Analogies between $\mathbb{Z}$ and $\mathbb{F}_q[t]$; elementary case studies

Paul Pollack

University of Georgia

July 2013

Integers vs. polynomials

Throughout, q denotes a prime power, and $\mathbb{F}_q$ denotes the finite field of order q (unique up to isomorphism).

The ring of integers $\mathbb{Z}$ and the ring of polynomials $\mathbb{F}_q[t]$ share a number of features. Both are:

- Euclidean domains (and so PIDs)
- Finite quotient domains (R/I is finite for nonzero I)
- Rings with only finitely many units.

Integers vs. polynomials

Throughout, q denotes a prime power, and $\mathbb{F}_q$ denotes the finite field of order q (unique up to isomorphism).

The ring of integers $\mathbb{Z}$ and the ring of polynomials $\mathbb{F}_q[t]$ share a number of features. Both are:

- Euclidean domains (and so PIDs)
- Finite quotient domains (R/I is finite for nonzero I)
- Rings with only finitely many units.

This means that much of the elementary theory carries over almost word-for-word — these parallels are stressed in many abstract algebra courses. Examples include unique factorization, Fermat's little theorem, and Wilson's theorem.

A brief dictionary

Integers	Polynomials
$\mathbb{Z}$, generic element n	$A = \mathbb{F}_q[t]$, generic element f
units: $\{\pm 1\}$	units: $\mathbb{F}_q^\times$
prime number	irreducible polynomial
positive integer	monic polynomial
absolute value	$\lvert f\rvert = q^{\deg f}$ (so $\lvert f\rvert = \lvert A/fA\rvert$)
dyadic interval $[x, 2x]$	polynomials of a given degree

Quadratic reciprocity

But the analogies run deeper than this. In this lecture, I want to dwell on a few of my favorite examples.

Recall that if p is an odd prime and $a \in \mathbb{Z}$, the Legendre symbol

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \mid a, \\ 1 & \text{if } a \equiv \square \pmod{p}, \\ -1 & \text{if } a \not\equiv \square \pmod{p}. \end{cases}$$

Theorem (Quadratic reciprocity law, Gauss)

For distinct odd primes p and q,

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

Quadratic reciprocity

What should quadratic reciprocity look like in $A = \mathbb{F}_q[t]$?

Suppose P is a monic irreducible element in $\mathbb{F}_q[t]$. Then A/P is a field of size $q^{\deg P}$. Hence, the nonzero squares form an index 2 subgroup of $(A/P)^\times$ whenever q is **odd**. So let's assume that.

Quadratic reciprocity

What should quadratic reciprocity look like in $A = \mathbb{F}_q[t]$?

Suppose P is a monic irreducible element in $\mathbb{F}_q[t]$. Then A/P is a field of size $q^{\deg P}$. Hence, the nonzero squares form an index 2 subgroup of $(A/P)^\times$ whenever q is **odd**. So let's assume that.

We can again define a Legendre symbol. If $f \in A$, set

$$\left(\frac{f}{P}\right) = \begin{cases} 0 & \text{if } P \mid f, \\ 1 & \text{if } f \equiv \square \pmod{P}, \\ -1 & \text{if } f \not\equiv \square \pmod{P}. \end{cases}$$

This is multiplicative in the top entry and "periodic" modulo P, in analogy with the usual Legendre symbol.

Quadratic reciprocity

Example

Let $q = 3$, so that $A = \mathbb{F}_3[t]$. Let $P = t^2 + 1 \in A$. Then A/P is the field with 3^2 elements, and so the unit group of A/P is the cyclic group of order 8. By direct computation, the $8 = \frac{1}{2} \cdot 4$ squares in $(A/P)^\times$ are represented by

$$1, \quad -1, \quad t, \quad 2t.$$

Continuing, suppose $Q = t^3 - t + 1$. Then $Q \equiv t + 1 \pmod{P}$, and so

$$\left(\frac{Q}{P}\right) = -1.$$

Quadratic reciprocity

Suppose P and Q are distinct monic irreducibles in A. Then the most naive guess for a quadratic reciprocity law would be

$$\left(\frac{P}{Q}\right)\left(\frac{Q}{P}\right) = (-1)^{\frac{|P|-1}{2}\frac{|Q|-1}{2}}.$$

Quadratic reciprocity

Suppose P and Q are distinct monic irreducibles in A. Then the most naive guess for a quadratic reciprocity law would be

$$\left(\frac{P}{Q}\right)\left(\frac{Q}{P}\right) = (-1)^{\frac{|P|-1}{2}\frac{|Q|-1}{2}}.$$

Theorem (Dedekind, 1857)

This is correct!

The proof of our theorem can be established completely analogously to Gausss fifth proof [of QR] and is based on [Gauss's lemma] ...its consequences, up to ...the proof of the theorem, are so similar to the ones in the cited treatise of Gauss that **no one can fail to find the complete proof**.

