n - 1$ which can't happen. And if $b \geq 2$ then $n \geq 2p$ so $\frac{n}{2} \geq p$ contradicting the fact that $\frac{n}{2} < p \leq n$. So we must have that $b = 1$ and $d_p(n) = 1 + (n - p) = n - (p - 1)$.

Now we suppose $\frac{n}{3} < p \leq \frac{n}{2}$. This time we notice that we must have $b > 1$ since if $b \leq 1$ then $n \leq (p - 1) + p = 2p - 1 \leq n - 1$. And if $b \geq 3$ then $n \geq 3p$ so $\frac{n}{3} \geq p$ contradicting the fact that $\frac{n}{3} < p \leq \frac{n}{2}$. So we must have that $b = 2$ and $d_p(n) = 2 + (n - 2p) = n - 2(p - 1)$.

We continue inductively to see that on each interval $(\frac{n}{j+1}, \frac{n}{j}]$, $d_p(n) = n - j(p-1)$. So the result follows. □

Definition 3.3.3. *Let's call*

$$B_1(n) = \sum_{\sqrt{n}\leq p\leq n} \frac{n(n-1)}{p-1} \log p, \qquad B_2(n) = (n-1) \sum_{j=1}^{\sqrt{n}-1} j \left[\sum_{\frac{n}{j+1}<p\leq\frac{n}{j}} \log p \right].$$

We see that to find asymptotics for $\sum_{\sqrt{n}\leq p\leq n} \frac{n-1}{p-1} d_p(n) \log p$, it suffices to find asymptotics for $B_1(n)$ and $B_2(n)$.

3.3.3 Asymptotics for $B_1(n)$

We will find an expression for $B_1(n) = \sum_{\sqrt{n}\leq p\leq n} \frac{n(n-1)}{p-1} \log p$ that consists of a main term and an error term, where the error term is $o(n^2)$. To do this, we'll analyse $B_1(n)$ by using partial summation and invoking the Prime Number Theorem (Theorem 1.11.3).

Theorem 3.3.4. $B_1(n) = \frac{1}{2}n^2\log(n) + O(n^2 e^{-\frac{c}{2}\sqrt{\log n}})$ *for an absolute constant* $c > 0$

We'll break the proof of Theorem 3.3.4 into a series of steps, first factoring out the $n(n-1)$ and considering $\sum_{\sqrt{n}\le p\le n}\frac{\log p}{p-1}$. We see that it will be easier to use summation by parts on $\sum_{\sqrt{n}\le p\le n}\frac{\log p}{p}$ than on $\sum_{\sqrt{n}\le p\le n}\frac{\log p}{p-1}$, so our first step will be to see that

$$\sum_{\sqrt{n}\le p\le n}\frac{\log p}{p-1} = \sum_{\sqrt{n}\le p\le n}\frac{\log p}{p} + \left(\frac{\log p}{p-1} - \frac{\log p}{p}\right)$$

and that the difference

$$\sum_{\sqrt{n}\le p\le n}\frac{\log p}{p-1} - \sum_{\sqrt{n}\le p\le n}\frac{\log p}{p}$$

is negligible.

Lemma 3.3.5. *We have that*

$$\sum_{\sqrt{n}\le p\le n}\frac{\log p}{p-1} - \sum_{\sqrt{n}\le p\le n}\frac{\log p}{p} = O\left(\frac{\log n}{\sqrt{n}}\right).$$

Proof.

$$\sum_{\sqrt{n}\le p\le n}\left(\frac{\log p}{p-1} - \frac{\log p}{p}\right) = \sum_{\sqrt{n}\le p\le n}\frac{\log p}{p(p-1)} \le \sum_{\sqrt{n}\le m\le n}\frac{\log m}{m(m-1)}$$

where we sum over all integers m between $\sqrt{n}$ and n.

We consider the following function and its derivative:

$$f(x) = \frac{\log x}{x(x-1)}, \qquad f'(x) = \frac{x - 1 - 2x\log x + \log x}{x^2(x-1)^2}.$$

Note $f'(x)$ is negative for large enough x because of the $-2x\log x$ term.

Therefore if we consider large enough n,

$$\begin{aligned}\sum_{\sqrt{n}\le m\le n}\frac{\log m}{m(m-1)} &\le \log n\int_{\sqrt{n}-1}^{n+1}\frac{1}{x(x-1)}\,dx \\ &\le \log n\int_{\sqrt{n}-1}^{n+1}\frac{1}{(x-1)^2}\,dx = O\left(\frac{\log n}{\sqrt{n}}\right).\end{aligned}$$

□

Now we look at the function $\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p}$. In the next Lemma, we use integration by parts and the Prime Number Theorem to gain an understanding of its asymptotic behaviour. Specifically we will see that this function behaves like $\frac{1}{2}\log n$. We will then use this information and look at the difference: $\left[\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p}\right]-\frac{1}{2}\log n$ to obtain a better asymptotic for $\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p}$ with a smaller error term.

Lemma 3.3.6. $\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p}=\frac{1}{2}\log(n)-\frac{1}{\log(n)}+O\left(\frac{1}{(\log n)^2}\right).$

Proof. By partial summation

$$\begin{aligned}\sum_{\sqrt{n}\leq p\leq n}\frac{\log p}{p}&=\int_{\sqrt{n}}^{n}\frac{\log t}{t}\,d\pi(t)\\&=\pi(n)\frac{\log n}{n}-\pi(\sqrt{n})\frac{\log(\sqrt{n})}{\sqrt{n}}-\int_{\sqrt{n}}^{n}\pi(t)\frac{d}{dt}\left(\frac{\log t}{t}\right)\,dt.\end{aligned}$$

By Prime Number Theorem, we have that $\pi(x)=\frac{x}{\log x}+\frac{x}{(\log x)^2}+O(\frac{x}{(\log x)^3})$. So

$$\begin{aligned}&\pi(n)\frac{\log n}{n}-\pi(\sqrt{n})\frac{\log(\sqrt{n})}{\sqrt{n}}\\&=\left(\frac{n}{\log n}+\frac{n}{(\log n)^2}+O\left(\frac{n}{(\log n)^3}\right)\right)\frac{\log n}{n}\\&\quad-\left(\frac{\sqrt{n}}{\log\sqrt{n}}+\frac{\sqrt{n}}{(\log\sqrt{n})^2}+O\left(\frac{\sqrt{n}}{(\log\sqrt{n})^3}\right)\right)\frac{\log\sqrt{n}}{\sqrt{n}}\\&=(1-1)+\left(\frac{1}{\log n}-\frac{2}{\log n}\right)+O\left(\frac{1}{(\log n)^2}\right)\\&=-\frac{1}{\log n}+O\left(\frac{1}{(\log n)^2}\right).\end{aligned}$$

This implies

$$\sum_{\sqrt{n}\leq p\leq n}\frac{\log p}{p}=-\frac{1}{\log n}+O\left(\frac{1}{(\log n)^2}\right)-\int_{\sqrt{n}}^{n}\pi(t)\frac{d}{dt}\left(\frac{\log t}{t}\right)\,dt.$$

So now we analyse $\int_{\sqrt{n}}^{n}\pi(t)\frac{d}{dt}\left(\frac{\log t}{t}\right)dt$. Again using the Prime Number Theorem, this time with $t=x$, we get:

$$\int_{\sqrt{n}}^{n}\pi(t)\frac{d}{dt}\left(\frac{\log t}{t}\right)\,dt=\int_{\sqrt{n}}^{n}\left(\frac{t}{\log t}+\frac{t}{(\log t)^2}+O\left(\frac{t}{(\log t)^3}\right)\right)\frac{d}{dt}\left(\frac{\log t}{t}\right)\,dt.$$

We can do the first integral by inspection:

$$\begin{aligned}\int_{\sqrt{n}}^{n} \frac{t}{\log t} \frac{d}{dt}\left(\frac{\log t}{t}\right)\, dt &= \log\left(\frac{\log t}{t}\right)\Bigg|_{\sqrt{n}}^{n} \\ &= \log(\frac{\log n}{n}\frac{\sqrt{n}}{\frac{1}{2}\log n}) \\ &= \log(\frac{2}{\sqrt{n}}) \\ &= \log 2 - \frac{1}{2}\log(n).\end{aligned}$$

Next we use $\frac{d}{dt}(\frac{\log t}{t}) = \frac{1-\log t}{t^2}$ to do the other integrals:

$$\begin{aligned}\int_{\sqrt{n}}^{n} \frac{t}{(\log t)^2} \frac{d}{dt}\left(\frac{\log t}{t}\right)\, dt &= \int_{\sqrt{n}}^{n} \frac{t}{(\log t)^2}\frac{1-\log t}{t^2}\, dt \\ &= \int_{\sqrt{n}}^{n} \left(\frac{1}{t(\log t)^2} - \frac{1}{t\log t}\right)\, dt \\ &= \left(-\frac{1}{\log t} - \log\log t\right)\Bigg|_{\sqrt{n}}^{n} \\ &= -\frac{1}{\log n} - \log\log n + \frac{2}{\log n} + \log\log\sqrt{n} \\ &= \frac{1}{\log n} - \log 2.\end{aligned}$$

We treat $\int_{\sqrt{n}}^{n} \frac{t}{(\log t)^3}\frac{d}{dt}(\frac{\log t}{t})dt$ similarly.

$$\begin{aligned}\int_{\sqrt{n}}^{n} \frac{t}{(\log t)^3} \frac{d}{dt}\left(\frac{\log t}{t}\right)\, dt &= \int_{\sqrt{n}}^{n} \frac{t}{(\log t)^3}\frac{1-\log t}{t^2}\, dt \\ &= \int_{\sqrt{n}}^{n} \left(\frac{1}{t(\log t)^3} - \frac{1}{t(\log t)^2}\right)\, dt \\ &= \left(-\frac{1}{2}\frac{1}{(\log t)^2} + \frac{1}{\log t}\right)\Bigg|_{\sqrt{n}}^{n} \\ &= -\frac{1}{2}\frac{1}{(\log n)^2} + \frac{1}{\log n} + \frac{1}{2}\frac{1}{(\log\sqrt{n})^2} - \frac{2}{\log n} \\ &= -\frac{1}{\log n} + O\left(\frac{1}{(\log n)^2}\right).\end{aligned}$$

So finally we add all of these to conclude

$$\int_{\sqrt{n}}^{n} \pi(t)\frac{d}{dt}\left(\frac{\log t}{t}\right)\, dt = \left(\log 2 - \frac{1}{2}\log n\right) + \left(\frac{1}{\log n} - \log 2\right) - \left(\frac{1}{\log n} + O(\frac{1}{(\log n)^2}\right)$$
$$= -\frac{1}{2}\log n + O\left(\frac{1}{(\log n)^2}\right).$$

So

$$\sum_{\sqrt{n}\leq p\leq n} \frac{\log p}{p} = \frac{1}{2}\log n - \frac{1}{\log n} + O\left(\frac{1}{(\log n)^2}\right).$$

□

Definition 3.3.7. *We define $\psi(n)$ the second Chebyshev function by $\psi(n) = \sum_{p^k\leq n} \log p$ where the sum is over all prime powers less than or equal to n*

Our goal is to use the second Chebyshev function $\psi(n)$ which we have better asymptotics for than the prime counting function $\pi(n)$ to provide us with a better asymptotic for $\sum_{\sqrt{n}<p\leq n} \frac{\log p}{p}$. We know from Lemma 3.3.6 that $\sum_{\sqrt{n}<p\leq n} \frac{\log p}{p} \approx \frac{1}{2}\log n$. Our idea is to look at the difference $\left[\sum_{\sqrt{n}<p\leq n} \frac{\log p}{p}\right] - \frac{1}{2}\log n$ and to apply summation by parts to bound it.

Proposition 3.3.8. *([17], section 2.1, page 63) (Summation by parts formula) Given two sequences $\{a_m\}$ and $\{b_m\}$, if we define $A_n = \sum_{m=1}^{n} a_m$, then we have the following formula:*

$$\sum_{m=k}^{n} a_m b_m = A_n b_n - A_{k-1}b_{k-1} - \sum_{m=k-1}^{n-1} A_m(b_{m+1} - b_m).$$

We rewrite $\left[\sum_{\sqrt{n}<p\leq n} \frac{\log p}{p}\right] - \frac{1}{2}\log n$ in such a way that this formula is applicable. We expect $\sum_{\sqrt{n}<p\leq n} \log p \approx \frac{1}{2}\log n$. So we consider

$$\left[\sum_{\sqrt{n}<p\leq n} \frac{\log p}{p}\right] - \frac{1}{2}\log n = \left[\sum_{\sqrt{n}<m\leq n} \frac{\log m}{m}\mathbb{I}_P(m)\right] - \frac{1}{2}\log n$$

where $\mathbb{I}_P(m)$ is the characteristic function of the set of primes and we sum over all integers inside $(\sqrt{n}, n]$. But

$$\left[\sum_{\sqrt{n}<m\leq n} \frac{\log m}{m}\mathbb{I}_P(m)\right] - \frac{1}{2}\log n = \sum_{\sqrt{n}<m\leq n} \left(\frac{\log m}{m}\mathbb{I}_P(m) - \frac{1}{m}\right) + O(1)$$
$$= \sum_{\sqrt{n}<m\leq n} \frac{1}{m}\left((\log m)\mathbb{I}_P(m) - 1\right) + O(1).$$

We can set $a_m = (\log m)\mathbb{I}_P(m) - 1$, $b_m = \frac{1}{m}$ and $k = \lfloor\sqrt{n}\rfloor + 1$ in this last expression and apply the summation by parts formula given in Definition 3.3.8 to find asymptotics for the difference $\left[\sum_{\sqrt{n}<p\leq n} \frac{\log p}{p}\right] - \frac{1}{2}\log n$.

