THE NUMBER OF NONZERO COEFFICIENTS OF POWERS OF A POLYNOMIAL OVER A FINITE FIELD

CAROLINE ELLISON

Abstract

Coefficients of polynomials over finite fields often encode information that can be applied in various areas of science; for instance, computer science and representation theory. The purpose of this project is to investigate these coefficients over the finite field F_p. We find four exact results for the number of nonzero coefficients in special cases of n and p for the polynomial $(1 + x + x^2)^n$. More importantly, we use Amdeberhan and Stanley's matrices to find what we conjecture to be an approximation for the sum of the number of nonzero coefficients of $P(x)^n$ over F_p. We also relate the number of nonzero coefficients to the number of base p digits of n. These results lead to questions in representation theory and combinatorics.

Introduction

The Sierpinski triangle, a well-known fractal, is made when the entries in Pascal's Triangle are taken $\pmod 2$. Pascal's Triangle, in turn, is made up of the coefficients of the polynomial $(1+x)^n$ listed in rows, one row for each value of n. The Sierpinski triangle raises the question: what if this were done for other polynomials? This paper examines this problem, looking at the coefficients of the polynomial $(1+x+x^2)^n$ and other polynomials over finite fields. For instance, in $F_2[x]$, all the even coefficients become zero, leaving only the odd coefficients as nonzero. In $F_3[x]$, coefficients not divisible by 3 are nonzero. By researching coefficients in $F_p[x]$, we look at the question of how many coefficients will be nonzero for any prime p.

There have already been results on these questions for $(1+x+x^2)^n$ for a few cases of p: F_2 and F_3 [2]. Both cases have formulas involving the base p representation of n. We looked at the problem for general p, hoping to go beyond F_2 and F_3. In working on this problem, we took two different approaches. The first was to look for an exact formula for $(1+x+x^2)^n$, which seems to be hard.

The second approach was to focus on an approximation. On initial instinct, it might seem that all coefficients are approximately equally likely. However, when working in finite fields, this is not the case. In F_p, $(a+b)^p = a^p + b^p$ because the other coefficients are divisible by p: this fact means that many more coefficients are

zero than might be expected if we thought that roughly $\frac{1}{p}$ of the coefficients would be divisible by p.

Amdeberhan and Stanley [1] have a formula allowing the number of coefficients in any polynomial $g(x)^n$ equal to a number α in F_p to be found as a product of certain matrices. Willson [3] has related the number of nonzero coefficients of $f(x)^n$ in $F_p[x]$ to the dimension of an additive cellular automaton, and computed this dimension using a matrix that turns out to be related to Amdeberhan and Stanley's matrices in a way that will be explained in the last section. The Sierpinski triangle is an example of a fractal that is created when the grid size of such a cellular automaton goes to zero. We use these related results to examine the asymptotic behavior of the number of nonzero coefficients of a polynomial to the n^{th} power.

Structure of the Article. First, in section 1, we will present some previous results. Then, in sections 2 and 3 we will present orignal results. Exact results in section 2 include these four results:

(1) generalization of the F_3 formula for all p with the polynomial $(1 + x + \ldots + x^{p-1})^n$

(2) a formula when the base p expression of n contains certain digits

(3) an answer for all p for selected values of n

(4) expressions for the coefficients when $1 + x + x^2$ can be factored mod p

In section 3, we analyze the number of nonzero coefficients of the n^{th} power of polynomials as n goes to infinity. Results include statement and proof of our main theorem, a formula when $n = p^k - 1$, and a conjecture for all values of n supported by computer evidence. We use the $p = 2$ case as an example.

Finally, we discuss further opportunities for research in section 5.

1. Background

Our results build on results by Amdeberhan and Stanley. For every polynomial $f(x)$, define $f_p(n)$ as

$$f_p(n) = \{\text{number of nonzero coefficients of } f(x)^n \text{ in } F_p\}.$$

For $p = 2$ and 3, they have proved the following [2]:

Proposition 1. *If*

$$2n = \sum_{i=0}^{k} a_i 3^i$$

is the base 3 expansion of $2n$, then the number of nonzero coefficients of $f(x) = (1 + x + x^2)^n$ over F_3 is equal to

$$f_3(n) = \prod_{i=0}^{k}(1 + a_i).$$

Proposition 2. *If we write n as $n = \sum_{i=1}^{r} 2^{j_i}(2^{k_i} - 1)$ where $j_i > k_{i-1} + j_{i-1}$, breaking it up into 1-strings, then the number of nonzero coefficients of $f(x) = (1 + x + x^2)^n$ in F_2 is equal to*

$$f_2(n) = \sum_{i=1}^{r} f(2^{k_i} - 1),$$

where

$$\begin{cases} f_2(2^k - 1) = \frac{2^{k+2}+1}{3} & \text{if } k \text{ is odd} \\ f_2(2^k - 1) = \frac{2^{k+2}-1}{3} & \text{if } k \text{ is even.} \end{cases}$$

These results provide motivation to explore the question for general p. The first proposition serves as the model for our generalization to $(1 + x + \ldots + x^{p-1})^n$.

