WXML Final Report: Primality of Polynomials

William Stein, Travis Scholl, Astrid Berge, Daria Micovic, Xiaowen Yang

Autumn 2016

1 Introduction

The density of certain types of primes is a classical question in number theory. The Prime Number Theorem statement says that the prime counting function $\pi(x)$, which gives the number of primes $\leq x$, is asymptotic to $\frac{x}{log(x)}$. It was conjectured that a better asymptotic approximation to $\pi(x)$ was the offset logarithmic integral $Li = \int\limits_2^x \frac{dt}{\log t}$. In 1962 Paul T. Bateman and Roger A. Horn provided a formula for the density for the positive integers at which a set of polynomials have prime values. The formula is as follows:

$$\pi_f(x) \sim \frac{C}{D} \int\limits_2^x \frac{dt}{\log(t)}$$

This formula differs from the prime-counting function by the constant $\frac{C}{D}$ where D is the degree of the polynomial, $n_p(f)$ is the number of solutions of $f(x) \equiv 0 \mod p$, and

$$C = \prod_p \frac{1 - \frac{n_p(f)}{p}}{1 - \frac{1}{p}}$$

where p is prime.

1.1 The initial problem

We began by experimenting with the conjecture by simply plugging in polynomials and finding the associated constant. This helped us to better understand what was going on and what the precise meaning of the conjecture was. We decided to explore the mysterious constant C further, focusing on maximizing it, as well as the error term associated with the conjecture.

2 Progress

2.1 Computational

2.1.1 C values

In order to improve the run time of our SageMath computation, we looked at ways of finding $n_f(p)$ using the Legendre symbol. Previously in our computation we had used a SageMath command in order to solve $f(x) \equiv 0 \mod p$ for the first 100 primes, which gave $n_f(p)$. Although this worked, the run time was not ideal. Focusing on quadratic polynomials, we then studied the law of quadratic reciprocity and the definition of a quadratic residue modulo p. This allowed us to use the Kronecker symbol instead, finding $(\frac{\Delta}{p})$=$n_f(p)$, where Δ is the discriminant of the polynomial as given by the quadratic formula and p is an odd prime number. We found that this gave the same value for C as our previous computation for quadratic polynomials but worked much faster.

2.1.2 Error term

The error term is approximated by:$|\pi_f(x) - \frac{C \cdot x}{D \cdot \log(x)}| < k \cdot \sqrt{x} \log(x)$ Where k is our error constant. For our computation of the error term constants we manually tested constants to find one that would give a good approximation.

polynomial	C-Value	error constant
$x^4 + 5x + 1$	2.3335	1/7
$x^4 + 7x + 1$	2.5729	1/7
$x^4 + 5x + 5$	3.3515	1/7
$x^4 + 15x + 5$	1.1224	1/17
$x^4 + 5x + 15$	0.8809	1/20
$x^{15} + 5x + 1$	3.1280	1/15

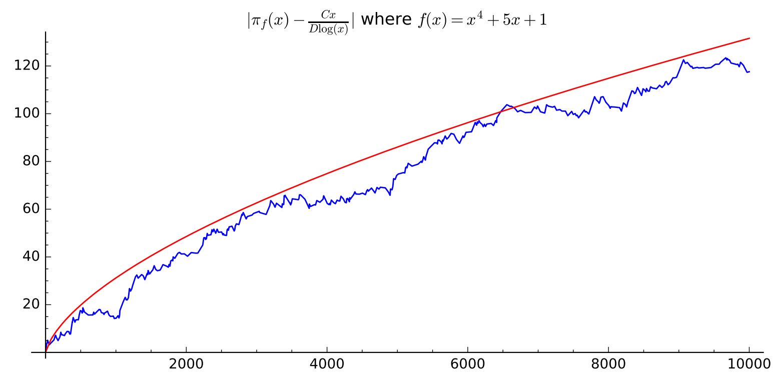

Figure 1: Error term for $\pi_f(x) < \frac{1}{7}\sqrt{x}\log(x)$

In this polynomial: $x^4 + 5x + 1$, we tried to find the best error constant which fits for the equation and we approximated 1/7. In the future, we could try to find a procedure to solve for the error constant or find a pattern.

2.2 Theoretical

2.2.1 Bunyakovsky condition

Initially we looked at the Bunyakovsky conjecture, which postulates that given a certain set of conditions, a polynomial will produce infinitely many primes. We attempted to create polynomials satisfying the Bunyakovsky condition. In order to satisfy the Bunyakovsky condition, a polynomial must have relatively prime coefficients, must not factor, and hold that $f(x)$ is not divisible by p for any prime p. We created polynomials that satisfied the Bunyakovsky condition and compared their C values to those that did not. We noticed that polynomials that factor gave a C value of 0.