Lemma 3.3.9.

$$\sum_{m\leq n}((\log m)\mathbb{I}_P(m) - 1) = \sum_{p\leq n}\log p - n = O\left(ne^{-c\sqrt{\log n}}\right).$$

Proof. We look at the second Chebyshev function $\psi(n)$ [26]

$$\psi(n) := \sum_{p^k\leq n}\log p \qquad \text{(where the sum is over all } p\text{th powers less than or equal to } n\text{)}$$
$$= \sum_{p\leq n}\log p + \sum_{\substack{p^k\leq n\\ k\geq 2}}\log p$$
$$= \sum_{p\leq n}\log p + O\left(\sum_{p\leq\sqrt{n}}\log p\right)$$
$$= \sum_{p\leq n}\log p + O\left(\sqrt{n}\log n\right).$$

We know by Prime Number Theorem that $\psi(n) - n = O\left(ne^{-c\sqrt{\log n}}\right)$. So putting these together we get that

$$\sum_{p\leq n}\log p = n + O\left(\sqrt{n}\log n\right) + O\left(ne^{-c\sqrt{\log n}}\right)$$

so that

$$\sum_{p\leq n} \log p - n = O\left(ne^{-c\sqrt{\log n}}\right).$$

□

Proof. of Theorem 3.3.4

So far, we've shown that $B_1(n) = n(n-1)\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p} + O\left(\frac{\log n}{\sqrt{n}}\right)$. We've also shown that the sum $\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p}$ is asymptotically equivalent to the function $\frac{1}{2}\log n$. All that remains to do is prove that the difference between $\sum_{\sqrt{n}<p\leq n}\frac{\log p}{p}$ and $\frac{1}{2}\log n$ is bounded by $O\left(e^{-c\sqrt{\log n}}\right)$. In the work above we've just rewritten this difference as

$$\sum_{\sqrt{n}<m\leq n}((\log m)\mathbb{I}_P(m) - 1)\frac{1}{m}$$

on which we'll use the summation by parts formula from Definition 3.3.8. In the formula, we set $a_m = (\log m)\mathbb{I}_P(m) - 1$, $b_m = \frac{1}{m}$ and $k = \lfloor\sqrt{n}\rfloor + 1$.

$$\begin{aligned}
&\sum_{\sqrt{n}<m\leq n}((\log m)\mathbb{I}_P(m) - 1)\frac{1}{m} \\
&\quad= \frac{1}{n}\sum_{m=1}^{n}((\log m)\mathbb{I}_P(m) - 1) - \frac{1}{\sqrt{n}}\left(\sum_{m=1}^{\sqrt{n}}(\log m\mathbb{I}_P(m) - 1)\right) \\
&\qquad\qquad - \sum_{j=\sqrt{n}}^{n-1}\left(\frac{1}{j+1} - \frac{1}{j}\right)\sum_{m=1}^{j}(\log m\mathbb{I}_P(m) - 1) + O\left(\frac{1}{\sqrt{n}}\right) \\
&\quad= O\left(\frac{1}{n}ne^{-c\sqrt{\log n}}\right) + O\left(\frac{1}{\sqrt{n}}\sqrt{n}e^{-c\sqrt{\log\sqrt{n}}}\right) \\
&\qquad\qquad + O\left(\sum_{j=\sqrt{n}}^{n-1}\frac{1}{j(j+1)}je^{-c\sqrt{\log j}}\right) + O\left(\frac{1}{\sqrt{n}}\right) \\
&\quad= O\left(e^{-c\sqrt{\log n}}\right) + O\left(e^{-c\sqrt{\log\sqrt{n}}}\right) + O\left(\sum_{j=\sqrt{n}}^{n-1}\frac{1}{j}e^{-c\sqrt{\log j}}\right) \\
&\quad= O\left(e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right) + O\left(\sum_{j=\sqrt{n}}^{n-1}\frac{1}{j}e^{-c\sqrt{\log j}}\right).
\end{aligned}$$

Now we see that

$$\sum_{j=\sqrt{n}}^{n-1} \frac{1}{j} e^{-c\sqrt{\log j}} = O\left(\int_{\frac{\sqrt{n}}{2}}^{n} \frac{1}{x} e^{-c\sqrt{\log x}}\, dx\right).$$

We can then evaluate this integral using the substitution $y = \sqrt{\log x}$ so that

$$\frac{dy}{dx} = \frac{1}{2x\sqrt{\log x}} = \frac{1}{2xy}$$

and

$$\begin{aligned}
\int_{\frac{\sqrt{n}}{2}}^{n} \frac{1}{x} e^{-c\sqrt{\log x}}\, dx &= \int_{\sqrt{\log \frac{\sqrt{n}}{2}}}^{\sqrt{\log n}} \frac{1}{x} 2xye^{-cy}\, dy \\
&= \int_{\sqrt{\log \frac{\sqrt{n}}{2}}}^{\sqrt{\log n}} 2ye^{-cy}\, dy \\
&= \frac{-2}{c} ye^{-cy} \Big|_{\sqrt{\log \frac{\sqrt{n}}{2}}}^{\sqrt{\log n}} + \int_{\sqrt{\log \frac{\sqrt{n}}{2}}}^{\sqrt{\log n}} \frac{2}{c} e^{-cy}\, dy \\
&= O\left(\sqrt{\log n} e^{-c\sqrt{\log \frac{1}{2}\sqrt{n}}} + \frac{2}{c^2} e^{-c\sqrt{\log \frac{\sqrt{n}}{2}}}\right) \\
&= O\left(\sqrt{\log n} e^{-c\sqrt{\log \frac{\sqrt{n}}{2}}}\right) \\
&= O\left(e^{\log(\sqrt{\log n}) - c\sqrt{\log \frac{\sqrt{n}}{2}}}\right) \\
&= O\left(e^{\frac{1}{2}\log\log n - c\sqrt{\frac{1}{2}\log n - \log \frac{1}{2}}}\right) \\
&= O\left(e^{-\frac{c}{2}\sqrt{\log n}}\right)
\end{aligned}$$

so we receive our desired error term and conclude our proof of the Theorem. □

3.3.4 Asymptotics for $B_2(n)$

We need to analyse

$$B_2(n) = (n-1) \sum_{j=1}^{\sqrt{n}-1} j \left[\sum_{\frac{n}{j+1} < p \leq \frac{n}{j}} \log p \right].$$

To do this, we simplify $B_2(n)$ and write it in terms of Chebyshev summatory functions, which because of the Prime Number Theorem, we know asymptotics for.

Theorem 3.3.10. *We have that*

$$(n-1)\sum_{j=1}^{\sqrt{n}-1} j\left[\sum_{\frac{n}{j+1}<p\leq\frac{n}{j}} \log p\right] = \frac{1}{2}n^2\log n + (\gamma-1)n^2 + O\left(n^2 e^{-\frac{c}{2}\sqrt{\log n}}\right).$$

Proof.

$$\begin{aligned}\sum_{j=1}^{\sqrt{n}-1} j\left[\sum_{\frac{n}{j+1}<p\leq\frac{n}{j}} \log p\right] &= \sum_{j=1}^{\sqrt{n}-1} j\left(\theta\left(\frac{n}{j}\right)-\theta\left(\frac{n}{j+1}\right)\right)\\ &= \left(\sum_{j=1}^{\sqrt{n}-1}\theta\left(\frac{n}{j}\right)\right) - (\sqrt{n}-1)\theta(\sqrt{n})\end{aligned}$$

where $\theta(m) = \sum_{p\leq m}\log p$ is the first Chebyshev summatory function [26].

Consider the second Chebyshev summatory function $\psi(x)$ where

$$\begin{aligned}\psi(x) :&= \sum_{p^k\leq x}\log p \qquad \text{(where the sum is over all } p\text{th powers less than or equal to } x\text{)}\\ &= \sum_{p\leq x}\log p + \sum_{\substack{p^k\leq x\\ k\geq 2}}\log p.\end{aligned}$$

We look at the function ψ instead of θ because the Prime Number Theorem gives better asymptotics for it. We can relate ψ to θ in the following way for large x:

$$\theta(x) = \sum_{p\leq x}\log p = \psi(x) + O(\sqrt{x}\log x). \tag{3.3.1}$$

This is because

$$\begin{aligned}\psi(x) &= \sum_{p\leq x}\log p + \sum_{\substack{p^k\leq x\\ k\geq 2}}\log p\\ &= \theta(x) + O\left(\sqrt{x}\log x\right).\end{aligned}$$

So

$$\theta(x) = \sum_{p\leq x}\log p = \psi(x) + O(\sqrt{x}\log x).$$

Thus

$$\begin{aligned}\sum_{j=1}^{\sqrt{n}-1}\theta\left(\frac{n}{j}\right) &= \sum_{j=1}^{\sqrt{n}-1}\psi\left(\frac{n}{j}\right)+O\left(\sum_{j=1}^{\sqrt{n}}\sqrt{\frac{n}{j}}\log\frac{n}{j}\right)\\ &= \sum_{j=1}^{\sqrt{n}-1}\psi\left(\frac{n}{j}\right)+O\left(\sqrt{n}\log n\sum_{j=1}^{\sqrt{n}}\frac{1}{\sqrt{j}}\right)\\ &= \sum_{j=1}^{\sqrt{n}-1}\psi\left(\frac{n}{j}\right)+O\left(\sqrt{n}\log n\int_{1}^{\sqrt{n}}\frac{1}{\sqrt{t}}\,dt\right)\\ &= \sum_{j=1}^{\sqrt{n}-1}\psi\left(\frac{n}{j}\right)+O\left(n^{3/4}\log n\right).\end{aligned}$$

Now the Siegel- Walfisz Theorem ([26], page 255, Theorem $8.2.5$) gives $\psi(x)=x+O(xe^{-c\sqrt{\log x}})$ for some positive constant c. Since $1\le j\le\sqrt{n}$, we have that $\sqrt{n}\le\frac{n}{j}\le n$. So in particular for $1\le j\le\sqrt{n}$, we see that

$$O\left(\frac{n}{j}e^{-c\sqrt{\log(\frac{n}{j})}}\right)=O\left(\frac{n}{j}e^{-\frac{c}{\sqrt{2}}\sqrt{\log(n)}}\right).$$

This gives

$$\begin{aligned}\sum_{j=1}^{\sqrt{n}-1}\psi\left(\frac{n}{j}\right) &= \sum_{j=1}^{\sqrt{n}-1}\left[\frac{n}{j}+O\left(\frac{n}{j}e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right)\right]\\ &= n\left(\sum_{j=1}^{\sqrt{n}}\frac{1}{j}\right)+O\left(ne^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\sum_{j=1}^{\sqrt{n}}\frac{1}{j}\right).\end{aligned}$$

Given the asymptotics for the harmonic series (Theorem 1.11.2), we know that this last expression is equal to

$$n(\log\sqrt{n}+\gamma)+O\left(n\log(n)e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right).$$

Finally we write $(\sqrt{n}-1)\theta(\sqrt{n})$ also in terms of ψ using equation 3.3.1 to get

$$\begin{aligned}(\sqrt{n}-1)\theta(\sqrt{n}) &= (\sqrt{n}-1)(\psi(\sqrt{n})+O(n^{1/4}\log n))\\ &= (\sqrt{n}-1)(\sqrt{n}+O(\sqrt{n}e^{-c\sqrt{\log n}}))\\ &= n+O\left(ne^{-c\sqrt{\log n}}\right).\end{aligned}$$

So putting all this together, we have that

$$\begin{aligned}\left(\sum_{j=1}^{\sqrt{n}-1} \theta\left(\frac{n}{j}\right)\right) - (\sqrt{n}-1)\theta(\sqrt{n}) &= n(\log\sqrt{n}+\gamma) + O\left(n\log(n)e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right) \\ &\quad - \left(n + O\left(ne^{-c\sqrt{\log n}}\right)\right) \\ &= \frac{1}{2}n\log n + (\gamma-1)n + O\left(n\log(n)e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right).\end{aligned}$$

And so

$$\begin{aligned}(n-1)\sum_{j=1}^{\sqrt{n}-1} j\left[\sum_{\frac{n}{j+1}<p\le\frac{n}{j}} \log p\right] &= \frac{1}{2}n^2\log n + (\gamma-1)n^2 + O\left(n^2\log(n)e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right) \\ &= \frac{1}{2}n^2\log n + (\gamma-1)n^2 + O\left(n^2 e^{\log\log(n)-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right) \\ &= \frac{1}{2}n^2\log n + (\gamma-1)n^2 + O\left(n^2 e^{-\frac{c}{2}\sqrt{\log n}}\right)\end{aligned}$$

as desired. □

3.3.5 Proving the Asymptotics for $B(n)$

Combining the estimates for $B_1(n)$ and $B_2(n)$ given by Theorems 3.3.4 and 3.3.10, Theorem 3.1.8 is proved:

We have that

$$\begin{aligned}B(n) &= B_1(n) - B_2(n) \\ &= \left(\frac{1}{2}n^2\log(n) + O(n^2e^{-\frac{c}{2}\sqrt{\log n}})\right) \\ &\qquad - \left(\frac{1}{2}n^2\log n + (\gamma-1)n^2 + O\left(n^2\log(n)e^{-\frac{c}{\sqrt{2}}\sqrt{\log n}}\right)\right)) \\ &= (1-\gamma)n^2 + O\left(n^2e^{-\frac{c}{2}\sqrt{\log n}}\right).\end{aligned}$$

3.4 Asymptotics for $C(n)$

3.4.1 Asymptotics for $\pi(n)n\log n$

We recall the indicator function on primes to be

$$\mathbb{I}_P(k) = \begin{cases} 1 & \text{if } k \text{ is prime} \\ 0 & \text{otherwise} \end{cases}$$

and the prime counting function to be $\pi(n) := \sum_{k\leq n} \mathbb{I}_P(k)$. We also recall the logarithmic integral function

$$\mathrm{li}(n) = \int_2^n \frac{1}{\log t}\, dt.$$

This section is devoted to finding asymptotics for the expression $\pi(n)n\log n$, where our main theorem is the following:

Theorem 3.4.1. *For a fixed integer m and all $n \geq 4$,*

$$\pi(n)n\log n = n\log n \sum_{k=1}^{m}(k-1)!\frac{n}{(\log n)^k} + O\left(2^{m+1}(m+1)!\frac{n^2}{(\log n)^m}\right).$$

By the Prime Number Theorem (Theorem 1.11.3), we can connect the prime counting function to the logarithmic integral.

Lemma 3.4.2. *For all positive integers $n, \pi(n) = \mathrm{li}(n) + O(n\exp(-c\sqrt{\log n}))$.*

This tells us that asymptotics for $\mathrm{li}(n)$ will give us asymptotics for $\pi(n)$. We state and prove the following lemma which gives us asymptotics for $\mathrm{li}(x)$ (that holds for all x, not just integers).

Theorem 3.4.3. *For fixed m and all $x \geq 4$, we have the following asymptotics for $\mathrm{li}(x)$*

$$\mathrm{li}(x) = \sum_{k=1}^{m}(k-1)!\frac{x}{(\log x)^k} + O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right)$$

where the O-constant is absolute and independent of m.

To prove (1), we'll first break the integral up to get

$$\mathrm{li}(x) = \int_2^x \frac{1}{\log t}\,dt = \int_2^{\sqrt{x}} \frac{1}{\log t}\,dt + \int_{\sqrt{x}}^x \frac{1}{\log t}\,dt$$

and show that contribution from the short interval $[2, \sqrt{x}]$ is negligible. Next we'll integrate by parts over $[\sqrt{x}, x]$, plug in the upper bound to get the finite sum

$$\sum_{k=1}^{m} (k-1)! \frac{x}{(\log x)^k}$$

and plug in the lower bound to get an error of size $O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right)$.

Lemma 3.4.4. *The contribution* $\int_2^{\sqrt{x}} \frac{1}{\log t}\,dt$ *is* $O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right)$ *where the O-constant is absolute and independent of* m*.*

Proof. Bounding the integrand above by the constant $\frac{1}{\log 2}$, we have that

$$\int_2^{\sqrt{x}} \frac{1}{\log t}\,dt = O(\sqrt{x}).$$

We need to show that this is $O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right)$. To prove this claim, it's enough to show that

$$\sqrt{x} \le 2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}$$

for all x, which is the same as showing that

$$2^{m+1}(m+1)!\frac{\sqrt{x}}{(\log x)^{m+1}} \ge 1$$

for all x. To do this, we consider the following function $f(x)$ and its derivative:

$$f(x) = 2^{m+1}(m+1)!\frac{\sqrt{x}}{(\log x)^{m+1}}, \quad f'(x) = 2^{m+1}(m+1)!\frac{1}{\sqrt{x}(\log x)^{m+2}}\left[\frac{1}{2}\log x - (m+1)\right].$$

We see that $f'(x) = 0$ at exactly one point: $x = e^{2(m+1)}$, which is a global minimum. Plugging in this value of x, we have

$$\begin{aligned} f\left(e^{2(m+1)}\right) &= 2^{m+1}(m+1)!\frac{e^{m+1}}{(2(m+1))^{m+1}} \\ &= \frac{e^{m+1}(m+1)!}{(m+1)^{m+1}}. \end{aligned}$$

Now appealing to Stirling's approximation (Theorem 1.11.4), we can bound $(m+1)!$ below by $(m+1)^{(m+1)}e^{-(m+1)}\sqrt{2\pi(m+1)}$, giving us a lower bound of $\sqrt{2\pi(m+1)}$ for $f\left(e^{2(m+1)}\right)$, which is always greater than 1. Since we got this by plugging in the minimum value, we conclude as desired for all $x \geq 4$ that

$$2^{m+1}(m+1)!\frac{\sqrt{x}}{(\log x)^{m+1}} \geq 1.$$

□

We now state and prove two more lemmas. In our first lemma, we evaluate the integral $\int_{\sqrt{x}}^{x} \frac{1}{\log t}\,dt$, getting our desired main term, one other term and our stated error term. In our second lemma, we'll show that the extra term we got from the first lemma can also be absorbed into the stated error term.