2. Formulae for $f_p(n)$

This section presents exact results for $f_p(n)$, looking at special cases of n and p .

2.1. Generalization to $(1 + x + \ldots + x^{p-1})^n$.

We look at the number of nonzero coefficients of the polynomial $(1 + x + \ldots + x^{p-1})^n$ over $F_p[x]$ in order to generalize the $p = 3$ case stated before. What we find is the following:

Proposition 3. *If*

$$n(p-1) = \sum_{i=0}^{k} a_i p^i$$

is the base p expansion of $n(p-1)$, then the number of non zero coefficients of $f(x) = (1 + x + \ldots + x^{p-1})^n$ is equal to

$$f_p(n) = \prod_{i=0}^{k} (1 + a_i).$$

Proof. Since for $g(x)$ in $F_p[x]$ we have $g(x)^p = g(x^p)$, we have

$$\begin{aligned}
&(1 + x + \ldots + x^{p-1})^n \\
&\qquad = \quad (\frac{1 - x^p}{1 - x})^n \\
&\qquad = \quad ((1 - x)^{p-1})^n \\
&\qquad = \quad (1 - x)^{n(p-1)} \\
&\qquad = \quad \sum_{k=0}^{np-n} (-1)^k \binom{np - n}{k} x^k.
\end{aligned}$$

By Luca's Theorem, if $np - n = \sum_{i=0}^{k} a_i p^i$ and $k = \sum_{i=0}^{l} b_i p^i$ then

$$\binom{np - n}{k} = \prod_{i=0}^{k} \binom{a_i}{b_i}.$$

The coefficients will be nonzero when all the terms in the product are nonzero. Each term is nonzero if $a_i \geq b_i$, so there are $a_i + 1$ ways for each term to be nonzero and $\prod_{i=0}^{k}(a_i + 1)$ nonzero coefficients. □

2.2. **Case Where n has Certain Digits.** As we have seen in Section 3, for the number of nonzero coefficients of the polynomial $f(x) = (1 + x + x^2)^n$ over $F_p[x]$, the smallest unsolved case was $p = 5$; looking at it leads us to an interesting general result involving all p. We looked at the number of nonzero coefficients of $f(x)^n$ in F_5. If $n = \sum_{i=0}^{k} a_i 5^i$ is the base 5 expansion of n, then

$$\begin{aligned}(1 + x + x^2)^n &= (1 + x + x^2)^{\sum_{i=0}^{k} a_i 5^i} \\ &\equiv \prod_{i=0}^{k} (1 + x^{5^i} + x^{2\cdot 5^i})^{a_i} \pmod 5.\end{aligned}$$

We would have liked to say that $f_5(n) = \prod_{i=0}^{k} f_5(a_i)$, because calculation would be very easy. However, this is not true. It turns out that it is true if $a_i \in \{0, 1, 2\}$. This led us to find more generally that this proposition is true if $f_p(n)$ denotes the number of non zero coefficients of $(1 + x + x^2)^n$.

Proposition 4. *If $n = \sum_{i=0}^{k} a_i p^i$ is the base p expansion of n, and if $a_i \in \{0, 1, \ldots, \frac{p-1}{2}\}$, then*

$$f_p(n) = \prod_{i=0}^{k} f_p(a_i).$$

Proof.

$$\begin{aligned}(1+x+x^2)^n &= (1+x+x^2)^{\sum_{i=0}^{k} a_i p^i} \\ &= \prod_{i=0}^{k}(1+x+x^2)^{a_i p^i} \\ &= \prod_{i=0}^{k}(1+x^{p^i}+x^{2p^i})^{a_i}.\end{aligned}$$