A short proof of quadratic reciprocity in $A = \mathbb{F}_q[t]$

We will prove quadratic reciprocity where the exponent on -1 looks a bit different. Of course, we only care about this exponent modulo 2.

Say P has degree d and Q has degree e. Then modulo 2,

$$\frac{|P|-1}{2} = \frac{q^d-1}{2} = \frac{q-1}{2}(1+q+q^2+\cdots+q^{d-1}) \equiv d\frac{q-1}{2}.$$

Similarly, $\frac{|Q|-1}{2} \equiv e\frac{q-1}{2}$. Thus,

$$\frac{|P|-1}{2}\frac{|Q|-1}{2} \equiv de\frac{q-1}{2} \pmod 2.$$

A short proof of quadratic reciprocity in $A = \mathbb{F}_q[t]$, ctd.

So we can replace the exponent of -1 with $de\frac{q-1}{2}$; this leads to the form of QR that we will actually prove:

Theorem

Let P and Q be distinct monic irreducibles in $A = \mathbb{F}_q[t]$, where q is odd. Say $\deg P = d$ and $\deg Q = e$. Then

$$\left(\frac{P}{Q}\right)\left(\frac{Q}{P}\right) = (-1)^{de\frac{q-1}{2}}.$$

The argument we will give is due essentially to F. K. Schmidt, with some fine tuning by L. Carlitz.

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Lemma

Let P be a monic irreducible in A. For every $f \in A$, we have

$$\left(\frac{f}{P}\right) \equiv f^{\frac{|P|-1}{2}} \pmod{P}.$$

This is clear if $f \equiv 0 \pmod{P}$, so suppose otherwise.

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Lemma

Let P be a monic irreducible in A. For every $f \in A$, we have

$$\left(\frac{f}{P}\right) \equiv f^{\frac{|P|-1}{2}} \pmod{P}.$$

This is clear if $f \equiv 0 \pmod{P}$, so suppose otherwise. If $f \equiv g^2 \pmod{P}$, then $f^{\frac{|P|-1}{2}} \equiv g^{|P|-1} \equiv 1 \equiv \left(\frac{f}{P}\right) \pmod{P}$.

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Lemma

Let P be a monic irreducible in A. For every $f \in A$, we have

$$\left(\frac{f}{P}\right) \equiv f^{\frac{|P|-1}{2}} \pmod{P}.$$

This is clear if $f \equiv 0 \pmod{P}$, so suppose otherwise. If $f \equiv g^2 \pmod{P}$, then $f^{\frac{|P|-1}{2}} \equiv g^{|P|-1} \equiv 1 \equiv \left(\frac{f}{P}\right) \pmod{P}$. Finally, suppose $f \not\equiv \square \pmod{P}$. Now there are at most $\frac{|P|-1}{2}$ solutions mod P to $X^{\frac{|P|-1}{2}} \equiv 1 \pmod{P}$, since A/P is a field. There are also $\frac{|P|-1}{2}$ squares mod P. So $f^{\frac{|P|-1}{2}} \not\equiv 1 \pmod{P}$. But $(f^{\frac{|P|-1}{2}})^2 \equiv f^{|P|-1} \equiv 1 \pmod{P}$, forcing $f^{\frac{|P|-1}{2}} \equiv -1 \equiv \left(\frac{f}{P}\right) \pmod{P}$.

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Idea of the proof: Find explicit expressions for $\left(\frac{P}{Q}\right)$ and $\left(\frac{Q}{P}\right)$ in terms of the roots of P and Q and then compare.

Let $\mathbb{F}$ stand for the algebraic closure of $\mathbb{F}_q$. Both P and Q split into distinct linear factors over $\mathbb{F}$, and we can write

$$P(t) = (t - \alpha)(t - \alpha^q) \cdots (t - \alpha^{q^{d-1}})$$

and

$$Q(t) = (t - \beta)(t - \beta^q) \cdots (t - \beta^{q^{e-1}}).$$

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Idea of the proof: Find explicit expressions for $\left(\frac{P}{Q}\right)$ and $\left(\frac{Q}{P}\right)$ in terms of the roots of P and Q and then compare.

Let $\mathbb{F}$ stand for the algebraic closure of $\mathbb{F}_q$. Both P and Q split into distinct linear factors over $\mathbb{F}$, and we can write

$$P(t) = (t-\alpha)(t-\alpha^q)\cdots(t-\alpha^{q^{d-1}})$$

and

$$Q(t) = (t-\beta)(t-\beta^q)\cdots(t-\beta^{q^{e-1}}).$$

We would like to evaluate $P^{\frac{|Q|-1}{2}}$ mod Q, since this gives $\left(\frac{P}{Q}\right)$.