Lemma 3.4.5. *We have that*

$$\int_{\sqrt{x}}^{x} \frac{1}{\log t}\,dt = \sum_{k=1}^{m}(k-1)!\frac{x}{(\log x)^k} - \sum_{k=1}^{m}(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} + O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right)$$

where the O-constant is absolute and independent of m.

Proof. Applying integration by parts m times, we arrive at

$$\int_{\sqrt{x}}^{x} \frac{1}{\log t}\,dt = \sum_{k=1}^{m}(k-1)!\frac{x}{(\log x)^k} - \sum_{k=1}^{m}(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} + \int_{\sqrt{x}}^{x}\frac{m!}{(\log t)^{m+1}}\,dt.$$

To get the stated result, we just need to bound $\int_{\sqrt{x}}^{x}\frac{m!}{(\log t)^{m+1}}\,dt$. This is done by

$$\begin{aligned}\int_{\sqrt{x}}^{x}\frac{m!}{(\log t)^{m+1}}\,dt &\leq m!\int_{\sqrt{x}}^{x}\frac{1}{(\log\sqrt{x})^{m+1}}\,dt\\ &= 2^{m+1}m!\int_{\sqrt{x}}^{x}\frac{1}{(\log x)^{m+1}}\,dt\\ &= 2^{m+1}m!\frac{x-\sqrt{x}}{(\log x)^{m+1}}\\ &= O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right).\end{aligned}$$

□

Lemma 3.4.6. *We have that*

$$\sum_{k=1}^{m}(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} = O\left(2^{m+1}(m+1)!\frac{x}{(\log x)^{m+1}}\right).$$

Proof. It suffices to show for every k that $(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} \leq 2^{m+1}m!\frac{x}{(\log x)^{m+1}}$. Summing over k from 1 to m on both sides will then give the desired result. This is the same as showing that

$$2^k(k-1)!\frac{\sqrt{x}}{(\log x)^k} \leq 2^{m+1}m!\frac{x}{(\log x)^{m+1}}$$

which is the same as showing

$$2^{m-k+1}\frac{m!}{(k-1)!}\frac{\sqrt{x}}{(\log x)^{m-k+1}} \geq 1.$$

Using the same trick as in the last lemma, we take the derivative. We see that the function $f(x) = 2^{m-k+1}\frac{m!}{(k-1)!}\frac{\sqrt{x}}{(\log x)^{m-k+1}}$ has a minimum at $x = e^{2(m-k+1)}$. Plugging in this minimum value, we see that showing $f(x) \geq 1$ amounts to showing that

$$2^{m-k+1}\frac{m!}{(k-1)!}\frac{e^{m-k+1}}{2^{m-k+1}(m-k+1)^{m-k+1}} \geq 1. \tag{3.4.1}$$

Again invoking Stirling's approximation (Theorem 1.11.4), we have that

$$(m-k+1)^{m-k+1} \leq \frac{(m-k+1)!e^{m-k+1}}{\sqrt{2\pi(m-k+1)}}$$

so to show that (3.4.1) is true, it suffices to show that

$$\frac{m!}{(k-1)!}\frac{\sqrt{2\pi(m-k+1)}}{(m-k+1)!} \geq 1.$$

However this last expression is equal to $\binom{m}{k-1}\sqrt{2\pi(m-k+1)}$, which is a product of terms which are all at least 1, proving the lemma. □

The last three lemmas together prove Theorem 3.4.3 and also Theorem 3.4.1.

3.4.2 Getting the Asymptotics for $C(n)$

Using the results of Theorems 3.1.7 3.1.8 and 3.4.1 we arrive at the asymptotics for $C(n)$, stated in Theorem 3.1.9.

Proof. Using Theorem 3.1.5 and Lemma 3.2.1 we have that

$$\begin{aligned}\frac{1}{2}n^2 + O(n\log n) &= A(n) - B(n)\\ &= \pi(n)n\log n + nC(n) - B(n)\\ &= \pi(n)n\log n + nC(n) - (1-\gamma)n^2 + O\left(n^2 e^{-\frac{c}{2}\sqrt{\log n}}\right)\\ &\qquad \text{(by Theorem 3.1.8).}\end{aligned}$$

So

$$\begin{aligned}nC(n) &= \frac{1}{2}n^2 - \pi(n)n\log n + (1-\gamma)n^2 + O\left(n^2 e^{-\frac{c}{c}\sqrt{\log n}}\right)\\ &= \left(\frac{3}{2}-\gamma\right)n^2 - n^2\log n\sum_{k=1}^{m+1}\frac{(k-1)!}{(\log n)^k} + O\left(2^{m+1}(m+1)!\frac{n^2}{(\log n)^{m+1}}\right)\\ &\quad + O\left(n^2 e^{-\frac{c}{2}\sqrt{\log n}}\right)\\ &= \left(\frac{1}{2}-\gamma\right)n^2 - n^2\sum_{k=1}^{m}\frac{k!}{(\log n)^k} + O\left(2^{m+1}(m+1)!\frac{n^2}{(\log n)^{m+1}}\right).\end{aligned}$$

□

3.5 An improved error term assuming the Riemann Hypothesis

If we assume the Riemann Hypothesis then we can improve the errors in our asymptotic formulas for $A(n)$, $B(n)$ and $C(n)$. Instead of only saving a log term, we're able to save power functions.

3.5.1 Improved Asymptotics for $A(n)$ and $B(n)$

On assuming the Riemann Hypothesis, we get the following asymptotic formula with a small error term for the first Chebyshev function from [25], Theorem 10.

$$\theta(n) := \sum_{p \leq n} \log p = n + O\left(\sqrt{x}(\log x)^2\right). \tag{3.5.1}$$

We recall that $B(n) = B_1(n) - B_2(n)$ with

$$B_1(n) = \sum_{\sqrt{n} \leq p \leq n} \frac{n(n-1)}{p-1} \log p, \qquad B_2(n) = (n-1) \sum_{j=1}^{\sqrt{n}-1} j \left[\sum_{\frac{n}{j+1} < p \leq \frac{n}{j}} \log p \right].$$

Lemma 3.5.1. *Assuming the Riemann Hypothesis*

$$B_1(n) = \frac{1}{2} n^2 \log n + O\left(n^{7/4} (\log n)^2\right).$$

Proof. We analyse $\sum_{\sqrt{n} \leq p \leq n} \frac{\log p}{p-1}$. We showed in Lemma 3.3.5 that

$$\sum_{\sqrt{n} \leq p \leq n} \frac{\log p}{p-1} = \sum_{\sqrt{n} \leq p \leq n} \frac{\log p}{p} + O\left(\frac{\log n}{\sqrt{n}}\right).$$

Next we see

$$\begin{aligned}
\sum_{\sqrt{n} \leq p \leq n} \frac{\log p}{p} &= \int_{\sqrt{n}}^{n} \frac{1}{x} d\theta(x), \qquad \text{where } \theta(x) = \sum_{p \leq x} \log p \\
&= \frac{1}{n}\theta(n) - \frac{1}{\sqrt{n}}\theta(\sqrt{n}) - \int_{\sqrt{n}}^{n} \theta(x) \frac{d}{dx}\left(\frac{1}{x}\right) dx \\
&= O\left(\frac{(\log n)^2}{n^{1/2}} + \frac{(\log n)^2}{n^{1/4}}\right) + \int_{\sqrt{n}}^{n} \frac{\theta(x)}{x^2} dx \qquad \text{by (3.5.1)} \\
&= \int_{\sqrt{n}}^{n} \frac{1}{x} dx + O\left(\frac{(\log n)^2}{n^{1/4}}\right) + O\left((\log n)^2 \int_{\sqrt{n}}^{n} \frac{1}{x^{3/2}} dx\right) \\
&= \frac{1}{2} \log n + O\left(\frac{(\log n)^2}{n^{1/4}}\right).
\end{aligned}$$

Upon multiplying by $n(n-1)$ and accounting for the error terms, the lemma is proven.

□

Lemma 3.5.2. *Assuming the Riemann Hypothesis*

$$B_2(n) = \frac{1}{2}n^2 \log n + (\gamma - 1)n^2 + O\left(\frac{(\log n)^2}{n^{1/4}}\right).$$

Proof.

$$\begin{aligned}
B_2(n) &= (n-1)\sum_{j=1}^{\sqrt{n}-1} j\left[\sum_{\frac{n}{j+1}<p\le\frac{n}{j}} \log p\right] \\
&= (n-1)\left[\sum_{j=1}^{\sqrt{n}-1} j\left(\theta\left(\frac{n}{j}\right) - \theta\left(\frac{n}{j+1}\right)\right)\right] \\
&= (n-1)\left[\sum_{j=1}^{\sqrt{n}-1} \theta\left(\frac{n}{j}\right) - (\sqrt{n}-1)\theta(\sqrt{n})\right].
\end{aligned}$$

We see

$$(n-1)(\sqrt{n}-1)\theta(\sqrt{n}) = n^2 + O\left(n^{7/4}(\log n)^2\right).$$

And using asymptotics for harmonic series,

$$\begin{aligned}
(n-1)\sum_{j=1}^{\sqrt{n}-1} \theta\left(\frac{n}{j}\right) &= (n-1)\left[\sum_{j=1}^{\sqrt{n}-1} \frac{n}{j} + O\left((\log n)^2 \sum_{j=1}^{\sqrt{n}-1}\left(\frac{n}{j}\right)^{1/2}\right)\right] \quad \text{by (3.5.1)} \\
&= \frac{1}{2}n(n-1)\log n + n(n-1)\gamma + \frac{\sqrt{n}}{2}(n-1) \\
&\qquad + O(n) + O\left((n-1)(\log n)^2 \sum_{j=1}^{\sqrt{n}-1}\left(\frac{n}{j}\right)^{\frac{1}{2}}\right) \\
&= \frac{1}{2}n^2\log n + n^2\gamma + O\left(n^{3/2}\right) + O\left(n^{3/2}(\log n)^2\left(\int_1^{\sqrt{n}} \frac{1}{x^{1/2}}\,dx + 1\right)\right) \\
&= \frac{1}{2}n^2\log n + n^2\gamma + O\left(n^{7/4}(\log n)^2\right).
\end{aligned}$$

□

Putting the previous two lemmas together, we have the following theorem:

Theorem 3.5.3. *Assuming the Riemann Hypothesis,*

$$B(n) = B_1(n) - B_2(n) = (1-\gamma)n^2 + O\left(n^{7/4}(\log n)^2\right). \tag{3.5.2}$$

Since we know from Theorem 3.1.5 that $\log \overline{G}(n) = \frac{1}{2}n^2 + O\left(n \log n\right)$ and we have by definition that $\log \overline{G}_n = A(n) - B(n)$, we now arrive at an asymptotic formula with a better error term for $A(n)$.

Theorem 3.5.4. *Assuming the Riemann Hypothesis,* $A(n) = \left(\frac{3}{2} - \gamma\right) n^2 + O\left(n^{7/4}(\log n)^2\right).$

3.5.2 Improved Asymptotics for $C(n)$

To find improved asymptotics for $C(n)$, we use two main ingredients. The first is that on assuming the Riemann Hypothesis, we get the following asymptotic formula with small error term for the prime counting function:

$$\pi(n) = \mathrm{li}(n) + O(\sqrt{n}\log n). \tag{3.5.3}$$

The second ingredient is a formula for the logarithmic integral that we'll get by modifying the formula in Theorem 3.4.3. The formula is as follows.

Lemma 3.5.5. *Taking a variable truncation at* $m_0 := \lfloor \frac{1}{2}\log x \rfloor$*, it holds for all* $x \geq 4$ *that*

$$\mathrm{li}(x) = \sum_{k=1}^{m_0(x)} (k-1)! \frac{x}{(\log x)^k} + O(\sqrt{x \log x})$$

where the O-constant is absolute.

Proof. We have that

$$\mathrm{li}(x) = \sum_{k=1}^{m_0}(k-1)!\frac{x}{(\log x)^k} + \sum_{k=1}^{m_0}(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} + O(\sqrt{x}) + O\left(\int_{\sqrt{x}}^{x} \frac{m_0!}{(\log t)^{m_0+1}}\,dt\right)$$

where $m_0 = \lfloor \frac{1}{2}\log x \rfloor$. First we look at $\sum_{k=1}^{m_0}(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k}$. We see by Stirling's approximation (Theorem 1.11.4) that

$$(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} \leq ek^k e^{-k}\sqrt{k}\frac{\sqrt{x}}{(\log\sqrt{x})^k}.$$

Noting that on $\left[1, \lfloor\frac{1}{2}\log x\rfloor\right]$, it holds that $k^k \leq \left(\frac{1}{2}\log x\right)^k$ we have

$$\begin{aligned}
\sum_{k=1}^{m_0}(k-1)!\frac{\sqrt{x}}{(\log\sqrt{x})^k} &\leq \sum_{k=1}^{m_0} e\left(\frac{1}{2}\log x\right)^k e^{-k}\sqrt{k}\frac{\sqrt{x}}{\left(\frac{1}{2}\log x\right)^k} \\
&= e\sqrt{x}\sum_{k=1}^{m_0}\frac{\sqrt{k}}{e^k} \\
&= O(\sqrt{x}).
\end{aligned}$$

Next we use Stirling's formula to bound $m_0!$, allowing us to bound $\int_{\sqrt{x}}^{x}\frac{m_0!}{(\log t)^{m_0+1}}\,dt$. We know that $m_0! \leq e m_0^{m_0}e^{-m_0}\sqrt{m_0}$, giving us

$$\begin{aligned}
\int_{\sqrt{x}}^{x}\frac{m_0!}{(\log t)^{m_0+1}}\,dt &\leq m_0!\int_{\sqrt{x}}^{x}\frac{1}{\left(\frac{1}{2}\log x\right)^{m_0+1}}\,dt \\
&\leq m_0!\frac{x}{\left(\frac{1}{2}\log x\right)^{m_0+1}} \\
&\leq e m_0^{m_0}e^{-m_0}\sqrt{m_0}\frac{x}{\left(\frac{1}{2}\log x\right)^{m_0+1}} \\
&= O\left(e^{-m_0}\sqrt{m_0}x\right) \\
&= O\left(\sqrt{x\log x}\right).
\end{aligned}$$

□

We're now ready to prove an improved asymptotic formula for $C(n)$.

Theorem 3.5.6. *Assuming the Riemann Hypothesis,*

$$C(n) = \left(\frac{1}{2}-\gamma\right)n - \sum_{k=1}^{\lfloor\frac{1}{2}\log n\rfloor} k!\frac{n}{(\log n)^k} + O\left(n^{3/4}(\log n)^2\right).$$

Proof. We have from Theorem 3.1.5 that

$$\frac{1}{2}n^2 + O(n\log n) = \pi(n)n\log n + nC(n) - B(n)$$

so that

$$\begin{aligned} nC(n) &= \frac{1}{2}n^2 - \pi(n)n\log n + B(n) + O(n\log n) \\ &= \frac{1}{2}n^2 - \left(\mathrm{li}(n) + O(\sqrt{n}\log n)\right) n\log n + B(n) + O(n\log n) \quad \text{by (3.5.3)} \\ &= \frac{1}{2}n^2 - \mathrm{li}(n)n\log n + B(n) + O\left(n^{3/2}(\log n)^2\right). \end{aligned}$$

Therefore assuming the Riemann Hypothesis, we have that

$$\begin{aligned} C(n) &= \frac{1}{2}n - \mathrm{li}(n)\log n + \frac{B(n)}{n} + O\left(n^{1/2}(\log n)^2\right) \\ &= \left(\frac{3}{2} - \gamma\right) n - \mathrm{li}(n)\log n + O\left(n^{3/4}(\log n)^2\right) \quad \text{by 3.5.2} \\ &= \left(\frac{1}{2} - \gamma\right) n - \sum_{k=1}^{\lfloor \frac{1}{2}\log n \rfloor} k! \frac{n}{(\log n)^k} + O\left(n^{3/4}(\log n)^2\right) \end{aligned}$$

where in this last line, we've made use of Lemma 3.5.5.