If this product is expanded, consecutive terms will be

$$(1+x^{p^i}+x^{2p^i})^{a_i}(1+x^{p^{i+1}}+x^{2p^{i+1}})^{a_{i+1}}.$$

The powers of x in the first term go from p^i to $2a_ip^i$. Similarly, the powers in the second term go from p^{i+1} to $2a_{i+1}p^{i+1}$. If $a_i \leq \frac{p-1}{2}$, then $2a_ip^i < p^{i+1}$. Because all the powers of x in the first term are less than those in the second term and the base p expansion of n is unique, there will be no interactions between terms. The number of nonzero coefficients in each term is $f_p(a_i)$, and with no interactions

$$f_p(n) = \prod_{i=0}^{k} f_p(a_i).$$

□

2.3. **Special Cases of n.** Long division has allowed us to obtain numbers of coefficients for specific values of n. For instance, the coefficients of $(1+x+x^2)^{p^k-1}$ alternate, the pattern being $1, -1, 0, 1, -1, 0, \ldots$ until it reaches the middle, at which point the coefficients are symmetric over reflection. This means that around two out

of every three coefficients are nonzero: specifically, the number of coefficients is $\frac{4p^k-1}{3}$ when $p^k \equiv 1 \bmod 3$ and $\frac{4p^k+1}{3}$ when $p^k \equiv 2 \bmod 3$.

We were able to discover formulas for $n = p^k - 2$, $p^k - 3$, and $p^k - 4$ as well, though the coefficients got more complicated.

If $n = p^k - 2$, then

$$\begin{cases} f_p(n) = 2p^k - 2p^{k-1} - 1 & \text{if } p \equiv 1 \pmod 3 \text{ or } k \text{ is odd} \\ f_p(n) = 2p^k - 2p^{k-1} + 1 & \text{if } p \equiv 2 \pmod 3 \text{ and } k \text{ is even.} \end{cases}$$

If $n = p^k - 3$, then

$$\begin{cases} f_p(n) = \frac{1}{3}(6p^k - 10p^{k-1} - 5) & \text{if } p \equiv 1 \pmod 3 \text{ or } k \text{ is odd} \\ f_p(n) = \frac{1}{3}(6p^k - 10p^{k-1} + 5) & \text{if } p \equiv 2 \pmod 3 \text{ and } k \text{ is even.} \end{cases}$$

If $n = p^k - 4$, then

$$\begin{cases} f_p(n) = 2p^k - 6p^{k-1} - 2 & \text{if } p \equiv 1 \pmod 3 \\ f_p(n) = 2p^k - 6p^{k-1} + 1 & \text{if } p \equiv 2 \pmod 3 \text{ and } k \text{ is even} \\ f_p(n) = 2p^k - 6p^{k-1} - 1 & \text{if } p \equiv 2 \pmod 3 \text{ and } k \text{ is odd.} \end{cases}$$

These were all found by the same method. Dividing by $1 + x + x^2$, we were able to discover the alternating pattern of the coefficients. This told us approximately how

many coefficents would be nonzero. Trying different cases of n and $p \pmod 3$, we calculated the exact formulas.

2.4. **Case when** $p \equiv 1 \pmod 3$**.** For some values of p, $(1 + x + x^2)$ can be factored. In these cases, we can compute expressions for the coefficients of the polynomial.

Proposition 5. *The polynomial* $(1 + x + x^2)$ *is reducible in* F_p *if and only if* p *is equivalent to* $1 \pmod 3$.

Proof. The polynomial $1 + x + x^2$ is reducible in Z_p if and only if $1 + x + x^2 = (x - a)(x - b) = x^2 - (a + b)x + ab$.

This gives that $ab = 1$, so $b = a^{-1}$ in the multiplicative group Z_p. We know that $-a-b = 1$, so $a+a^{-1} = -1$. We also know that $1+a+a^2 = 0$, so $a^{-1} = a^2 \to a^3 = 1$. We can see if there exists an a such that $a^3 = 1$ then a is a root of $1+x+x^2$. Therefore, $1 + x + x^2$ is reducible in Z_p if and only if $a^3 = 1$ in Z_p for some a.

We know the multiplicative group Z_p is cyclic with order $p - 1$. If $m|p - 1$, then there exists an a such that $a^m = 1$. So if 3 divides $p - 1$, then there exists an a such that $a^3 = 1$.

Therefore, if $p \equiv 1 \bmod 3$, then $1 + x + x^2$ is reducible in Z_p and for all p such that $1 + x + x^2$ is reducible in Z_p, $p \equiv 1 \bmod 3$. □

For those primes, it is possible to find an expression for the coefficients. Let a be a root of the polynomial $(1+x+x^2)^n$. Then for $d<n$,

$$a_d = (-1)^d \sum_{k=0}^{d} \binom{n}{k}\binom{n}{d-k} a^{2d-k},$$

and for $d>n$,

$$a_d = (-1)^d \sum_{k=d-n}^{n} \binom{n}{k}\binom{n}{d-k} a^{2d-k}.$$

We find these expressions because in F_p, $(1+x+x^2)^n$ factors to $(x-a)^n(x-a^2)^n$. Then for every k between 0 and d, there are $\binom{n}{k}$ ways to choose k factors of a and $\binom{n}{d-k}$ ways to choose $d-k$ factors of a^2, giving a product of $a^k a^{2(d-k)} = a^{2d-k}$.