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Idea of the proof: Find explicit expressions for $\left(\frac{P}{Q}\right)$ and $\left(\frac{Q}{P}\right)$ in terms of the roots of P and Q and then compare.

Let $\mathbb{F}$ stand for the algebraic closure of $\mathbb{F}_q$. Both P and Q split into distinct linear factors over $\mathbb{F}$, and we can write

$$P(t) = (t-\alpha)(t-\alpha^q)\cdots(t-\alpha^{q^{d-1}})$$

and

$$Q(t) = (t-\beta)(t-\beta^q)\cdots(t-\beta^{q^{e-1}}).$$

We would like to evaluate $P^{\frac{|Q|-1}{2}}$ mod Q, since this gives $\left(\frac{P}{Q}\right)$. We compute $P^{\frac{|Q|-1}{2}}$ mod $t-\beta^{q^i}$ for each i, starting with $P^{\frac{|Q|-1}{2}}$ mod $t-\beta$ (the case $i=0$).

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Using that P has coefficients belonging to $\mathbb{F}_q$, we see that

$$P(t)^{\frac{|Q|-1}{2}} = P(t)^{\frac{q^e-1}{2}} = P(t)^{(1+q+\cdots+q^{e-1})\frac{q-1}{2}}$$
$$= (P(t)P(t^q)\cdots P(t^{q^{e-1}}))^{\frac{q-1}{2}}.$$

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

Using that P has coefficients belonging to $\mathbb{F}_q$, we see that

$$P(t)^{\frac{|Q|-1}{2}} = P(t)^{\frac{q^e-1}{2}} = P(t)^{(1+q+\cdots+q^{e-1})\frac{q-1}{2}}$$
$$= (P(t)P(t^q)\cdots P(t^{q^{e-1}}))^{\frac{q-1}{2}}.$$

Modulo $t-\beta$, this is congruent to

$$(P(\beta)P(\beta^q)\cdots P(\beta^{q^{e-1}}))^{\frac{q-1}{2}}.$$

Remembering that $P(t) = \prod_{i=0}^{d-1}(t-\alpha^{q^i})$, we get

$$P(t)^{\frac{|Q|-1}{2}} \equiv \left(\prod_{j=0}^{e-1}\prod_{i=0}^{d-1}(\beta^{q^j}-\alpha^{q^i})\right)^{\frac{q-1}{2}} \pmod{t-\beta}.$$

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

OK, so we have now that

$$P(t)^{\frac{|Q|-1}{2}} \equiv \left(\prod_{j=0}^{e-1} \prod_{i=0}^{d-1} (\beta^{q^j} - \alpha^{q^i}) \right)^{\frac{q-1}{2}} \pmod{t - \beta}.$$

How does the right hand side change if we replace the modulus $t - \beta$ with $t - \beta^{q^\ell}$?

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

OK, so we have now that

$$P(t)^{\frac{|Q|-1}{2}} \equiv \left(\prod_{j=0}^{e-1}\prod_{i=0}^{d-1}(\beta^{q^j} - \alpha^{q^i})\right)^{\frac{q-1}{2}} \pmod{t-\beta}.$$

How does the right hand side change if we replace the modulus $t-\beta$ with $t-\beta^{q^\ell}$? It doesn't! Hence,

$$\prod_{j=0}^{e-1}\left(\prod_{i=0}^{d-1}(\beta^{q^j} - \alpha^{q^i})\right)^{\frac{q-1}{2}} \equiv P(t)^{\frac{|Q|-1}{2}} \equiv \left(\frac{P}{Q}\right) \pmod{Q(t)}.$$

Both sides are constants (elements of $\mathbb{F}$); this implies

$$\left(\frac{P}{Q}\right) = \left(\prod_{j=0}^{e-1}\prod_{i=0}^{d-1}(\beta^{q^j} - \alpha^{q^i})\right)^{\frac{q-1}{2}}.$$

A short proof of QR in $A = \mathbb{F}_q[t]$, ctd.

So in $\mathbb{F}$, we have the equation

$$\left(\frac{P}{Q}\right) = \left(\prod_{j=0}^{e-1}\prod_{i=0}^{d-1}(\beta^{q^j} - \alpha^{q^i})\right)^{\frac{q-1}{2}}.$$

Similarly: $\left(\frac{Q}{P}\right) = \left(\prod_{j=0}^{e-1}\prod_{i=0}^{d-1}(\alpha^{q^i} - \beta^{q^j})\right)^{\frac{q-1}{2}}$. Thus,

$$\left(\frac{P}{Q}\right) = (-1)^{de\frac{q-1}{2}}\left(\frac{Q}{P}\right), \quad \text{whence} \quad \left(\frac{P}{Q}\right)\left(\frac{Q}{P}\right) = (-1)^{de\frac{q-1}{2}}.$$

The final identity is true not only in $\mathbb{F}$ but also in $\mathbb{Z}$, since both sides are ± 1. Done!