□

3.6 A variant of the problem: Base b Radix Expansions

We ask ourselves what happens to the quantities $A(n)$ $B(n)$ and $C(n)$ when instead of restricting to primes p, we sum over all the integers smaller than n.

Definition 3.6.1. *In this section, we let*

$$A'(n) = \sum_{2 \le b \le n} \frac{2\log b}{b-1} S_b(n) \qquad B'(n) = \sum_{2 \le b \le n} \frac{n-1}{b-1} d_b(n) \log b$$

and

$$C'(n) = \sum_{2 \le b \le n} \frac{2\log b}{b-1} f_b(\log_b(n))$$

where we're taking the sum over the same expressions as before, but this time we're summing over all positive integers greater than 1 *and less than or equal to* n*, instead of just the primes.*

Remark 3.6.2. *We exclude the integer $b = 1$ from the sums for $A'(n), B'(n)$ and $C'(n)$ for two reasons:*

- *The denominator $b - 1$ means that the sums are not well-defined at $b = 1$.*
- *$d_1(n)$ is not well defined, since 1 raised to any power is still 1 and there are no positive integers less than 1, so there's no way to choose the a_i in the radix expansion.*

Remark 3.6.3. *The result in Theorem 3.1.5, equation* (3.1.2) *stated over primes does not hold when we sum over all numbers, including composites. I.e.*

$$\log \overline{G}_n \neq A'(n) - B'(n).$$

This is because the proof of Theorem 3.1.5, where we just summed over primes relies heavily on the fact that for any integer n and any prime p, the p-adic valuation of $n!$ is given by the formula $\sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor$. However for a non-prime base b, it's very much not the case that the b-adic valuation of $n!$ is $\sum_{i=1}^{\infty} \left\lfloor \frac{n}{b^i} \right\rfloor$.

The main results in this section are the following asymptotic formulas for $A'(n), B'(n)$ and $C'(n)$. In this case, it turns out that once we obtain asymptotics for $B'(n)$, the asymptotics for $A'(n)$ and $C'(n)$ are straightforward to write down. We recall the theorems given in the introduction.

Theorem 3.6.4. *For all $n \geq 4$*

$$A'(n) = \left(\frac{3}{2} - \gamma\right) n^2 \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right) n^2 + O(n^{3/2} \log n).$$

Theorem 3.6.5. *For all $n \geq 4$*

$$B'(n) = (1 - \gamma)\, n^2 \log n + (\gamma + \alpha - 1)\, n^2 + O(n^{3/2} \log n).$$

Theorem 3.6.6. *For all $n \geq 4$*

$$C'(n) = \left(\frac{1}{2} - \gamma\right) n \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right) n + O(\sqrt{n} \log n).$$

Remark 3.6.7. *We note that the coefficients of the main terms in these expressions are the same as in Theorems 3.1.7, 3.1.8 and 3.1.9 respectively. However, now that we're summing over all integers smaller than n, the orders of magnitude of the main terms in these new expressions are a multiple of $\log n$ larger than the main terms for $A(n)$, $B(n)$ and $C(n)$.*

3.6.1 Asymptotics for $B'(n)$

To prove Theorem 3.6.5, we break the sum into two pieces, one piece summing up to $\sqrt{n}$ (which has negligible contribution) and the other summing between $\sqrt{n}$ and n which we analyse. I.e. we write $B'(n) = B'_1(n) + B'_R(n)$ where

$$B'_1(n)_= \sum_{\sqrt{n}<b\leq n} \frac{n-1}{b-1} d_b(n)\log b \qquad B'_R(n) = \sum_{b\leq\sqrt{n}} \frac{n-1}{b-1} d_b(n)\log b.$$

This decomposition is analogous to the one in the proof of Theorem 3.1.8.

We'll first show that $B'_R(n)$ is negligible compared to the main term.

Lemma 3.6.8. $B'_R(n) = O\left(n^{3/2}\log n\right)$.

Proof. We use the definition of digit sum expansions to note that for every b,

$$d_b(n) \leq (b-1)\left(\frac{\log n}{\log b} + 1\right).$$

This allows us to say

$$\begin{aligned}
B'_R(n) &= \sum_{b\leq\sqrt{n}} \frac{n-1}{b-1} d_b(n)\log b \\
&\leq \sum_{b\leq\sqrt{n}} (n-1)\log n + \sum_{b\leq\sqrt{n}} (n-1)\log b \\
&\leq n^{3/2}\log n + (n-1)\sum_{b\leq\sqrt{n}} \log b \\
&= O\left(n^{3/2}\log n + n\int_1^{\sqrt{n}} \log x\,dx\right) \\
&= O\left(n^{3/2}\log n\right).
\end{aligned}$$

□

Next we see that just as in the prime case, for $1 < j \leq \sqrt{n}$ and $b \in \left(\frac{n}{j+1}, \frac{n}{j}\right]$, we have that $d_b(n) = n - j(b-1)$. This allows us to break down $B_1'(n)$ into $B_{11}'(n) - B_{12}'(n)$ where

$$B_{11}'(n) = n(n-1) \sum_{\sqrt{n} < b \leq n} \frac{\log b}{b-1}, \qquad B_{12}'(n) = (n-1) \sum_{j=1}^{\lfloor \sqrt{n} \rfloor} j \left(\sum_{\frac{n}{j+1} < b \leq \frac{n}{j}} \log b \right)$$

and we analyse each part separately.

Lemma 3.6.9. $B_{11}'(n) = \frac{3}{8} n^2 \log^2 n + O\left(n^{3/2} \log n\right).$

Proof. We first see that $\sum_{\sqrt{n} < b \leq n} \frac{\log b}{b-1} = \sum_{\sqrt{n} < b \leq n} \frac{\log b}{b} + O\left(\frac{\log n}{\sqrt{n}}\right)$. This is because $\sum_{\sqrt{n} < b \leq n} \frac{\log b}{b-1} = \sum_{\sqrt{n} < b \leq n} \frac{\log b}{b} + \sum_{\sqrt{n} < b \leq n} \frac{\log b}{b(b-1)}$ and we have that

$$\begin{aligned} \sum_{\sqrt{n} < b \leq n} \frac{\log b}{b(b-1)} &= O\left(\log n \sum_{b \geq \sqrt{n}} \frac{1}{b^2}\right) \\ &= O\left(\log n \int_{\sqrt{n}}^{\infty} \frac{1}{x^2}\, dx\right) \\ &= O\left(\frac{\log n}{\sqrt{n}}\right). \end{aligned}$$

Next we note that the error in approximating this sum with the integral is small, namely that

$$\begin{aligned} \left| \sum_{\sqrt{n} < b \leq n} \frac{\log b}{b} - \int_{\sqrt{n}}^{n} \frac{\log x}{x}\, dx \right| &\leq \frac{\log \sqrt{n}}{\sqrt{n}} - \frac{\log n}{n} \\ &= O\left(\frac{\log n}{\sqrt{n}}\right) \end{aligned}$$

where we've used the error term from Riemann sum approximations:

$$|f(b) - f(a)| \Delta x. \tag{3.6.1}$$

This is valid because for $n \geq 9$, we have that $\frac{\log b}{b}$ is a decreasing function on the interval $[\sqrt{n}, n]$. Therefore we have

$$\begin{aligned}\sum_{\sqrt{n}<b\leq n} \frac{\log b}{b} &= \int_{\sqrt{n}}^{n} \frac{\log x}{x}\, dx + O\left(\frac{\log n}{\sqrt{n}}\right)\\ &= \frac{1}{2}\left(\log x\right)^2 \Big|_{\sqrt{n}}^{n} O\left(\frac{\log n}{\sqrt{n}}\right)\\ &= \frac{3}{8}\log^2 n + O\left(\frac{\log n}{\sqrt{n}}\right)\end{aligned}$$

and so the lemma is proved. □

Next we analyse

$$B'_{12}(n) = (n-1)\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\sum_{\frac{n}{j+1}<b\leq\frac{n}{j}} \log b\right).$$

Again using the error term in the Riemann sum approximations for monotone functions: (3.6.1) we get

$$\begin{aligned}B'_{12}(n) &= (n-1)\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\sum_{\frac{n}{j+1}<b\leq\frac{n}{j}} \log b\right)\\ &= (n-1)\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\int_{\frac{n}{j+1}}^{\frac{n}{j}} \log x\, dx + O\left(\log\left(\frac{n}{j}\right) - \log\left(\frac{n}{j+1}\right)\right)\right)\\ &= (n-1)\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\int_{\frac{n}{j+1}}^{\frac{n}{j}} \log x\, dx + O\left(\frac{1}{j}\right)\right)\\ &= (n-1)\left(\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\int_{\frac{n}{j+1}}^{\frac{n}{j}} \log x\, dx\right)\right) + O(n^{3/2})\\ &= (n-1)\left[\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\frac{n}{j}\log\left(\frac{n}{j}\right) - \frac{n}{j+1}\log\left(\frac{n}{j+1}\right)\right) + \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\frac{n}{j+1} - \frac{n}{j}\right)\right]\\ &\qquad + O\left(n^{3/2}\right).\end{aligned}$$

We call

$$B'_{12a}(n) = \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\frac{n}{j} - \frac{n}{j+1}\right)$$

and

$$B'_{12b}(n) = \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\frac{n}{j}\log\left(\frac{n}{j}\right) - \frac{n}{j+1}\log\left(\frac{n}{j+1}\right)\right)$$

so that $B'_{12}(n) = (n-1)(B_{12b}(n) - B'_{12a}(n)) + O(n^{3/2})$.

Lemma 3.6.10. $B'_{12a}(n) = \frac{1}{2}n\log n + (\gamma - 1)n + O(\sqrt{n})$.

Proof. We can simplify $\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\frac{n}{j} - \frac{n}{j+1}\right) = n\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\frac{1}{j+1} = n\sum_{j=2}^{\lfloor\sqrt{n}\rfloor+1}\frac{1}{j}$. This is the truncated harmonic series, for which the asymptotics, obtained by Euler-Maclaurin summation formula are given in Theorem 1.11.2. We get that

$$n\sum_{j=2}^{\lfloor\sqrt{n}\rfloor+1}\frac{1}{j} = n\left(\log(\lfloor\sqrt{n}\rfloor + 1) + (\gamma - 1) + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

All that's left to note is that

$$\left|\log(\lfloor\sqrt{n}\rfloor + 1) - \log\sqrt{n}\right| \leq \frac{1}{\sqrt{n}}$$

giving us the stated asymptotic formula. □

Next we look at $B'_{12b}(n) = \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j\left(\frac{n}{j}\log\left(\frac{n}{j}\right) - \frac{n}{j+1}\log\left(\frac{n}{j+1}\right)\right)$. In order to analyse $B'_{12b}(n)$, we will need the following estimate.

Lemma 3.6.11. *We have that*

$$\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\frac{\log(j+1)}{j+1} = \frac{1}{8}\log^2 n + \alpha + O\left(\frac{\log n}{\sqrt{n}}\right).$$

Here α is the constant defined in (3.1.3) by

$$\alpha := \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{\log k}{k} - \int_1^n \frac{\log t}{t}\,dt\right).$$

Proof. We reindex the sum as $\sum_{j=2}^{\lfloor\sqrt{n}\rfloor+1}\frac{\log(j)}{j}$. The Euler-Maclaurin summation formula

(Theorem 1.11.1) gives us:

$$\begin{aligned}\sum_{j=2}^{\lfloor\sqrt{n}\rfloor+1}\frac{\log(j)}{j} &= \int_1^{\lfloor\sqrt{n}\rfloor+1}\frac{\log t}{t}\,dt+\frac{1}{2}\left(\frac{\log(\lfloor\sqrt{n}\rfloor+1)}{\lfloor\sqrt{n}\rfloor+1}\right)-\frac{1}{12}\left(\frac{1-\log(\lfloor\sqrt{n}\rfloor+1)}{(\lfloor\sqrt{n}\rfloor+1)^2}-1\right)\\ &\quad+\frac{1}{120}\left(\frac{-6\log(\lfloor\sqrt{n}\rfloor+1)+11}{(\lfloor\sqrt{n}\rfloor+1)^4}-11\right)-\int_1^{\lfloor\sqrt{n}\rfloor+1}B_4(t)\left(\frac{-50+24\log t}{t^5}\right)dt\\ &= \frac{1}{2}\log^2 t\bigg|_1^{\lfloor\sqrt{n}\rfloor+1}+\frac{1}{12}-\frac{11}{120}-\int_1^{\lfloor\sqrt{n}\rfloor+1}B_4(t)\left(\frac{-50+24\log t}{t^5}\right)dt\\ &\quad+O\left(\frac{\log n}{\sqrt{n}}\right)\\ &= \frac{1}{2}(\log^2(\lfloor\sqrt{n}\rfloor+1))+\frac{1}{12}-\frac{11}{120}-\int_1^{\lfloor\sqrt{n}\rfloor+1}B_4(t)\left(\frac{-50+24\log t}{t^5}\right)dt\\ &\quad+O\left(\frac{\log n}{\sqrt{n}}\right).\end{aligned}$$

We note by the difference of two squares that

$$\begin{aligned}\log^2(\lfloor\sqrt{n}\rfloor+1)-\log^2(\sqrt{n}) &= (\log(\lfloor\sqrt{n}\rfloor+1)-\log\sqrt{n})(\log(\lfloor\sqrt{n}\rfloor+1)+\log\sqrt{n})\\ &= O\left(\frac{1}{\sqrt{n}}\log n\right).\end{aligned}$$

Therefore

$$\begin{aligned}\sum_{j=2}^{\lfloor\sqrt{n}\rfloor+1}\frac{\log(j)}{j} &= \frac{1}{2}\log^2(\sqrt{n})+\frac{1}{12}-\frac{11}{120}-\int_1^{\lfloor\sqrt{n}\rfloor+1}B_4(t)\left(\frac{-50+24\log t}{t^5}\right)dt\\ &\quad+O\left(\frac{\log n}{\sqrt{n}}\right).\end{aligned}$$

Noting that when $j=1$, $\frac{\log j}{j}=0$ and letting n tend to infinity, we obtain

$$\begin{aligned}\alpha &= \lim_{n\to\infty}\left(\sum_{j=2}^{\lfloor\sqrt{n}\rfloor+1}\frac{\log(j)}{j}-\frac{1}{2}\log^2(\sqrt{n})\right)\\ &= \frac{1}{12}-\frac{11}{120}-\int_1^{\infty}B_4(t)\left(\frac{-50+24\log t}{t^5}\right)dt.\end{aligned}$$

where the first of these two lines comes from the limit definition of α in (3.1.3) as the first Stieltjes constant. We just need to note finally that

$$\int_{\sqrt{n}}^{\infty}B_4(t)\left(\frac{-50+24\log t}{t^5}\right)dt=O\left(\frac{1}{n}\right)$$

to arrive at the desired result. □

We're now ready to give asymptotics for $B'_{12b}(n)$:

Lemma 3.6.12. $B'_{12b}(n) = \frac{3}{8} n \log^2 n + \left(\gamma - \frac{3}{2}\right) n \log n - \alpha n + O(\sqrt{n} \log n)$

Proof. We'll first rewrite the expression for $B'_{12b}(n)$ so that it's easier to analyse.