3. Approximation for $f_p(n)$

This section uses Amdeberhan and Stanley's theorem [1] about $N_\alpha(n)$, the number of coefficients of the n-th power of a polynomial $f(x)$ that are equal to α in F_p. They prove the existence of matrices matrices $A_0, \ldots, A_{p-1}$ of size $p^{deg(f)+1}$, row vector u and a column vector v (the former depending on α) such that if

$$\sum_{i=0}^{r} a_i p^i$$

is the base q expansion of n, then

$$N_\alpha(n) = uA_{a_0} \ldots A_{a_r} v.$$

For the $p = 2$ case Willson [3] came up with a matrix A and vectors u and v that compute the behavior of the sum of nonzero coefficients in $f(x)^n$. This matrix is strongly related to the matrices A_0 and A_1 cited before, however the size happens to be significantly smaller (less or equal to $p^{deg(f)} - 1$). In section 4 we will focus on the size of these matrices and explain the relations among them. This matrix led him to prove the conjecture stated in this section; we were unaware of his work when this research was done and the conjecture was made.

3.1. **Approximation.** As we state in the introduction, our main result concerns the behavior of nonzero coefficients for $f(x)^n$. We give a formula for r_n, which we define as

$$r_n = \sum_{m=0}^{n-1} f_p(m),$$

in the case $n = p^k$.

Theorem 6. *Let $f(x)$ be a polynomial in $F_p[x]$, and let*

$$r_n = \sum_{m=0}^{n-1} f_p(m)$$

be the sum of nonzero coefficients of $f(x)^n$. Then

$$r_{p^k} \underset{k\to\infty}{\sim} (p^k)^b$$

where

$$b = \log_p(\rho(A))$$

and $\rho(A)$ is an algebraic integer.

Proof. Let $m = \sum_{i=0}^{k} a_i p^i$ be the base p expansion of m. According to the theorem mentioned [1], the number of coefficients equal to α in $g(x)^n$, where g is a polynomial, is

$$u_\alpha A_{a_0} \dots A_{a_k} v.$$

In this expression, u_α is a row vector, v is a column vector, and all A are square matrices. Since $n = p^k$, m takes every value between 0 and $p^k - 1$, meaning that every possible combination of $k+1$ base p digits results.

For instance if $p = 2$,

$$\sum_{m=0}^{2^k-1} f(m) = u(A_0 + A_1)^k v,$$

because when $(A_0 + A_1)^k$ is expanded, every combination of matrices that correspond to digits results. If $p > 2$, in order to find the number of nonzero coefficients we must sum

$$u_\alpha (A_0 + \dots + A_{p-1})^k v$$

over all α, since each term represents the sum of coefficients that are equal to α. Summing over α gives

$$(u_1 + \dots + u_{p-1})(A_0 + \dots + A_{p-1})^{k-1} v.$$

The sum $A := (A_0 + \cdots + A_{p-1})$ is another square matrix; giving

$$r_{p^k} = (u_1 + \cdots + u_{p-1})A^k v.$$

The behavior of r_{p^k}, as k goes to ∞, will be determined by one eigenvalue of the matrix A. Since A is over the integers and $\rho(A)$ is an eigenvalue, $\rho(A)$ is an algebraic number. We have

$$r_{p^k} \underset{k\to\infty}{\sim} \rho(A)^k.$$

Taking $\log_{p^k}$ of both sides, we get

$$b = \log_{p^k}(r_{p^k}) = \log_{p^k}(\rho(A)^k) = \log_p(\rho(A)).$$

□

Based on this theorem, we have a conjecture for all p.

Conjecture 7. $\log_n(r_n)$ *approaches* b *as* n *goes to infinity.*

For $p = 2$, this appears to be the case. Computer simulations find that $\rho(A)$ is approximately 3.25 when n is large, while $1+\sqrt{5}$, an eigenvalue of the relevant matrix, is approximately 3.24.

3.2. **Example: $p = 2$.** For instance, we can look at the case where $p = 2$. The matrices in this case are 8×8. Calculation gives

$$A_0 = \begin{pmatrix} 2 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad A_1 = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$A = A_0 + A_1$: call its characteristic polynomial $d(x)$. Calculation shows that $\rho(A)$ is approximately $1 + \sqrt{5}$, which is a root of $d(x)$. In fact $d(x)$ contains the factor $x^2 - 2x - 4$, and $1 + \sqrt{5}$ is its largest root.