Fermat's last theorem

Perhaps the most celebrated mathematical success story in recent memory is the resolution of the following longstanding conjecture of Fermat.

Theorem (Wiles and Taylor, 1995)

Let $n > 3$. Then there are no integer solutions to

$$x^n + y^n = z^n$$

with $xyz \neq 0$.

Fermat's last theorem, ctd.

One could formulate the exact same conjecture for polynomials.

Conjecture

If $n > 3$, there are no solutions to $x^n + y^n = z^n$ with $x, y, z \in A = \mathbb{F}_q[t]$ and $xyz \neq 0$.

But this is **false**. For example, there might well be constant solutions. Even worse, whenever $x + y = z$ in A, then $x^{p^k} + y^{p^k} = z^{p^k}$, where $p = \mathrm{char}(F)$.

Fermat's last theorem, ctd.

One could formulate the exact same conjecture for polynomials.

Conjecture

If $n > 3$, there are no solutions to $x^n + y^n = z^n$ with $x, y, z \in A = \mathbb{F}_q[t]$ and $xyz \neq 0$.

But this is **false**. For example, there might well be constant solutions. Even worse, whenever $x + y = z$ in A, then $x^{p^k} + y^{p^k} = z^{p^k}$, where $p = \mathrm{char}(F)$.

Conjecture (modified)

If $n \geq 3$ and $p \nmid n$, then there are no coprime solutions to $x^n + y^n = z^n$ with $x, y, z \in A = \mathbb{F}_q[t]$, $xyz \neq 0$, and x, y, z not all constant.

Fermat's last theorem, ctd.

Conjecture (modified)

If $n \geq 3$ and $p \nmid n$, then there are no coprime solutions to $x^n + y^n = z^n$ with $x, y, z \in A = \mathbb{F}_q[t]$, $xyz \neq 0$, and x, y, z not all constant.

Theorem (Liouville – Korkine – Greenleaf)

The modified conjecture is true!

There are various ways to prove this. Perhaps the simplest proof uses Mason's theorem.

Mason's theorem

For a polynomial f over a field F, let $R(f)$ be the product of the distinct monic irreducibles dividing f (the **squarefree part** of f), and let $r(f) = \deg R(f)$.

Theorem (Mason, 1984)

Let F be any field. Suppose $f, g, h \in F[t]$ are nonzero and that that there is no irreducible dividing all of f, g, and h. Suppose that $f + g = h$ and that it is **not** *the case that $f' = g' = h' = 0$. Then*

$$\max\{\deg f, \deg g, \deg h\} \leq r(fgh) - 1.$$

Deduction of FLT for $\mathbb{F}_q[t]$

Theorem (Mason, 1984)

Let F be any field. Suppose $f, g, h \in F[t]$ are nonzero and that that there is no irreducible dividing all of f, g, and h. Suppose that $f + g = h$ and that it is **not** *the case that $f' = g' = h' = 0$. Then*

$$\max\{\deg f, \deg g, \deg h\} \leq r(fgh) - 1.$$

Now we return to Fermat's last theorem for polynomials. Suppose $x^n + y^n = z^n$ with x, y, z nonzero elements of $\mathbb{F}_q[t]$, coprime, not all constant.
Suppose also that $p \nmid n$. We have to show $n < 3$.

We can assume x, y, and z are not all polynomials in t^p; otherwise, take pth roots of the equation $x^n + y^n = z^n$ (repeat as necessary).

Deduction of FLT, continued

Now $f = x^n$, $g = y^n$, and $h = z^n$ satisfy the relation $f + g = h$.

Moreover, not all of $f', g', h' = 0$, since $p \nmid n$ and not all of x, y, and z are polynomials in t^p.

So Mason's theorem applies and shows that

$$\begin{aligned} n \max\{\deg x, \deg y, \deg z\} &\leq r(x^n y^n z^n) - 1 \\ &= r(xyz) - 1 \\ &< \deg{(xyz)} \\ &\leq 3 \max\{\deg x, \deg y, \deg z\}. \end{aligned}$$

Hence, $n < 3$.

Proof of Mason's theorem

We give a proof due to Noah Snyder (1999).

Recall that $R(f)$ denotes the product of the distinct monic irreducibles dividing f and that $r(f) = \deg R(f)$.

Lemma

Let f be a nonzero polynomial in $F[t]$. Then

$$f/R(f) \mid \gcd(f, f').$$

Proof of Mason's theorem, ctd.

Lemma

Let f be a nonzero polynomial in $F[t]$. Then

$$f/R(f) \mid \gcd(f, f').$$

Exercise: Check that $R(f)$ and $\gcd(f, f')$ do not change under extensions of F.

Proof of Mason's theorem, ctd.