$$\begin{aligned} B'_{12b}(n) &= \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} j \left(\frac{n}{j} \log\left(\frac{n}{j}\right) - \frac{n}{j+1} \log\left(\frac{n}{j+1}\right)\right) \\ &= n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \left(\log n - \log j - \frac{j}{j+1} \log n + \frac{j}{j+1} \log(j+1)\right) \\ &= n \log n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \frac{1}{j+1} + n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \left(\frac{j}{j+1} \log(j+1) - \log j \right). \end{aligned}$$

From the Theorem 1.11.2 we know the asymptotics for the first of these sums:

$$n \log n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \frac{1}{j+1} = n \log n \left(\frac{1}{2} \log n + (\gamma - 1) + O\left(\frac{1}{\sqrt{n}}\right)\right).$$

For the second sum we have

$$\begin{aligned} n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \left(\frac{j}{j+1} \log(j+1) - \log j \right) &= n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} (\log(j+1) - \log j) - n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \frac{1}{j+1} \log(j+1) \\ &= n \log(\lfloor\sqrt{n}\rfloor + 1) - n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \frac{1}{j+1} \log(j+1) \\ &= \frac{1}{2} n \log n - n \sum_{j=1}^{\lfloor\sqrt{n}\rfloor} \frac{1}{j+1} \log(j+1) + O(\sqrt{n}). \end{aligned}$$

Putting all this together, we obtain:

$$\begin{aligned}
B'_{12b}(n) &= n\log n\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\frac{1}{j+1} + n\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\left(\frac{j}{j+1}\log(j+1) - \log j\right)\\
&= n\log n\left(\frac{1}{2}\log n + (\gamma-1) + O\left(\frac{1}{\sqrt{n}}\right)\right) + \frac{1}{2}n\log n\\
&\qquad - n\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\frac{1}{j+1}\log(j+1) + O(\sqrt{n})\\
&= \frac{1}{2}n\log^2 n + \left(\gamma - \frac{1}{2}\right)n\log n - n\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\frac{1}{j+1}\log(j+1) + O(\sqrt{n}\log n)\\
&= \frac{1}{2}n\log^2 n + \left(\gamma - \frac{1}{2}\right)n\log n - \frac{1}{8}n\log^2 n - \alpha n + O(\sqrt{n}\log n)\\
&= \frac{3}{8}n\log^2 n + \left(\gamma - \frac{1}{2}\right)n\log n - \alpha n + O(\sqrt{n}\log n).
\end{aligned}$$

□

Proof of Theorem 3.6.5

Since

$$\begin{aligned}
B'(n) &= B'_{11}(n) - B'_{12}(n) + O(n^{3/2}\log n)\\
&= B'_{11}(n) - (n-1)(B'_{12b}(n) - B'_{12a}(n)) + O(n^{3/2}\log n)
\end{aligned}$$

and we have from Lemma 3.6.9 that

$$B'_{11}(n) = \frac{3}{8}n^2\log^2 n + O\left(n^{3/2}\log n\right).$$

we can rewrite this as

$$B'(n) = \frac{3}{8}n^2\log n + O(n^{3/2}\log n) - (n-1)(B'_{12b}(n) - B'_{12a}(n)).$$

Finally, we can use the asymptotics for $B'_{12a}(n)$ and $B'_{12b}(n)$ found in Lemmas 3.6.10 and 3.6.12 to conclude the given asymptotics for $B'(n)$. □

3.6.2 Asymptotics for $A'(n)$ and $C'(n)$

We now find the asymptotics for $A'(n)$. Given the size of $B'(n)$, we conjecture that $A'(n)$ will have terms of order $n^2 \log n$ and n^2 as well as a power-saving error term. Directly from its definition, we have

$$\begin{aligned} A'(n) &= \sum_{2 \le b \le n} \frac{2 \log b}{b-1} S_b(n) \\ &= \sum_{2 \le b \le n} \frac{2 \log b}{b} S_b(n) + \sum_{2 \le b \le n} \frac{2 \log b}{b(b-1)} S_b(n). \end{aligned}$$

To find asymptotics for $A'(n)$, we'll make use of Drazin and Griffith's theorem, which gives an upper bound for the running digit sum function:

Theorem 3.6.13. *[12] (Drazin and Griffith) Let $b \ge 2$ be an integer. Then for all $n \ge 1$,*

$$S_b(n) \le \frac{b-1}{2} n \log_b n$$

and equality holds if and only if $n = b^k$ for $k \ge 0$.

Lemma 3.6.14. *We have that*

$$A'(n) = \sum_{2 \le b \le n} \frac{2 \log b}{b} \sum_{j=0}^{n} d_b(j) + O(n \log^2 n).$$

Proof. From Drazin and Griffith's theorem (Theorem 3.6.13), we know that

$$S_b(n) \le \frac{b-1}{2} n \frac{\log n}{\log b}.$$

Applying it, we have

$$\begin{aligned} A'(n) &= \sum_{2 \le b \le n} \frac{2 \log b}{b} S_b(n) + O\left(\sum_{2 \le b \le n} \frac{2 \log b}{b(b-1)} \frac{b-1}{2} n \frac{\log n}{\log b} \right) \\ &= \sum_{2 \le b \le n} \frac{2 \log b}{b} S_b(n) + O\left(n \log n \sum_{2 \le b \le n} \frac{1}{b} \right) \\ &= \sum_{2 \le b \le n} \frac{2 \log b}{b} S_b(n) + O(n \log^2 n) \\ &= \sum_{2 \le b \le n} \frac{2 \log b}{b} \sum_{j=0}^{n} d_b(j) + O(n \log^2 n). \end{aligned}$$

□

Noting that $d_b(0) = 0$ and $d_b(1) = 1$ for any $b \geq 2$ we have that

$$A'(n) = \sum_{2 \leq b \leq n} \frac{2 \log b}{b} \sum_{j=2}^{n} d_b(j) + O(n \log^2 n)$$

leaving us to estimate the sum $\sum_{2 \leq b \leq n} \frac{2 \log b}{b} \sum_{j=2}^{n} d_b(j)$. We'll do this by switching the order of summation and using what we know about the quantities $B'(j)$ in the following:

$$\begin{aligned}
\sum_{2 \leq b \leq n} \frac{2 \log b}{b} \sum_{j=2}^{n} d_b(j) &= \sum_{j=2}^{n} \sum_{2 \leq b \leq n} \frac{2 \log b}{b} d_b(j) \\
&= \sum_{j=2}^{n} \left(\sum_{2 \leq b \leq j} \frac{2 \log b}{b} d_b(j) + \sum_{j+1 \leq b \leq n} \frac{2 \log b}{b} d_b(j) \right) \\
&= \sum_{j=2}^{n} \frac{2}{j-1} B'(j) - \sum_{j=2}^{n} \sum_{2 \leq b \leq j} \frac{2 \log b}{b(b-1)} d_b(j) \\
&\qquad + \sum_{j=2}^{n} \sum_{j+1 \leq b \leq n} \frac{2 \log b}{b} j
\end{aligned}$$

where we've used the identity $\frac{2 \log b}{b} = \frac{2 \log b}{b-1} - \frac{2 \log b}{b(b-1)}$ and gotten the third sum from the fact that for $b > j$, the base b digit expansion of j has just one digit: j itself.

We'll write

$$\begin{aligned}
A'_1(n) &= \sum_{j=2}^{n} \frac{2}{j-1} B'(j), \\
A'_R(n) &= \sum_{j=2}^{n} \sum_{2 \leq b \leq j} \frac{2 \log b}{b(b-1)} d_b(j), \\
A'_2(n) &= \sum_{j=2}^{n} \sum_{j+1 \leq b \leq n} \frac{2 \log b}{b} j
\end{aligned}$$

where we'll show that the contribution from $A'_R(n)$ is negligible and conclude that

$$A'(n) = A'_1(n) + A'_2(n) + O(n^{3/2} \log n).$$

First we consider $A'_1(n) = \sum_{j=2}^{n} \frac{2}{j-1} B'(j)$.

Lemma 3.6.15. $A'_1(n) = (1 - \gamma) n^2 \log n + \left(\frac{3}{2}\gamma + \alpha - \frac{3}{2} \right) n^2 + O(n^{3/2} \log n)$.

Proof. By Theorem 3.6.5 we can write $\sum_{j=2}^{n} \frac{2}{j-1} B'(j)$ as

$$\sum_{j=2}^{n} \frac{2}{j-1} B'(j) = 2 \sum_{j=2}^{n} \frac{1}{j-1} \left((1-\gamma) j^2 \log j + (\gamma + \alpha - 1) j^2 + O(j^{3/2} \log j) \right).$$

We'll approximate $\frac{j^2}{j-1}$ by j where the error in approximation is of magnitude $O(1)$, giving us

$$\begin{aligned} A_1'(n) &= 2 \sum_{j=2}^{n} ((1-\gamma) j \log j + (\gamma + \alpha - 1) j) + O\left(\sum_{j=2}^{n} \log j \right) + O(n^{3/2} \log n) \\ &= 2 \sum_{j=2}^{n} (1-\gamma) j \log j + (\gamma + \alpha - 1) n^2 + O(n^{3/2} \log n). \end{aligned}$$

Next we'll note that for $x \geq 2$, the function $x \log x$ is strictly increasing so that we may approximate remaining sum via its corresponding integral, with an error no larger than $n \log n$. Therefore

$$\begin{aligned} A_1'(n) &= 2(1-\gamma) \int_2^n x \log x \, dx + O(n \log n) + (\gamma + \alpha - 1) n^2 + O(n^{3/2} \log n) \\ &= 2(1-\gamma) \left(\frac{1}{2} x^2 \log x \Big|_2^n - \int_1^n \frac{1}{2} x \, dx \right) + (\gamma + \alpha - 1) n^2 + O(n^{3/2} \log n) \\ &= 2(1-\gamma) \left(\frac{1}{2} n^2 \log n - \frac{1}{4} n^2 \right) + (\gamma + \alpha - 1) n^2 + O(n^{3/2} \log n) \\ &= (1-\gamma) n^2 \log n + \left(\frac{3}{2} \gamma + \alpha - \frac{3}{2} \right) n^2 + O(n^{3/2} \log n). \end{aligned}$$

□

Next we need to consider $A_2'(n) = \sum_{j=2}^{n} \sum_{j+1 \leq b \leq n} \frac{2 \log b}{b} j$.

Lemma 3.6.16. $A_2'(n) = \frac{1}{2} n^2 \log n - \frac{1}{4} n^2 + O(n \log^2 n)$.

Proof. As in the proof of Lemma 3.6.15, we'll approximate sums via integrals. We first use the fact that $\frac{2 \log x}{x}$ is decreasing for $x \geq e$ giving us that

$$\begin{aligned} \sum_{j+1 \leq b \leq n} \frac{2 \log b}{b} &= \int_{j+1}^{n} \frac{2 \log x}{x} dx + O\left(\frac{\log j}{j} \right) \\ &= \log^2 n - \log^2 (j+1) + O\left(\frac{\log j}{j} \right). \end{aligned}$$

Considering the difference of two squares

$$\begin{aligned}\log^2(j+1)-\log^2(j) &= (\log(j+1)+\log j)(\log(j+1)-\log j)\\ &= O\left(\log j \log\left(\frac{j+1}{j}\right)\right)\\ &= O\left(\frac{\log j}{j}\right).\end{aligned}$$

This now allows us to write

$$\begin{aligned}\sum_{j=2}^{n}\sum_{j+1\le b\le n}\frac{2\log b}{b}j &= \sum_{j=2}^{n} j\left(\log^2 n-\log^2 j+O\left(\frac{\log j}{j}\right)\right)\\ &= \frac{n(n+1)\log^2 n}{2}-\sum_{j=2}^{n} j\log^2 j+O(n\log n)\end{aligned}$$

where we'll again do an integral approximation for the remaining sum, getting us

$$\begin{aligned}A_2'(n) &= \frac{1}{2}n^2\log^2 n-\int_2^n x\log^2 x\,dx+O(n\log^2 n)\\ &= \frac{1}{2}n^2\log^2 n-\frac{1}{2}x^2\log^2 x\Big|_2^n+\int_2^n x\log x\,dx+O(n\log^2 n)\\ &= \frac{1}{2}x^2\log x\Big|_2^n-\int_2^n\frac{1}{2}x\,dx+O(n\log^2 n)\\ &= \frac{1}{2}n^2\log n-\frac{1}{4}n^2+O(n\log^2 n).\end{aligned}$$

□

Finally we need to show that the contribution from $A_R'(n)=\sum_{j=2}^{n}\sum_{2\le b\le j}\frac{2\log b}{b(b-1)}d_b(j)$ is negligible.

Lemma 3.6.17. $A_R'(n)=O(n\log^2 n)$.

Proof. We can bound $\sum_{j=2}^{n}\sum_{2\le b\le j}\frac{2\log b}{b(b-1)}d_b(j)$ above by

$$\sum_{2\le b\le n}\sum_{j=2}^{n}\frac{2\log b}{b(b-1)}d_b(j).$$

However, this last expression is $O\left(\sum_{2\leq b\leq n}\frac{2\log b}{b(b-1)}S_b(n)\right)$, which using the bound from Drazin and Griffith's theorem (Theorem 3.6.13) is $O\left(n\log n\sum_{2\leq b\leq n}\frac{1}{b}\right)$ and this is just $O\left(n\log^2 n\right)$.

□

Putting everything together from Lemmas 3.6.15, 3.6.16 and 3.6.17 we can prove Theorem 3.6.4. We recall that

$$A'(n) = A_1'(n) + A_2'(n) + A_R'(n) + O(n^{7/4})$$

so utilising the asymptotics we obtained in the lemmas we have

$$\begin{aligned}
A'(n) &= \left((1-\gamma)n^2\log n + \left(\frac{3}{2}\gamma + \alpha - \frac{3}{2}\right)n^2 + O(n^{3/2}\log n)\right) \\
&\quad + \left(\frac{1}{2}n^2\log n - \frac{1}{4}n^2 + O(n\log^2 n)\right) \\
&= \left(\frac{3}{2} - \gamma\right)n^2\log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right)n^2 + O(n^{3/2}\log n).
\end{aligned}$$

Delange's Theorem (Theorem 3.1.2) (which holds in base b as well as base p) tells us

$$\begin{aligned}
A'(n) &= \sum_{2\leq b\leq n} n\log n + nC'(n) \\
&= n(n-1)\log n + nC'(n)
\end{aligned}$$

giving us the result of Theorem 3.6.6:

$$C'(n) = \left(\frac{1}{2} - \gamma\right)n\log n + \left(\frac{3}{2}\gamma + \alpha - \frac{7}{4}\right)n + O(\sqrt{n}\log n).$$

CHAPTER IV

Stories from my Classroom: Reflections and Narratives from Different Types of Mathematicians

4.1 Introduction

A thesis reflecting my graduate work would not be complete without a chapter about my teaching experiences. I've been extremely lucky to have taught in multiple different settings over the past six years. My wonderful experiences in the classroom have been challenging, fulfilling and extremely informative. Being a teacher and mentor, especially to those students who seem at first a little different from the stereotypical mathematician, has grown to be an integral part of my professional life and personal identity.

Starting in fall 2014, during all but two semesters, I've been in charge of my own class in Precalculus, Calculus I or Calculus II at the University of Michigan. In addition, I've worked for Canada/USA Mathcamp and Bridge to Enter Advanced Mathematics, as well as founded my own math club for middle school students.

My purpose in writing this chapter is to record some of my students' and my own experiences, so that the lessons I've learned from them may stay with me through time. I'll explore some of the challenges we've overcome together as well as the challenges that remain for them, as non-conventional mathematicians or non-conventional learners. This chapter doesn't claim to be formal 'education research' and isn't written to identify any broad trends in math education, nor to give instruction to other educators. It's simply a collection of my reflections, that may also have a chance of being useful for readers to reflect on.