3.3. **Correlation between $f_p(n)$ and base p digits.** Computation of $f_p(n)$ is related to the base p expansion of n. This implies that there might be some correlation between $f_p(n)$ and the number of nonzero digits in the base p expansion of n.

Let $D(n)$ be the number of 1's in the binary expansion of n. Define

$$r_n(t) := \sum_{m=0}^{2^k-1} f_2(n) t^{D(n)}.$$

Proposition 8. $\log_k r_{2^k}(t)$ *approaches* $\log_2 \frac{t+\sqrt{9t^2+10t+1}+1}{2}$.

Proof. Fix a polynomial over F_2. Then Amdeberhan and Stanley's formula implies that

$$r_{2^k}(t) = u(A'_0 + tA'_1)^k v.$$

This formula allows us to find the correlation between $f_2(2^k)$ and $D(2^k)$ using eigenvalues of the matrix $A'_0+tA'_1$, which are algebraic functions of t. Computer calculation finds that the largest eigenvalue of the matrix is $\frac{t+\sqrt{9t^2+10t+1}+1}{2}$.

□

When $t = 1$, this expression is equal to $1+\sqrt{5}$, the eigenvalue found in the example above.

4. Size of the matrices and relation to representation theory

By looking at the previous example we wondered whether or not it is possible to find smaller matrices A'_0, A'_1 and vectors u' and v' that compute the number of nonzero coefficients of $(1 + x + x^2)^n$ in the same way the one provided by Amdeberhan and Stanley do. The fact that the behavior of r_n is determined by the root of a quadratic polynomial suggests that we might be able to find such smaller matrices. In this

particular example we actually found $u' = (1, 1, 2)$, $v' = (0, 1, 0)$ and matrices

$$A'_0 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad A'_1 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Note that in particular the characteristic polynomial of $A' := A'_0 + A'_1$ contains the same factor x^2-2x-4. As we mentioned in the introduction, Willson also gives a way of constructing a single matrix A, that determines the behavior of r_n. However what is interesting is that his construction provides a smaller matrix. In this particular case his matrix is already 3 by 3, and is conjugate to A'. We would like to understand how the degree of the minimum polynomial for $\rho(A)$ is related to the size of the matrices. This leads to interesting questions in representation theory.

4.1. **Relation to Representation Theory.** By looking at the polynomial $1+x+x^3$ over F_2, we have seen that the exponent b from theorem 6 is approximately 1.6942. This number is a root of a factor of the characteristic polynomial $d(x)$ of $A := A_0+A_1$ of degree 4 (call this factor $d'(x)$). For this polynomial, while Amdeberhan and Stanley's matrix A is 16×16, Willson's matrix (corresponding to the sum $A_0 + A_1$) is only 7×7. However, the degree of $d'(x)$ being 4 suggests the existence of some smaller matrix A.

In general, the characteristic polynomial of $A_0 + A_1$ is usually a product of many low degree factors with integer coefficients. This property can be explained by the

small size and many irreducible representations of very small dimension of the algebra generated by A_0 and A_1. This results in the degree of the algebraic integer $\rho(A)$ being much smaller than expected. It is not clear what happens for different polynomials and for different p. However, one may expect that representation theory can be used to gain a better understanding of the numbers $\rho(A)$ (and in particular to find their degrees), and to find exact formulas for $f_p(n)$.

5. Further Research

These results open up new paths for further study. Related to the results on formulae, we hope that, using findings for specific cases of n, the $p = 5$ case can be solved. Also, the expressions for coefficients that we found in reducible cases yield identities, and we would like to determine whether or not they are trivial.

Related to the approximation approach we took in the second part of the project, as we said in the previous section, we would like to understand the relation between the size of the matrices A_i, for $i = 0, \ldots, p-1$, and the representation theory of the algebra $\mathcal{A} =< A_i, i = 0, \ldots, p-1 >$. We would like to see how much the size of the matrices can be reduced.

Another question is whether or not there exist constants c_1 and c_2 such that

$$c_1 \leq \frac{r_n}{n^{\log_p \alpha}} \leq c_2,$$

where α is an algebraic integer. Finally, we find the correlation between $f_p(n)$ and base p digits for the polynomial $1 + x + x^2$. We hope that further study will find general results on this question for every polynomial.

References

[1] Amdeberhan, Tewodros, and Richard Stanley. "Polynomial Coefficient Enumeration." 2008.

[2] Stanley, Richard. *Enumerative Combinatorics.* 2010.

[3] Willson, Stephen J. "Computing Fractal Dimensions for Additive Cellular Automata." 1986.