Lemma

Let f be a nonzero polynomial in $F[t]$. Then

$$f/R(f) \mid \gcd(f, f').$$

Exercise: Check that $R(f)$ and $\gcd(f, f')$ do not change under extensions of F.

Hence, we can assume F is algebraically closed. Write $f = c\prod(t-\alpha_i)^{e_i}$. By the product rule, $(t-\alpha_i)^{e_i-1} \mid f'$ for each i, and hence

$$\prod(t-\alpha_i)^{e_i-1} \mid \gcd(f, f').$$

The left hand side is $f/R(f)$.

Proof of Mason's theorem, ctd.

Using $f + g = h$, one checks $h'g - g'h = fg' - f'g$.

This common element is divisible by $\gcd(f, f')$, $\gcd(g, g')$, and $\gcd(h, h')$. Thus, it is divisible by the (coprime!) elements $f/R(f), g/R(g)$, and $h/R(h)$.

Hence, $h'g - g'h$ is divisible by

$$fgh/(R(f)R(g)R(h)) = fgh/R(fgh).$$

Proof of Mason's theorem, ctd.

We want to show all of $\deg f, \deg h, \deg h$ are smaller than $r(fgh) = \deg R(fgh)$.

Assume for the sake of contradiction that $\deg f \geq \deg R(fgh)$. Then

$$\begin{aligned}\deg(fgh/R(fgh)) &= \deg gh + (\deg f - \deg R(fgh)) \\ &\geq \deg (gh) \\ &> \deg (h'g - g'h).\end{aligned}$$

Since $fgh/R(fgh) \mid h'g - g'h$, these inequalities imply that $h'g - g'h = 0$. But then $h \mid h'$, so $h' = 0$. Since $h'g - g'h = 0$, we get $g' = 0$. Since $f = h - g$, we get $f' = h' - g' = 0$.

Proof of Mason's theorem, ctd.

So all of $f', g', h' = 0$. But we assumed that this was not the case! So the only possibility left is that

$$\deg f \leq \deg R(fgh) - 1 = r(fgh) - 1.$$

But f and g play symmetric roles, since $f + g = h$. So the same bound on the degree holds for $\deg g$.

Finally, since $f + g = h$, we conclude that the same bound holds for $\deg h$.

This completes the proof of Mason's theorem (and so also of FLT).

abc?

As we have just seen, Mason's theorem allows one to give a very short proof of Fermat's last theorem for polynomials.

There is an analogous conjecture for integers, known as the *abc*-conjecture.

Conjecture (Oesterlé–Masser)

For every $\epsilon > 0$, there are only finitely many triples of coprime positive integers a, b, c, satisfying $a + b = c$ and having

$$c > \Big(\prod_{p \mid abc} p\Big)^{1+\epsilon}.$$

Quite recently, Mochizuki has claimed a proof. This would have many important arithmetic consequences.

Sums of two squares

Fermat knew that every prime $p \equiv 1 \pmod 4$ was a sum of two squares. A complete characterization of which integers are sums of two squares is attributed to Euler.

Theorem

The positive integer n is a sum of two squares if and only if every prime $p \equiv 3 \pmod 4$ shows up to an even exponent (possibly zero) in the prime factorization of n.

Sums of two squares, ctd.

OK, which elements of $A = \mathbb{F}_q[t]$ are sums of two squares?

When q is even — i.e., $p = 2$ — then everything that is a sum of two squares is a square itself. So let's assume that q is odd.

A natural guess, after Euler's result, might be the following.

Conjecture

Let $f \in A$. Then f can be written as a sum of two squares in A if and only if every prime P with $|P| \equiv 3 \pmod 4$ shows up to an even exponent in the prime factorization of A.

Sums of two squares, ctd.

OK, which elements of $A = \mathbb{F}_q[t]$ are sums of two squares?

When q is even — i.e., $p = 2$ — then everything that is a sum of two squares is a square itself. So let's assume that q is odd.

A natural guess, after Euler's result, might be the following.

Conjecture

Let $f \in A$. Then f can be written as a sum of two squares in A if and only if every prime P with $|P| \equiv 3 \pmod 4$ shows up to an even exponent in the prime factorization of A.

Theorem (Leahey, 1967)

This is true!

Sums of two squares, ctd.

A more general theorem was proved by Joly (1970).

Theorem

Let F be a field of characteristic $\neq 2$. Suppose that -1 is not a square in F, but that every element of F is a sum of two squares. Then the following are equivalent:

1. *f is a sum of two squares,*
2. *if P is an irreducible dividing f for which -1 is not a square in $F[t]/(P)$, then P appears to an even power in the prime factorization of f.*

For the proofs, Leahey and Joly use the arithmetic of $F[t][i] = F[i][t]$. This is analogous to studying sums of two squares as norms from $\mathbb{Z}[i]$.

Higher powers

Instead of considering sums of squares, let's consider sums of kth powers.