In my stories, instead of assigning fake names to students, I've used random letters to denote their names. This was an intentional choice I made stemming from the fact that names often carry connotations, which can colour the way a story is perceived before it's even told. Just as I think of my students as individuals, it is my hope that their stories can serve as a spotlight for meeting each student where they're at, rather than making assumptions based on any groups they appear to belong to.

4.2 Reflection through Conversation

During the summer of 2019, I interviewed three young mathematicians about their mathematical life and journey [1]. All three are people of colour, from Bronx, NY and are currently undergraduate college students. Although at first glance, we seem different, I wanted to identify some of the things we have in common, as non-conventional mathematicians.

4.2.1 Barriers to Success

I know first hand what it's like to feel doubt and discouragement about my place in math. Throughout my childhood, I always attended public schools, where I was never by any means the 'math kid' in the class. It was especially difficult during my middle school years, being a girl of colour at an all-white school, where with my cultural identity I really struggled to find my place. Once I decided that I liked math and wanted to pursue a math degree, I often felt like I struggled to fit in or be taken seriously since I'd begun so late and other students knew so much more than I did. One of the hardest and also most rewarding things I've ever done is find the path that's right for my mathematical journey. Throughout

[1] The interviews were conducted under the eResearch ID HUM00167368 and given a determination of 'Not Regulated' status by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board.

this process, I've learned to find peace with learning at the pace that's suitable for me, instead of feeling intimidated by those who had a head start, seek guidance from excellent mentors, ask questions unashamedly and treasure each opportunity that's been presented to me.

As non-conventional mathematicians who haven't always necessarily fitted in inside math spaces, my interviewees spoke about external barriers that made it difficult for them to get help or to work with other people. These barriers mainly centred around not receiving encouragement from authority figures and also not having the means to take advantage of resources. A few examples of what they said are given below:

- 'I couldn't go to office hours during high school since I travelled more than an hour for a specialised school. I couldn't stay after school since I had to leave immediately to take care of my younger siblings. I felt discouraged: they didn't have time to help me and didn't accommodate for my needs. My only incentive was to get good grades to get to a good college.'
- 'My teachers said to us that if you don't feel confident in your math ability, then you shouldn't be in an advanced math class. The message ran close to home for me: it's been passed around in my community since elementary school that math isn't for you. A lot of black and latino kids I knew at school could have handled high level math, but didn't sign up for those classes because the messaging told them they couldn't handle it. Kids in upper math classes were from richer backgrounds. They could get a tutor who could help them. Because I felt like I didn't have time to develop understanding, it discouraged me from asking for help. I've always had to be independent, so was never comfortable asking for help.'
- 'Most college STEM classes are filled with white males. It's difficult for me. Sometimes I know what I'm talking about but people don't listen. Once in class I knew the

answer and had an explanation, but explaining out loud is sometimes difficult for me. I was with a group of boys who wouldn't listen to me- they assumed I was wrong. I was frustrated and then started doubting myself. When they finally got it, they said 'good job' to someone else. They didn't accept me for what I was saying. I thought it was because I'm a woman, or because I'm a person of colour they talked over me. It made me feel like a minority. '

- 'My college counsellors who were supposed to help me were not encouraging. They'd encourage you to transfer school if you struggled with a class but asked white kids how they could help them. The first thing they did when I was struggling with a personal issue was ask me to transfer schools. '

I remember struggling earlier in my mathematical journey as a result of feeling out of my depth. Sometimes, just like my interviewees, I'd feel discouraged by not having the background or vocabulary to express my mathematical ideas, or by how other people expressed their perceptions of me.

4.2.2 Asking for Help

My interviewees and I talked about how external factors would feed into our perceptions of ourselves and of the math space around us and how such feelings of isolation would make external factors hit harder and so make it difficult to ask for help even when we needed it:

- 'I was the only black girl in advanced math class again at my school. Sometimes you just want someone who gets where you're coming from. None of my friend group did that level of math, so I felt isolated. As a black girl, even if you know you're right, you still bring yourself down and don't feel ok raising your hand in class.'
- 'There were two other boys from my school at a summer math programme I went to.

Everyone knew them to be the best. I was the only girl from my school. I felt like I had to live up to them. They would get things quicker than me. I'd try my best, but wasn't where they were at. The teachers didn't know I was struggling. I wanted to keep it to myself. I felt that once they knew they'd start pitying me and I didn't want that. Growing up in the Bronx, I learned to do things by myself so I wanted to do math by myself. Even to this day, I want to do things for myself. '

- 'I took a stats class in senior year of high school. I was very lost. We were using a computer programme and I couldn't understand how to do it. I didn't feel comfortable asking questions because I'd fallen off in school. I didn't feel motivated in school-I didn't feel it would matter in the long run. I had a problem asking questions because I didn't feel like I deserved attention and I felt like I should just be able to do it myself. I liked the teacher, but I did find him intimidating. I'm intimidated by authority in general.'
- 'I knew working by myself that I couldn't feel like I didn't fit in or make other people mad. In math I don't necessarily have the right vocab, so I don't feel comfortable talking.'

The underlying theme that my interviewees really helped vocalise for me through these reflections, is that as a teacher, I need to be constantly aware of how different learners will ask for help and support. By the time I have a student in my class, their learning journey will already have taken a unique, complex series of twists and turns. These will have shaped them into the incredible individual that I meet on the first day of class. My job as a teacher is to work with them, starting from where they are and not from where one may assume they 'should be'.

In the next section, I talk about many explicit examples of what working with different types of students has looked like for me. Here I'll give just a couple of brief examples of very talented young mathematicians who have approached asking for and accepting help in unique ways.

1. I was really excited to work with C because I'd seen some of her past work and knew that she'd achieved some outstanding test scores. However she came into my class very quiet and reserved, never drawing attention to herself. It was a class where almost every student was taking full advantage of office hours, but C had never shown up. One day I went to find her. I asked how class was going and if she had any questions. She said no, so I asked if she was comfortable with us just being in the same space as we got work done: she'd work on her homework and I'd work on my grading. C said she was, so I took out my grading and worked away at it. A few times, C looked over at me and after about half an hour, she spoke up: 'Actually, maybe I do have a question...' We had a happy math conversation that day and from then on, she was a regular attendee of office hours, where she asked many great questions and really showcased her amazing mathematical ability.

2. On the surface, it seemed like M didn't care. She was defiant about following instructions and would often interrupt to ask for the answer. When I upheld my boundary of never giving out solutions (since it doesn't help students with actually understanding the material), but made it clear I would help her problem-solve, she told me that she had no time for this. I pushed a little for when she would have time and she suggested a time very early one morning. Once we met, I learned that M was a student from a low-income background who spent most of her time outside classes at work. She struggled to juggle her work-study lifestyle and had availability at odd hours that she wasn't used to being accommodated. To work with M's needs, I held many study sessions at odd hours during the course of the semester, that actually ended up being very well-attended by other students too. M really opened up during these sessions,

becoming a friendly, communicative member of the class and a very diligent student that went above and beyond in exploring the subject matter independently.

An analogy that comes to mind when thinking about students who struggle to ask for help is that of one of my favourite animals: the rabbit. Something I learned from spending time with rabbits is that if I chase them with carrots, they'll always run and hide, but that if I appear as quiet and non-threatening as possible, their curiosity will always take over and they'll come closer to explore. Things that make students curious and excited about exploring vary so much on an individual basis. However I've found that this curiosity comes out best when I give students space to be true to themselves. When I put all my focus into listening to a student, it doesn't just mean hearing the words they're saying. To me, it means constantly recalibrating my energy levels with what they're showing me. It means honouring those hardships they encountered when doing what they're 'supposed to do' didn't work out for them and the subsequent rebellion this may have sparked. It means acknowledging that to many students, asking for help in a math class feels like a frightening privilege, rather than a simple right for everyone and taking concrete steps to bring that privilege to them. In sum, it means to me that connecting with my students as humans is just as essential as believing in them as mathematicians.

4.2.3 Positive Mentorship

By listening to the simple, but sometimes unique things that my students need to succeed and acting on them, I've watched them rise beyond my highest expectations. Indeed, my interviewees talked to me about how life-changing it was for them to just encounter one or two positive mentors.

- 'My first vivid math experience was in second grade. My teacher was very excited about math and we just explored what she thought was cool. We sang math songs!

It made me think numbers are cool. I'm very competitive and like knowing I'm doing my best. I like to compare myself to others to see where I am now and think about how I could grow. Math allows me to do that: there's always so much more to know and so many aspects to math and different ways to do problems. Logic based stuff appeals to me more than straight-out calculus. I learned to play through math programmes where I learned there was a variety of math.'

- 'In middle school, I saw my 7th grade math teacher as my 'dad' at school. He was always there for us within math class. If you felt bad and didn't want to do something, he'd find out what was wrong instead of just disciplining you. He made sure we had one to one attention and we understood why the math worked, not just the procedures. As much as he could, he introduced you to different kinds of math. He took us to a college to introduce us to college students in STEM and told us that students of colour could succeed in STEM. He also showed us a commercial encouraging girls to do engineering. He was white. I think if he was black, he'd explain it from his own experience. My teacher couldn't do that, but he was always aware of his position and never overstepped the boundary, which was awesome. Sure, the message may have been a little more impactful from a person of colour, but I did still take his words with me. '
- 'My math teacher in 7th grade was a woman from Jamaica. I was so excited! I'd never had a woman of colour as a math teacher before. She pushed me really hard to be in an advanced class. She'd also give me math puzzles and didn't want us to be bored. She tried to make sure everyone was trying their best. '
- 'In college, I really enjoyed my linear algebra class. I could relate to my teacher because she was a woman teaching a college class, which is great. She also actually acknowledged me, understood that I knew stuff and kept her office hours consistent.

She was always patient with me, spent a long time talking to me and let me figure things out. She's the best math teacher I've had in college so far. Whenever I wasn't getting a topic, she'd recommend websites, notes and more problems. She was always open to giving advice for how to improve. '

- 'When I was a kid, I always looked up to my older brother. I was his shadow! His friends were the only people I wanted to hang out with. He thought I was annoying sometimes, but I did eventually make friends of my own. My brother and I did puzzles together. He went to a math summer programme and I did his puzzles that he brought home. He got me hooked on math puzzles, so when I was old enough, I applied for the same programme.'

Just like my interviewees, there have been a few mentors along my mathematical path who really believed in me. It's from their work ethic and positivity that I've gained the self-belief to persevere through concepts I've found challenging. I've learned to focus on the good things in the mathematical community around me and to work on math with other people. Despite encountering many negative experiences that caused me to doubt myself, the encouragement from my role models has really shone through. It's been a huge factor in me continuing to pursue the wonderful subject of math and to pass on what I know to my students. It's true that there are very real limits to what one teacher or one professor can do, but in my life and in my interviewees' lives, those interactions have made all the difference in the world. As a teacher, I've learned to accept those limitations, but also to mentor with optimism and with so much hope for my students' futures. By looking back into my life and the lives of my interviewees, I try my best to look forward into my students' mathematical lives and to impact them with all the positivity that I'm capable of.

4.3 Reflection through Teaching

This section isn't intended as any sort of guide on how to teach, or what interactions with students should look like. It's just a collection of stories: my stories and my students' stories, during times when we've inspired each other.

4.3.1 Stories of discipline

The math classroom is a human place. Just like any other human place, it's filled with people who have different needs, emotions and triggers. Because of this, things may not fall into place and students may not fall in line the way one could initially expect. In this section, I discuss examples of when I've needed to discipline a student. If I were to pick out one underlying theme in these stories, it would be the effectiveness in providing space for a student's emotions before trying to implement any sort of change.

Story 1

It was a summer's day when I took a teenage boy named K for a walk. He'd punched another student and much to his teachers' dismay wouldn't admit he was wrong. I didn't ask any questions. I simply made sure he had his water bottle with him because of the warm weather and we walked side by side in silence. His silence felt quizzical, as if he was testing me. I stayed silent until I was sure of what I wanted to say.

'You know, K', I told him, 'We all have things that are important to us. I remember an incident when I was in middle school...I felt I had no option at that time but to punch the other kid as hard as I could.' Abruptly K stopped walking and gave me a hug so tight it knocked the wind out of me. 'You understand!' he declared.

After that moment K opened up completely. We talked candidly about the internal forces that drove him to his decision. We problem-solved and came up with strategies for the future together. As his teacher, once we had mutual understanding and dialogue, I did hold K accountable and enforce consequences for him using violence on another student.

Something that I understood immediately when it became my task to work with K was that it shouldn't be about winning any debate with K about right or wrong. Rather, it should be simply about winning K over: not to my point of view or to agree with my values, but to the notion that the two of us are on the same team. I feel that too often, when teachers discipline a student, they jump too fast to explaining why the student was wrong. For K, the only step of the journey he needed my guidance on was from a place of defensiveness to a place of curiosity. The empathy and responsibility he took on next were all decisions he made independently.

Story 2

From the moment I began teaching her, I really liked B. Despite being young in years, it was clear B had a great awareness of who she was and what she wanted. The outward message B always projected (as one of my colleagues so eloquently put it to me once), was 'If you're respectful, I'll be really sweet. But if you mess with me, I'm going to f**k you up'.

One occasion I remember so clearly was during a class period where I wasn't teaching. B had disagreed with another teacher on how a learning activity should be run and refused to back down. The situation became messier and messier and eventually B had been dismissed from the class and I had been tasked with getting through to her.

As soon as I met with B, it was incredibly clear that she felt so very angry: she was standing in the hallway kicking the wall and cursing rapidly under her breath. Accordingly, I said 'Hi B' to alert her of my presence. I said nothing else for the time being. Instead, I sat on the floor of that hallway, began grading papers and just allowed the two of us to feel what needed to be communicated but didn't need to be said: that her energy levels were dropping gradually from boiling anger to a place where she felt more in control and communicative and that I was happy to meet her there when she was ready.

Eventually once B had reached a level of calm and engagement, she slid down to sit on the floor next to me. At this point, I pulled out a loose sheet of paper from my pile and began a few games of tic-tac-toe with her. When B cracked a smile at having beaten me for the second time in a row I asked her: 'B, You know that I think you're fantastic and that I have your back, right?' I waited for her to nod before continuing. 'The incident that just happened in class, what's up with that?' I asked for her version. No sooner had the words left my mouth that B's expression darkened again and she uttered a few choice expletives.

I paused a few seconds to let the atmosphere relax a little more again before continuing: 'It's totally ok to feel the way you do' I reassured her, 'But we need to figure out together what we can do next.' B again expressed adamantly that she didn't want to do the activity the teacher's way. 'Ok,' I thought for a few seconds, 'Well, what if we do it *slightly different way that still meets the teacher's learning objectives but is a bit closer to what B wanted*.' 'No,' she responded, this time thoughtfully, 'But we could.....' And so a constructive dialogue began.

After our conversation, B went back and talked things through with her teacher. From what I was happy to hear, it was a productive conversation that gave them both insight into collaborating more successfully in that class. B taught me through her non-verbal signals about the importance of giving students the space they need, while still being present for them. I felt a temptation initially to get an explanation from B for her behaviour, but then came to understand that I never needed one after all. So often, understanding a student isn't about forcing them to explain complex feelings and decisions that they themselves don't even fully get. It's simply about the two of us teaming up and saying 'what's next?'

4.3.2 Stories in different languages

I often tell my students that learning math is equivalent to learning a different language. It gives us a whole new vocabulary to discuss the beautiful patterns and phenomena we observe in the world around us with pinpoint precision. The language of mathematics has been evolving over many centuries into an extensive set of conventions now embraced through academia, dictating how mathematics should be written, thought about and expressed.