2.2.2 C values

One of our goals was to find a polynomial that generates large C values, indicating that it is particularly prime-dense. We used the Chinese Remainder Theorem in conjunction with the Kronecker Symbol$(\frac{\Delta}{p})$ to find non-quadratic residues mod p. Through this we were able to reverse-engineer a polynomial that produced a high C-value. This polynomial was of the form x^2 minus an integer that was about 100 digits long. We generated a very non-compact polynomial and decided to change the direction of our exploration of the C

value. From here we did brute force trial-and-error computation of C values where we wrote Sage algorithms that tested different polynomials both sequentially and randomly. We looked at the trends of C values they generated. Doing so allowed us to make observations about how quadratic polynomials with the same discriminant generated the same C values, which naturally follows from our computation that relied on Δ to compute $n_f(p)$. This can be seen in Figure 1 where the y-axis shows the C values and the x-axis shows the index of polynomials is of the form $x^2 + bx + c$ where the range for b and c shown:

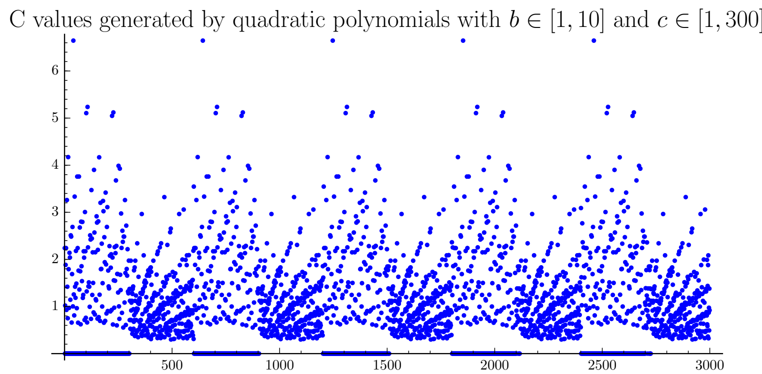

Figure 2: Looking for patterns in C

We were also able to draw a correlation from the constant term of the polynomial to the C value. We noticed from our data that the polynomials that generated the highest C values had one thing in common: the constant term was always prime. We found that the mean C value for the set of polynomials with prime constants was approximately 1.9494, compared to the C value for all the polynomials we tested which was approximately 1.3263.

An interesting result is that, although we would expect the mean C value to be 1 (since $\pi(x)$ has a C value of 1), we noticed that the mean C value for this sample of over 33,000 quadratic polynomials was approximately 1.3263. A possible reason for this is the fact that we limited our study to monic polynomials.

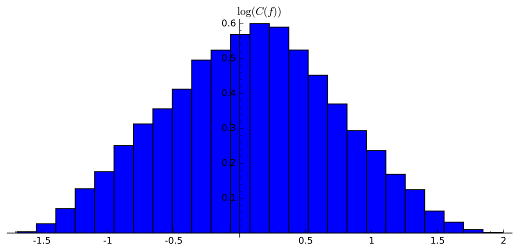

Figure 3: $\log(C(f))$ for 33,488 randomly produced quadratic polynomials

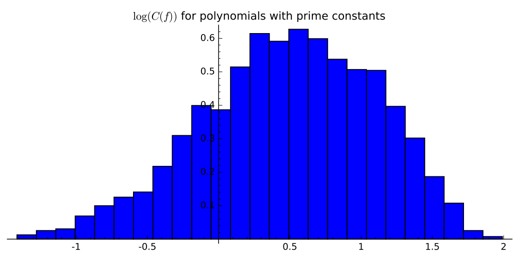

Figure 4: $\log(C(f))$ for polynomials whose constant term is pseudoprime

3 Future directions

Some of our future goals include refining our study of the error term as well as looking into C values for polynomials of degree greater than 2. Additionally, while we know it can be shown that the infinite product generating C converges, we would like to present our own proof for it.

References

[1] P. T. Bateman and R. A. Horn. A heuristic asymptotic formula concerning the distribution of prime numbers. *Math. Comp.*, 16:363–367, 1962.

[2] P.-J. Cahen and J.-L. Chabert. *Integer valued polynomials.* American Math. Soc., 1997.

[3] I. Rivin. Some experiments on Bateman-Horn. *ArXiv e-prints*, Aug. 2015.

[4] W. Stein. *Elementary number theory: primes, congruences, and secrets.* Undergraduate Texts in Mathematics. Springer, New York, 2009. A computational approach.

[5] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 6.10)*, 2016. `http://www.sagemath.org`.