For each k, let $\Sigma(k, \mathbb{Z})$ be the set of integers that can be written as a finite sum of kth powers of elements of $\mathbb{Z}$.

It is easy to see that if k is odd, then $\Sigma(k, \mathbb{Z}) = \mathbb{Z}$, while when k is even, $\Sigma(k, \mathbb{Z}) = \mathbb{Z}_{\geq 0}$. The following conjecture was made by Edward Waring (1770).

Conjecture

Every element of $\Sigma(k, \mathbb{Z})$ can be written as the sum of at most $w(k, \mathbb{Z})$ kth powers, where $w(k, \mathbb{Z}) < \infty$.

For example, Lagrange's theorem shows that $w(2, \mathbb{Z}) = 4$ is acceptable.

Waring's problem

As another example, notice that

$$(t+1)^3 + (-t)^3 + (-t)^3 + (t-1)^3 = 6t.$$

So every multiple of 6 is a sum of four cubes in $\mathbb{Z}$. Since $n - n^3$ is always a multiple of 6, we see that $w(3, \mathbb{Z}) = 5$ is admissible.

Waring's problem

As another example, notice that

$$(t+1)^3 + (-t)^3 + (-t)^3 + (t-1)^3 = 6t.$$

So every multiple of 6 is a sum of four cubes in $\mathbb{Z}$. Since $n - n^3$ is always a multiple of 6, we see that $w(3, \mathbb{Z}) = 5$ is admissible.

The first proof of the existence of a finite $w(k, \mathbb{Z})$ for every k is due to Hilbert.

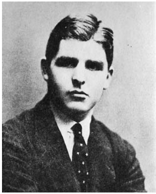
Theorem (Hilbert, 1909)

Waring was right!

All known proofs of this theorem are fairly intricate.

Paley's theorem

If R is a ring (always understood to be commutative, with 1), we let $\Sigma(k, R)$ be the set of elements of R that have an expression as a finite sum of kth powers.

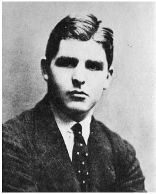


Theorem (Paley, 1932)

Let $A = \mathbb{F}_q[t]$. Then every element of $\Sigma(k, A)$ can be written as a sum of at most $w(k, A)$ kth powers, where $w(k, A) < \infty$.

Paley's theorem

If R is a ring (always understood to be commutative, with 1), we let $\Sigma(k, R)$ be the set of elements of R that have an expression as a finite sum of kth powers.



Theorem (Paley, 1932)

Let $A = \mathbb{F}_q[t]$. Then every element of $\Sigma(k, A)$ can be written as a sum of at most $w(k, A)$ kth powers, where $w(k, A) < \infty$.

In fact, we will show that $w(k, A)$ can be chosen to depend **only** on k (and **not** on q). Rather than follow Paley, we give an argument using methods of Vaserstein (1987).

Preliminaries: two results of Tornheim (1938)

Analogies between $\mathbb{Z}$ and $\mathbb{F}_q[t]$

Paul Pollack

A dictionary

Reciprocity

Fermat's last theorem

Mason's theorem

Sums of two squares

Waring's problem

Theorem

Let F be a field of positive characteristic. Then $\Sigma(k, F)$ is a subfield of F.

Proof: By definition, $\Sigma(k, F)$ is closed under $+$. It is also closed under $\cdot$, since

$$(\sum_i \alpha_i^k)(\sum_j \beta_j^k) = \sum_{i,j} (\alpha_i \beta_j)^k.$$

It is closed under taking additive inverses, since (e.g.)

$$-\sum_i \alpha_i^k = \underbrace{\left(\sum_i \alpha_i^k + \cdots + \sum_i \alpha_i^k\right)}_{p-1 \text{ times}}.$$

First theorem, ctd.

Finally, it is closed under taking multiplicative inverses: Suppose $0 \neq \alpha \in \Sigma(k, F)$. Then $\alpha^{-k} \in F^K \subset \Sigma(k, F)$, and $\alpha^{k-1} \in \Sigma(k, F)$ (since we already proved closure under multiplication).

Thus, using closure under $\cdot$ once again,

$$\alpha^{-1} = \alpha^{-k}\alpha^{k-1} \in \Sigma(k, F).$$

Preliminaries: two results of Tornheim (1938)

Theorem

Let $F = \mathbb{F}_q$ be a finite field. Then every element of $\Sigma(k, F)$ is expressible as a sum of k kth powers in F.

It is helpful to introduce some notation from additive number theory. If B and C are subsets of an additive group, we let

$$B \oplus C = \{b + c : b \in B, c \in C\}.$$

We define the ℓ-fold sumset of B to be

$$\ell B = \underbrace{B \oplus B \oplus \cdots \oplus B}_{\ell \text{ times}}.$$

Now let B be the set of kth powers in the field $F = \mathbb{F}_q$.