If we think in terms of languages of the world, whether a student's first language is English or Swahili or Korean has no bearing on their intelligence or ability. I believe the same is true about the language in which a student expresses mathematical thinking. How well-versed a student is in the conventions of academia and how easily they pick up on these conventions varies vastly depending on the individual. In this subsection, I discuss learning journeys with specific students who were strong mathematicians but who initially spoke math in different languages.

Story 1

At face value, E seemed to fit the stereotype of 'difficult student' perfectly. He was unengaged and uncooperative in class and wasn't turning in homework assignments. When I reached out to his other teachers, they uniformly reported that this behaviour was also typical in their classes.

Following up with E about homework was like pulling teeth. He shrugged, gave one word responses and asked repeatedly if he could leave. After pairing E with multiple members of the class and never seeing his productivity increase, it was time to dig deep and really try to understand him. 'E,' I approached him at the end of class one day, 'Could we meet up this afternoon? I'd like us to check in.' Cue shuffling and nervous glances everywhere but at me. 'Ok' he replied.

When E arrived, I asked him to explain how to do the first problem on the latest homework assignment. As he began to talk, in a tone that wasn't without enthusiasm, I noticed clear logical reasoning, good attention to detail and signs that he had in fact given the problem previous thought. I twisted the problem on the fly and asked E how he would approach the new modified problem. Again, his answers were coherent. He was reflective and coming up with fresh new ideas himself.

'Good job, E!' I high-fived him. 'Could you write down what you just told me on your worksheet?' E picked up a pen and looked blankly at his worksheet for a while. 'I don't know how,' he declared hopelessly. 'Well what did you just tell me? What's the first step to beginning the problem?' Slowly, painstakingly we arrived together at a written answer. During this process, I gained a little insight into how E's mind worked. Although an imaginative mathematician and a very able problem solver, E struggled with the multitasking process of holding complex layers of ideas in his head while simultaneously turning them into words to transcribe onto paper.

'Would you like to turn in future assignments as oral presentations?' I asked E at the end of our meeting. For a brief second, I saw his eyes light up as I'd never seen before. Outwardly, he simply shrugged and said 'Ok, I can do that.' From that day onwards, E came to every single one of my office hours without me ever needing to ask him to. Each time, he declared understatedly that he'd come to give his presentation. After reasoning through each problem out loud, we'd work to write at least one full answer down together. I became more mindful to direct E's time during class to coming up with ideas with his partner and just jotting them down as mind maps, allowing him to turn those into a more coherent written form after class, away from time pressure or fear of judgement from his peers. As time went on, E's in-class participation improved dramatically.

For E and I, it was never about avoiding the problem. It was about tackling it head on, while still honouring the type of learner E was and the type of mind he had. Once we found an environment that was right for him, E's dedication and resilience were incredible to witness.

Story 2

When people ask me why I love to teach, one story I often share is S's story. S is a student I'll never forget, because of his unwavering positivity, because of his shining mathematical talent, because of his generosity in helping other students, but above all, because of his incredibly unique background and style of communication...which actually held him back for so long.

The first time I talked to S outside of class was after the first quiz of the semester, which he had bombed: he hadn't even been able to get ten percent of the points. My goal was to find out if S had somehow been misplaced into the class, or if there was some other circumstance holding him back. I remarked gently that this quiz probably hadn't been his finest mathematical moment and handed him a piece of chalk. 'I'm just not good at math' S said dejectedly. 'Well, could we try the first problem again? I want to see a bit more of your thinking and I promise I'll help you if you get stuck.' S stared at the question for a long moment and told me that he just didn't know how to begin. I read the problem to S a few different ways. On the final reading, his eyes seem to light up and he made a tiny motion towards the board, only to draw back with an 'I'll probably get it wrong.' 'Well if that happens, we'll fix it together,' I prodded. To my amazement, S shyly picked up the chalk, began talking and within a minute had a perfect solution. We proceeded similarly through the rest of the quiz. 'Feeling a bit better about this quiz?' I asked S. 'Lots better,' he nodded.

After this one to one conversation, I began paying attention to S in class through a slightly different lens. Often I'd approach him and ask what question he was solving, since it was unrecognisable through looking at his work. He'd show me the question and I'd step back and remark: 'S, you're clearly solving something, but it's not this.'

As I got to know S's background, I realised we had a lot in common. He'd grown up speaking English (as had I), but, just like for me, it was a different form of English than white American English that textbooks and exam questions are written in. During my first semester living and teaching in the US, one of the biggest culture shocks for me was how differently math problems were worded than what I was used to-different enough that I'd often have trouble parsing what they were asking. The saving grace for me was that I knew all the math content I was teaching very solidly, I just had to get at what the questions were asking. For S, he was required to do both at the same time. The language that he spoke at home and with his friends simply never matched up to the language he saw on official exams at school. Through all his struggles, his mind had never quite identified the problem, nor figured out how to bridge the gap. To make matters worse, S's confidence took such a battering that often, even when he did understand the question, he'd still second-guess himself, convince himself he was wrong and either leave it blank or write something else entirely. And so, S had always struggled on tests. He told me that he wasn't good at math, but really, once he understood the English he was outstanding mathematically.

I began assessing S orally. It was essentially the same as giving him written assessments. He didn't need mathematical hints. What he did need was for us to read through the problems together before he began and to be able to ask a couple of clarifying questions about wording. Just by making this change, his scores began showing dramatic improvement. Even better, since parsing questions was a collaborative exercise each time, he began getting better and better at doing this himself. By the end of the semester, his written assignments and tests were unrecognisable from what he'd been turning in at the beginning. As his grades improved, so did his confidence. More and more, I started sending confused students to S, because he was growing into such an excellent resource to his classmates. It was clear he never forgot where he started off, because the patience and level of encouragement he displayed with his peers were simply inspirational. Over the course of the semester, S's tests improved by multiple letter grades and S himself grew into a mathematician who explained sophisticated ideas with clarity and self-belief.

Why do we need to make math a more diverse and inclusive space? There are many good answers to question, but one answer I'd give personally is 'because of young men like S.' There are many languages to express math in and many ways students can show flashes of brilliance. For those flashes to grow into a more constant light, different young mathematicians require different forms of mentorship.

4.3.3 Stories of getting to know each other

I wrote this section because I think it's of vital importance to first think of students as humans before considering their role in a math classroom and of myself as a human before my role as a teacher. I don't think that making my classroom a human place should involve pushing that any given level of 'openness' is the optimal one. What I do believe in is that it is just as important to get to know each of my students as people, as it is to get to know them as mathematicians. To me, this means honouring their boundaries, showing a little bit of vulnerability where appropriate and creating a space where learning is fun. I think in the end, we're all more authentic and more curious when we feel free to be ourselves.

Story 1

After working with several hundred uniquely wonderful students, whenever someone asks me for a story about a student who's inspired me, A is one who jumps out immediately. When I first met A, she was a quiet girl who flew under the radar. At the time, I was teaching a large, advanced math class in which A was one of only four female students. She never spoke in class, which prompted me to try to get to know her outside of class.

A and I understood each other straightaway. We found early on that we had multiple shared non-mathematical interests and also that we both felt the struggle of being an introverted woman of colour in math. At first our interactions were mainly talking through homework problems during office hours but they quickly grew into more broad math mentorship through walks on campus and ice cream sessions. The thing I appreciated mathematician A for the most was her amazing versatility: her ability to jump effortlessly from the discrete (as in topics from abstract algebra), to the continuous (as she explored real analysis problems). A showed herself to be a formidable on-the spot problem solver, who could also sit down for hours to untangle the complex details of someone else's jumbled proof. One of my favourite moments ever as a teacher is the first time, after several weeks of working together outside of class, A raised her hand during an in-class discussion.

Around the time when A began teaching herself topology and wowing me with how fast she was progressing, I started trying to get A more mathematical mentorship. However, whenever I brought her up to my colleagues, I received the same answer (in slightly different words) each time: 'I don't really know A. She's just so quiet and she never talks to me.' Over the next several months, I advocated for A for various opportunities, both verbally and through letters of recommendation. As A took hold of her opportunities, her mathematical confidence grew. Much later, she came back to TA for one of my classes and did an outstanding job, where she really listened hard to what each student had to say. Through our own experiences, A and I both knew what it was like to be a mathematician who wasn't heard. I'll forever treasure those precious people that listened to me when I first tentatively began and set me up on a journey to discover and fall in love with math. Years later, standing in my mentors' positions and listening to A was one of the biggest privileges I've ever had. If I could ask one thing of math teachers and professors, perhaps I'd phrase it as: 'give each of your students an opportunity to surprise you.'

Story 2

From my first class with him, I could tell that H was bright, talented and wanted to succeed. I could also tell that he was unhappy in the class. He was very much used to math being a solo activity where he had to do everything himself and so was resistant to asking for or giving help. H didn't participate in class discussions. Almost every time I heard H speak, it was to vent his frustrations about not understanding something, or not being able to solve a problem. The uncertainty of math problems being too challenging to solve by oneself, but becoming more manageable as a collaborative effort was an idea H struggled to come to terms with. I summoned H to a couple of meetings outside of class where we tried talking through making the transition and what this could look like. However I could still feel resistance from H.

It was around this time that I learned from another teacher that H had signed up for a small extra-curricular activity with only a few other students in it. Immediately, I saw this as my opportunity for an 'in' with him and so I signed up to TA for this activity, which actually proved to be fun and instructive in its own right. During this activity, I got to know H in a setting completely removed from a math classroom. Slowly H began to open up and I really got to know him as a person. From stories about his family, to which video games he liked to play, H transitioned from a student who would give one word responses to my questions in math class, to someone who was animated, excitable and very talkative.

As I hung out with H regularly during the extracurricular activity, I saw his engagement in class improve correspondingly. He began to participate in group discussions and if he was unhappy about a concept or a learning activity, he became more forthcoming in articulating what would make him more comfortable.

I really saw the turning point of this class for H to be midway through the semester, when I'd given out a worksheet with a particularly difficult challenge problem. There was something appealing in the problem to H and he became instantly hooked on it. I'd see him thinking about it and scribbling in his notebook between classes. Often he'd come to me to share new progress or to ask a clarifying question. It was clear that he was frustrated by the problem, but this frustration was different: it was productive, collaborative and rooted in curiosity. After a few days of hard work, H had solved the problem entirely. Since he'd now had some experience of speaking out loud about math to me and getting more comfortable with being stuck, I challenged him to step up as an informal TA and help other students with the problem. When the semester first began, H wouldn't work with anyone, but what he did at this point was outstanding. He thought about the problem from different angles and handled working with his classmates in such an encouraging, positive way. If their approach was different from his, instead of dismissing them, he really put in effort to understand their thought process and to see it through with them. This problem set launched H into a much more active class participant, who worked very well with multiple other students for the second half of the semester.

For me, I think there's nothing wrong with a student being an introvert, or often enjoying math in a setting of solitary contemplation. Indeed, this is something many mathematical researchers enjoy. However, I really believe that different students get to know authority figures at different speeds and the setting in which they're comfortable doing so may vary a lot. I would have been doing H a disservice if I had made negative assumptions about him, without providing him the setting he needed to challenge those assumptions. Getting to know H was wonderfully rewarding, both for myself and for what a huge positive contribution he grew to make as a member of my class.

4.3.4 Stories of not speaking, but simply understanding

Something that I think is absolutely wonderful about working with teenagers is that they're incredibly communicative and very real about how they're feeling. However, this communication is varied in style and medium and unravelling it can take profound listening and reflection. Just as students can communicate in different languages, they can also communicate without words. Sometimes finding the right words can be difficult, certain emotions seem not to have words and sometimes being required to speak can feel like too much to a student. In such scenarios, I think forcing the student to verbalise in any given way almost always does more harm than good to the their well-being. In fact, each time it's been necessary, I've found it a challenging but massively rewarding puzzle, to wordlessly communicate everything which needs to be said. Cracking such a puzzle in the past has led my students and I to the huge upside of an improved mutual connection.

Story 1

W was never a particularly vocal or outspoken student, but there was a certain charm about him that made him popular and well-liked among his peers. Often, I'd see him chilling with a small group of students, eyes twinkling, as if thinking about some private joke. In my classes, W was always that kid who sat quietly in the back, never causing trouble, never drawing attention to himself. When assessing his contribution to group work or his written assignments, they were always fair: never outstanding, never poor.

W never raised his hand in class, nor chose to attend office hours. Once I established from his other teachers that this was a pattern, I took it upon myself to just poke him once in a while and talk math. And so around once a week when I ran into him, I'd ask casually: 'Hey W! Want to talk math soon?' or 'How's this week's assignment for you? People have been telling me it's tricky. Want to stop by and tell me how it's going for you?' W would always accept such invitations quite cheerfully, but would never initiate such interactions himself.

Just like this, I had the privilege of getting to know W. I never felt I knew him too deeply, but I was very happy to know him at the level he seemed comfortable opening up. I got to know him as a young mathematician through sitting and listening to him explain his ideas to homework once a week. Once I put him at the board, he was surprisingly comfortable there and I learned that he's a student who catches onto new ideas remarkably quickly: much faster than I had known by just looking at his assignments. In addition, I got to know him a little as a person as he talked about his hobbies and once in a while, the ins and outs of his friendship group.

When I wrapped up my final class of the semester, my students very generously gave me a round of applause. A few students remarked that they'd miss the class. W didn't say anything. He put his things in his backpack and it seemed as if he'd simply walk out. However as W passed my desk, he stopped, made eye contact with me and stretched out his arms and hugged me. It was a long hug that lasted maybe 20-30 seconds, where I simply stood and allowed it to continue, where W non-verbally communicated the things he wanted to say. Prior to that interaction, W had never hugged me. After it was over, W walked silently out of my classroom and we haven't talked since. I feel really blessed to have been W's teacher. I really saw through him that student-teacher conversations don't always have to be with words or answered questions-sometimes just being present and authentic is enough.

Story 2

The first time I met R, she was crying. Another teacher had asked her what was intended to be a routine question, but this question had caused such anxiety that she'd shut down. Unable to speak, she'd retreated to a corner where she was unresponsive and further prodding had simply led to her cry harder. Since the teacher had multiple other students to attend to at the time, I volunteered to work with R. I walked over to her and said 'R, let's get out of here for a bit,' simply removing her temporarily from the physical setting she'd felt so much pressure in.

I took R to a pretty part of campus and we walked around that area in laps. At the beginning of the first lap, I told her: 'I'm not going anywhere. I promise we have all the time in the world. You do you.' And so, we walked for a long time. Once every few minutes, I'd make a friendly remark, ensuring it was never a question, or something R would feel an obligation to respond to before she was ready. I told her that I taught math and a little bit about the sort of math I found interesting. I pointed out one of the campus buildings that I knew. During this time, R had started to present as a little more relaxed. She made eye contact with me increasingly often. Eventually she declared without prompting: 'I think I'm ready to go back now.'

Later, I filled in R's teacher on how our interaction had gone. Because I'd been able to build a bit of rapport, R's other teachers enlisted my help a few more times in communicating with her. As I spent more time with her, I learned a little more about who she was. R wasn't a shy student by any means. She got very good results academically and loved to raise her hand in class and participate in discussions. However a couple of scenarios terrified her.

For R, making decisions was often a frightening and stressful process. When asked a question that required a decision to be made, she'd often shut down. And so, R and I would make decisions together. Often, I'd simply sit with her and help her list out what the options were. When I saw her grow more tense and speak less, I relied on her non-verbal cues. Observing her micro-expressions as I slowly read through each option helped us eliminate certain things. Another thing that helped when talking became more difficult was writing things down. I remember several silent conversations I've had with R where we just scribbled notes to each other on scratch paper. Some of these notes consisted of words, but other times, they would just be pictures: a face with an expression on it communicating how R felt at the time, or cartoon animals to lighten the mood and take a break from our task when things got too stressful.

Another aspect of classroom life that R found difficult was when she encountered a difficult math problem that she couldn't reason through on her own. The shame she experienced with this made her panic so much that eventually, it shut down her ability to continue working or communicating productively. In this case, having her write down what was confusing, instead of having to talk about it helped her articulate. I found that she also found it intimidating to just work with a a teacher one on one when this was going on, so I enlisted help from her friends. This just involved me sitting with R and one other student and asking that other student to give a few thoughts on the problem. Hearing a peer talk through things gently helped ground R to focusing on the problem at hand. It also often helped the other student understand the problem just that bit more deeply.

I really enjoyed all my conversations with R. She often had fresh, interesting ideas on how to think about math concepts. I also discovered she was a very well-rounded academic, often telling me exciting things she'd learned in other classes and how she'd embarked on independent mini-projects taking these concepts further. While many of our conversations were non-verbal and communicated via drawing pictures or through body language, I appreciated her candour and how she always tried her best to be open with me. So often, as teachers, we can fall into the trap of pushing a student to give an immediate answer to a question, or of communicating in a way we're expecting. I was privileged to meet in R an interesting, thoughtful young woman whose only difference from her peers was that sometimes she spoke without words. Through her, I learned a different way of talking.

4.3.5 Stories of setting differences aside

I find that my classroom is a more curious, accepting and connected place when there is room for disagreement. Some of the most interesting and important lessons I've learned as a teacher have been from students who have very different views from me on learning or on the world around them. Putting myself in these students' shoes, I've reflected on how I can get the best out of them despite our differences. I've worked to help them find their place and sense of belonging in my class and do good productive work, within a culture of mutual respect. Throughout this process, I've really internalised that the more I give my students the space they need to decide on their own opinions, the more willing they become to adjust those in light of what they learn from me and from their peers.

Story 1

What causes racism? Many incredibly wise people have already tried to answer this question and so, I don't necessarily feel qualified to give an answer. In my opinion, although racism is widespread and often systemic, just like anything else in the world, each individual's opinions are informed by their unique circumstances and then by the choices they make inside those circumstances.

A student that I'll never forget is T, a white boy who I tried my best with as a teacher and who I really believe eventually tried his best too. When T first came into my class, it was clear he was taken aback to have me, a woman of colour, as his math teacher. Even when I gave an instruction that had been understood by all his classmates, he'd sometimes accuse me (a native English speaker) of not speaking English well enough for him to follow. He often challenged my authority. Several times, he was outwardly aggressive and I remember one occasion when he responded to a difference in opinion with me by throwing chalk at me. T's disrespect for his classmates of colour was clear. When assigned to a group with a student of colour, he would refuse to work with them. When I inquired why, he told me that cooperation was impossible because the student in question (despite actually being a native English speaker) didn't speak English.

Despite T's viewpoints and his outward hostility, I picked up early on that this was a student who really wanted to do well in my class. I also realised that having an ethics debate with him would only antagonise the situation. I needed to challenge T's viewpoints not through words, but through actions. And so the first thing I did was meet with T individually, where I communicated exactly what my expectations were for his conduct and what the precise consequences would be in cases when they weren't met. I identified a mutual starting point: 'T, do you agree that the classroom should be a space where every student has the right to feel safe?' and pointed out a few specific behaviours I'd observed from him, without voicing any opinions or judgements. 'Here's what I'm going to do going forward,' I informed him. 'If you raise your voice or throw objects when speaking to me or to another student during class, I'll send you out for a short five minute time out. There's no judgement here, just a chance for you to reset and think about how you want the rest of the class period to go. If it happens again, you're out for the rest of that class. We'll talk after class about whether it's appropriate for you to come back the following class and what changes need to be made.'

After our conversation, it was very important to act on my words. Whenever T overstepped boundaries, I always responded immediately by enforcing time out. To his credit, he responded well, always returning from time out ready to work and I never had to dismiss him for an entire class period. Although sometimes it felt like one step forward, two steps back, things did improve slowly. One thing I noticed at this point was that T was having trouble finding his footing socially in the class and his lack of support compounded his frustrations. After considering my roster for a very long time, I switched up seating in that class. The student I placed T next to was mild-mannered and mellow individual, who clearly exuded a quiet, unshakeable confidence. I was optimistic, but by no means certain it would work well so I was definitely excited when it did.

T started to do good work with this classmate and through him began having more positive interactions with other students. They'd come to office hours together and slowly T started to open up. T came from a small, rural midwestern town where he'd never had a non-white classmate before, let alone a person of colour as an authority figure. His family struggled financially and so T valiantly balanced schoolwork and taking on odd jobs to help his family out. He saw a good education as a ticket out of poverty and so put a lot of pressure on himself to succeed academically. I really came to admire T's courage and persistence and I feel that I earned his respect over time by consistently communicating and holding him to high standards behaviourally, while never telling him which opinions he 'should have' or how he 'should feel.'

Once T had a small group of classmates he worked well together with, I started slowly helping him integrate into the class community as a whole, getting to know those students who he'd originally believed were so different from him. I started small: if T asked me a question, I'd call on another student to give a one-line hint. This built up to having that student sit with T for a couple of minutes under my supervision and explain their solution and eventually the two of them working on a difficult problem together from beginning to end. Throughout this I felt T's viewpoints on race shift little by little, just from gaining first-hand experience of being in a diverse group of people, helping and motivating each other inside a math classroom.

T definitely had opinions that I found challenging and strongly disagreed with. However, I learned that unfamiliarity and fear of the unknown played a bit part in forming those opinions. When I look back, I feel like it was one of my most rewarding teaching experiences finding ways to challenge T's world views and broaden his perspective and in that process, actually broaden my own perspective. I learned that my math classroom can be about learning so much more than just math. Through choices and actions it can be a space where students take a step forwards from who they are as a person towards who they can become.

Story 2

From the first day of class, it was clear that J didn't want to be there. He was uncooperative and mutinous in so many ways. When asked to work on problem X, he'd make comments just loud enough to be heard that problem Y would be a better choice. When asked to stand, he'd remain in his seat and during times when asked to sit at his desk, he'd suddenly have the most important reason to walk across the classroom. When asked to work in a pair, he'd inevitably form a group of four or five and then waste no time telling his whole group that he thought the whole exercise was pointless.

After an unsuccessful first quiz, I met one to one with J and asked him to do his corrections in front of me. J laughed sheepishly, a first sign of vulnerability and insight into his classroom attitude so far. 'You see' he said 'It's been a while since I did math and when I did it before, it was different from this.' 'How so?' I asked him. I came to understand that while J had done well at his previous math classes, they had all been very much in the flavour of memorising formulas and solving a very narrow set of problems on tests just like the ones in class. 'I'm on your side and I want you to do well,' I reassured J, 'But I wonder...do you have any ideas why we do things the way we do in our class?' We talked a little about growth mindset and mathematical independence before completing the rest of J's corrections.

After that initial talk, J's attitude in class was noticeable better and his quiz scores began showing encouraging early signs of improvement. However what I noticed from J was that the confidence still wasn't there. He was still overall a passive participant in his own learning, often just nodding along as he watched the math happen. One day, as I passed J and his partner in class, I suggested they switch over writing so that J held the chalk and put their ideas on the board. Although initially reluctant, I noticed that he did great and as the class went on, I saw him happier and more confident than I'd ever observed before. However this didn't last: after a class or two, J reverted back to passivity.

Determined to hold onto the progress we'd made, I wrote J a note one day and slipped it onto his desk before class started. My note simply said 'Did you know that I ask you to write on the board because you're great at it? I really enjoy seeing your awesome ideas in action.' As J read the note, his expression changed. He got out of his seat, turned to his partner and said 'Give me the chalk!' From that moment onwards, J took control of his own learning. Often he and his partner would both be working at the board together, each brainstorming their ideas and bringing them together to get a final polished set of solutions. However J never relinquished his chalk, nor his initiative to think hard and to put his own ideas on the board ever again.

By the end of the semester, J had grown into an excellent young mathematician: confident and assertive, curious and exploratory. He brought a struggling friend from a different class to our final class together. When his friend inquired why I wasn't telling them how to solve the problem, J's response made me so proud to have taught him. He said: 'That's not how we do things in this class. Here, let me show you!' before taking his now-familiar spot at the blackboard and continuing on right where he'd left off last class with his mathematical exploration.

Change is scary to all of us, as teachers and as students. Sometimes there can be quite a bit of pain in growth and I've come to appreciate just how important it is to honour that when asking students to adapt. There are so many students like J out there: who can rise to expectations higher than they'd ever imagined, just as long as those expectations are built out of concrete tangible goals. 'J, I need you to take more initiative' is a hard unattainable goal that's in the sky, but 'J, I would like you to brainstorm your ideas for this problem on the board' is very attainable and gave both of us a solid starting point.

4.4 Conclusion

For a long time, I thought about how I should conclude this chapter. I guess the hardest part was that it didn't feel right to conclude something which is by no means over for me. So in the end, maybe instead of a conclusion, I should think of this as a new beginning. It's a beginning where I'll take the memories from my wonderful, inspiring students I've taught in the past into my teaching future.

As I leave graduate school, the place where I discovered the infinite joy in being a teacher, I'll carry forward with me the wise words of one of my teaching mentors: 'Teaching is about living in the grey area.' Through time, I've reflected and come up with my own interpretations to these thought-provoking words. If I were to summarise here, I would say it means letting each individual student's personality come out and allowing each collective class to shape what they want the culture of their learning environment to look like. It means very occasionally throwing plans out of the window and even more occasionally needing to break hard rules. In sum, I think of the 'grey area' as avoiding telling my students 'you should...' and instead saying 'why don't you show me?'

When I set out to write this chapter, I had no lofty aspirations of judging or talking about math education as a whole, I simply wanted to give myself space to reflect. I've reflected on my classrooms, on office hours, on taking walks with my students, on eating dessert, dancing, making paper airplanes, cooking, writing silly poetry with them and on those raw, authentic moments when I've walked the line between being an authority figure and a human mentor with vulnerabilities, not so far from my students' own. Similarly, as I now think about what my 'grey area' looks like, I think of a space with infinite possibility and only a few absolutes. My absolutes are the following:

1. Students will always communicate, but it may be at an unexpected level, or in an unexpected language. The best I can do is to hear them in the way they're asking to be heard.
2. Students will always rise to high expectations, just as long as we're ascending together, staying grounded in reality and starting from a place that they know.
3. Students will always be amazing, in their own fresh, unique, individualistic ways. The best I can do is to help them tap into that, so that they can make decisions which they themselves believe in.

I've learned to hold the unconditional belief that my students are the best, regardless of how much math they know at the beginning of the semester, or how their struggles outside the classroom affect them as learners. This doesn't mean that I should turn a blind eye to discipline issues, or that I should always agree with them or always give them the highest grades. Rather, it means holding them to the highest standards behaviourally as members of my classroom community and being consistent in communicating what my expectations are. It means challenging them to become more independent, more empathic and more curious about their academics and about the world around them. It means taking action to hold clear boundaries. Above all, it means really hearing my students when they disagree with me and letting them make their own mistakes and experience consequences (good or bad!) for their choices. Through all this, it’s my job to instil in my students the conviction that I have their backs and that as a classroom community, we’re always on the same team. My students have shown me time and time again just how high they can climb and how much our belief in each other fuels them to succeed. They’ve also shown me that when we look back, such success nearly always began with just one conversation, or one worksheet that seemed so insignificant at the time.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] E. Bank and L. Bary-Soroker. Prime polynomial values of linear functions in short intervals. *Journal of Number Theory*, **151**:263–275, (2015).

[2] P. Belmans. Algebraic geometry fun facts for the festivities: genera of complete intersection curves. https://pbelmans.ncag.info/blog, (updated March 1, 2020).

[3] A. Bertram. Complex algebraic geometry: Smooth curves. https://www.math.utah.edu/ bertram/, (updated Spring 2020).

[4] J. Bober, D. Fretwell, G. Martin, and T. D. Wooley. Smooth values of polynomials. *Journal of the Australian Mathematical Society*, **(108)**:245–261, (2020).

[5] Z. I. Borevich and I. R. Shafarevich. Number theory. Academic Press Inc., (1966).

[6] L.E. Bush. An asymptotic formula for the average sums of digits of integers. *American Mathematical Monthly*, **47**:177–236, (1940).

[7] K. Conrad. Proofs by descent. https://kconrad.math.uconn.edu, (updated March 9, 2020).

[8] C. de la Vallée Poussin. Recherches analytiques sur la théorie des nombres premiers. *Annales de la Société Scientifiques de Bruxelles*, **20B**:183–256, (1896).

[9] C. de la Vallée Poussin. Sur la fonction zeta de riemann et le nombre des nombres premiers inferieur a une limite donnée. *Mémoires Couronnés de l'Academie de Belgique*, **59**:1–74, (1899).

[10] H. Delange. Sur la fonction sommatoire de la fonction: Somme des chiffres. *L'Enseignment Mathématique*, **21**:31–47, (1975).

[11] P. Deligne. La conjecture de Weil: I. *Publications mathématiques de l'I.H.É.S.*, **43**:273–307, (1974).

[12] M. P. Drazin and J. S. Griffith. On the decimal representation of integers. *Proceedings of the Cambridge Philosophical Society*, **48**:555–565, (1952).

[13] S. Gannon. Autoethnography. Oxford Research Encyclopedia of Education, (2017).

[14] S. Ghorpade and G. Lachaud. Etale cohomology, Lefschetz theorems and number of points of singular varieties over finite fields. *Moscow Mathematical Journal*, **2**:589–631, (2002).

[15] J. Hadamard. Sur la distribution des zeros de la fonction $\zeta(s)$ et ses consequences arithmetiques. *Bulletin de la Société Mathématique de France*, **24**:199–220, (1896).

[16] K. Ireland and M. Rosen. A classical introduction to modern number theory. Springer, (1982).

[17] A. Kumar and S. Kumaresan. A basic course in real analysis. CRC Press, (2014).

[18] J. C. Lagarias. Euler's constant: Euler's work and modern developments. *Bulletin of the American Mathematical Society*, **50**:527–628, (2013).

[19] J. C. Lagarias and H. Mehta. Products of binomial coefficients and unreduced Farey fractions. *International Journal of Number Theory*, **12**:57–91, (2016).

[20] M. Levine. A mind at a time. Simon and Schuster, (2002).

[21] J. J. Y. Liang and J. Todd. The stieltjes constants. *Journal of Research of the National Bureau of Standards-Mathematical Sciences*, **76B**:161–178, (1972).

[22] P. Pollack. A polynomial analogue of the twin prime conjecture. *Proceedings of the American Mathematical Society*, **136**:3775–3784, (2008).

[23] V. Prasolov. Polynomials. Springer, (1999).

[24] Andrzej Schinzel. On two theorems of Gelfond and some of their applications. *Acta Arithmetica*, **13**:177–236, (1967).

[25] L. Schoenfeld. Sharper bounds for the chebyshev functions $\theta(x)$ and $\psi(x)$. *Mathematics of Computation*, **30**:337–360, (1976).

[26] G. Tenenbaum. Introduction to analytic and probabilistic number theory. Cambridge studies in Advanced Mathematics, (1995).

[27] D. Wan. Generators and irreducible polynomials over finite fields. *Mathematics of Computation*, **66**:1195–1212, (1997).

[28] A Weil. Numbers of solutions of equations in finite fields. *Bulletin of American Mathematical Society*, **55**:497=508, (1949).

[29] A. Winterhof. Character sums, primitive elements and powers in finite fields. *Journal of Number Theory*, **91**:153–163, (2001).

[30] M. Matchett Wood. On the probabilities of local behaviors in abelian field extensions. *Compositio Mathematica*, **146**:102–128, (2008).