Preliminaries: two results of Tornheim (1938)

Since $0 \in B$, we have a sequence of inclusions

$$0B = \{0\} \subset B \subset 2B \subset 3B \subset \dots.$$

We look for the first positive integer i for which $(i+1)B = iB$. In that case,

$$(i+2)B = (i+1)B + B = iB + B = (i+1)B,$$

and so the sequence of sumsets stabilizes:

$$iB = (i+1)B = (i+2)B = \cdots = \Sigma(k, F).$$

Key observation: $(i+1)B \setminus iB$ is stable under multiplication by $(F^{\times})^k$.

Preliminaries: two results of Tornheim (1938)

Consequently, whenever $(i+1)B$ properly contains B, the set-difference $(i+1)B \setminus iB$ is a union of cosets of $(F^\times)^k$. The total number of cosets of $(F^\times)^k$ in $F^\times$ is

$$\gcd(q-1,k) \leq k.$$

Consequently. there can be at most $\gcd(q-1,k) \leq k$ strict inclusions in the sequence

$$\{0\} = 0A \subset A \subset 2A \subset 3A \subset \dots.$$

Thus, every element of $\Sigma(k,F)$ is a sum of at most $\gcd(q-1,k) \leq k$ kth powers.

Since 0 is a kth power, we can use **exactly** k such powers in the representation, if we wish.

Back to Waring's problem for polynomials over finite fields

Recall that our goal is to prove the following theorem.

Theorem

Let $A = \mathbb{F}_q[t]$. Then every element of $\Sigma(k, A)$ can be written as a sum of at most $w(k, A)$ kth powers, where $w(k, A) < \infty$ can be chosen to depend only on k.

First, we show that we can assume $p \nmid k$. Suppose that the theorem is proved under this extra assumption.

Say $k = p^e k'$, where $p \nmid k'$. If $f \in \Sigma(k, A)$, then

$$f = \sum f_i^k = \left(\sum f_i^{k/p^e}\right)^{p^e}.$$

Let

$$g = \sum f_i^{k/p^e} \in \Sigma(k/p^e, A).$$

Reduction to the case when $p \nmid k$

We have $f = g^{p^e}$, where

$$g = \sum f_i^{k/p^e} \in \Sigma(k/p^e, A).$$

Since $p \nmid \frac{k}{p^e}$, we know that every element of $\Sigma(k/p^e, A)$ is a sum of $w(k/p^e, A)$ (k/p^e)th powers.

In particular, g is a sum of $w(k/p^e, A)$ (k/p^e)th powers. Thus, $f = g^{p^e}$ is a sum of $w(k/p^e, A)$ kth powers.

So the theorem follows with

$$w(\mathbb{F}_q[t], k) = w(\mathbb{F}_q[t], k/p^e).$$

Waring's problem for polynomials

Theorem

Let $A = \mathbb{F}_q[t]$. Then every element of $\Sigma(k, A)$ can be written as a sum of at most $w(k, A)$ kth powers, where $w(k, A) < \infty$ can be chosen to depend only on k.

Case 1: $p > k$. In this case, we show that $\Sigma(k, A) = A$ and that one can take $w(k, A) = k^2$. Choose distinct elements $\alpha_1, \ldots, \alpha_k$ of $\mathbb{F}_p$. Consider the $k \times k$ Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \cdots & \alpha_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \cdots & \alpha_k^{k-1} \end{pmatrix}.$$

Waring's problem for polynomials

Since the matrix is invertible, we can solve the system

$$\sum_{i=1}^{k} \beta_i \alpha_i^s = \begin{cases} 0 & \text{if } s = 0, 1, 2, \dots, k-2, \\ k^{-1} & \text{if } s = k-1 \end{cases}$$

for $\beta_1, \dots, \beta_k \in \mathbb{F}_p$.

It follows that in $\mathbb{F}_p[y]$,

$$\sum_{i=1}^{k} \beta_i (y + \alpha_i)^k = y + \gamma, \quad \text{where} \quad \gamma = \sum_{i=1}^{k} \beta_i \alpha_i^k \in \mathbb{F}_p.$$

Thus,

$$\sum_{i=1}^{k} \beta_i (y + (\alpha_i - \gamma))^k = y.$$

Waring's problem for polynomials

We have (for constants α_i, β_i, γ all in $\mathbb{F}_p$)

$$\sum_{i=1}^{k} \beta_i (y + (\alpha_i - \gamma))^k = y.$$

We can expand each β_i as a sum of k kth powers in $\mathbb{F}_p$. This gives y as a sum of k^2 kth powers in $\mathbb{F}_p[y]$.

Replacing y with an arbitrary element f of $A = \mathbb{F}_q[t]$, we get that every $f \in A$ is a sum of k^2 kth powers in $\mathbb{F}_p[f] \subset \mathbb{F}_q[t]$. This completes the proof of Case 1.

Waring's problem for polynomials

Case 2: $p \leq k$

We observe that the argument given for Case 1 in fact proves the following result (with $F = \mathbb{F}_p$).

Lemma

Let F be a field of characteristic coprime to k and where F has more than k elements. Then y can be written in the form

$$\sum \beta_i \ell_i(y)^k,$$

where each $\beta_i \in F$ and each $\ell_i(y)$ is a (linear) polynomial with coefficients from F.

Waring's problem for polynomials, case 2

Lemma

Let F be a field of characteristic coprime to k and where F has more than k elements. Then y can be written in the form

$$\sum \beta_i \ell_i(y)^k,$$

where each $\beta_i \in F$ and each $\ell_i(y)$ is a (linear) polynomial with coefficients from F.

We choose $F = \Sigma(k, \mathbb{F}_p(t))$. Using that each $\beta_i \in \Sigma(k, \mathbb{F}_p(t))$, we obtain that y is a finite sum of kth powers in $\mathbb{F}_p(t)[y]$.

Now we clear denominators. Multiplying by $D(t)^k \in \mathbb{F}_p[t]$ for a suitable $D(t)$, we get an identity

$$M(t)y = (\text{finite sum of } k\text{th powers in } \mathbb{F}_p[t][y]),$$

where $M(t) = D(t)^k$.

Waring's problem for polynomials, case 2

We get an identity in $\mathbb{F}_p[t][y]$:

$$M(t)y = (\text{finite sum of } k\text{th powers in } \mathbb{F}_p[t][y]).$$

We can now characterize $\Sigma(k, A)$, where $A = \mathbb{F}_q[t]$.

Lemma

An element $f \in A$ is a sum of kth powers in A if and only if its reduction mod M is a sum of kth powers in $A/(M)$.

If f is a sum of kth powers, then it is a sum of kth powers mod M. In the other direction, if $f \equiv f_1^k + \cdots + f_s^k \pmod{M}$, then

$$f(t) - (f_1(t)^k + \cdots + f_s(t)^k) = M(t)q(t)$$

for some $q(t) \in \mathbb{F}_q[t]$. Plug $y = q(t)$ into our identity above.

Waring's problem for polynomials, case 2

We still have to show that an $f \in \Sigma(k, A)$ is a sum of $O_k(1)$ kth powers in A.

So suppose $f \in \Sigma(k, A)$. We have just seen that to write f as a sum of kth powers, it suffices to first write f mod M as a sum of kth powers in $A/(M)$, say

$$f \equiv f_1^k + \cdots + f_s^k \pmod{M},$$

and then apply the identity

$$M(t)y = (\text{finite sum of } k\text{th powers in } \mathbb{F}_p[t][y]).$$

to write $f - (f_1^k + \cdots + f_s^k)$ as a sum of kth powers. The identity depends only on p and k, and since $p \leq k$, the number of terms in the identity is bounded solely in terms of k.

Waring's problem for polynomials, case 2

So it remains only to show we can always choose $s = O_k(1)$.

In other words, we have reduced the proof of the theorem to the following lemma.

Lemma

Let $A = \mathbb{F}_q[t]$, and let M be a nonzero element of A. Then every element of $\Sigma(k, A/(M))$ can be written as a sum of at most $w(k, A/(M))$ kth powers, where $w(k, A/(M))$ is bounded solely in terms of k.

In fact, we will show that we can take $w(k, A/(M)) = k + 1$.

By the Chinese remainder theorem, it suffices to prove this stronger claim when M is a power of an irreducible polynomial, say $M = P^e$.

Waring's problem for polynomials, case 2

So suppose that $f \bmod P^e$ is a sum of kth powers modulo P^e. Then $f \bmod P$ is a sum of kth powers mod P.

Since $\Sigma(k, A/(P))$ is a field, it is also true that $f - 1 \bmod P$ is a sum of kth powers mod P.

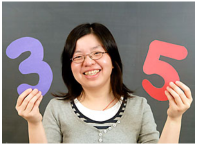
Since $A/(P)$ is a finite field, Tornheim says we only need k kth powers: We can write $f - 1 \equiv f_1^k + \cdots + f_k^k \pmod{P}$. Thus,

$$f - (f_1^k + \ldots f_k^k) \equiv 1 \pmod{P}.$$

Using once more that $p \nmid k$, Hensel's lemma implies that $f - (f_1^k + \ldots f_k^k) \equiv f_{k+1}^k \pmod{P^e}$. Hence,

$$f \equiv f_1^k + f_2^k + \cdots + f_{k+1}^k \pmod{P^e}.$$

The state of the art on Waring for polynomials




Liu and Wooley (2007) have shown that one can take

$$w(k, \mathbb{F}_q[t]) \leq (1 + o(1))k \log k$$

as $k \to \infty$, uniformly in q.