Today.

Polynomials.

Secret Sharing.

Correcting for loss or even corruption.

Secret Sharing.

Share secret among n people.

Secrecy: Any $k-1$ knows nothing.
Roubustness: Any k knows secret.
Efficient: minimize storage.

The idea of the day.

Two points make a line.
Lots of lines go through one point.

Polynomials

A **polynomial**

$$P(x) = a_d x^d + a_{d-1}x^{d-1} \cdots + a_0.$$

is specified by **coefficients** $a_d, \ldots a_0$.

$P(x)$ **contains** point (a,b) if $b = P(a)$.

Polynomials over reals: $a_1, \ldots, a_d \in \Re$, use $x \in \Re$.

Polynomials $P(x)$ with arithmetic modulo p: [1] $a_i \in \{0, \ldots, p-1\}$ and

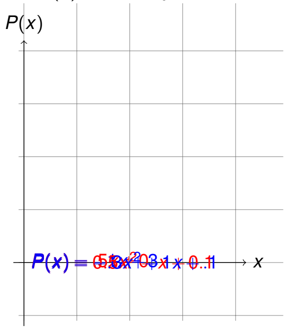
$$P(x) = a_d x^d + a_{d-1}x^{d-1} \cdots + a_0 \pmod p,$$

for $x \in \{0, \ldots, p-1\}$.

[1] A field is a set of elements with addition and multiplication operations, with inverses. $GF(p) = (\{0, \ldots, p-1\}, + \pmod p, * \pmod p)$.

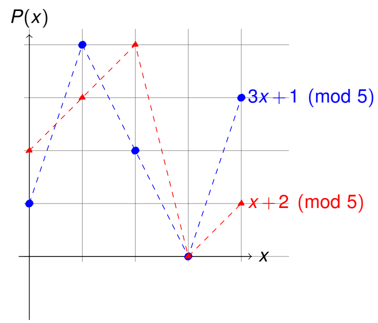
Polynomial: $P(x) = a_d x^4 + \cdots + a_0$

Line: $P(x) = a_1 x + a_0 = mx + b$



Parabola: $P(x) = a_2 x^2 + a_1 x + a_0 = ax^2 + bx + c$

Polynomial: $P(x) = a_d x^4 + \cdots + a_0 \pmod p$



Finding an intersection.
$x + 2 \equiv 3x + 1 \pmod 5$
$\implies 2x \equiv 1 \pmod 5 \implies x \equiv 3 \pmod 5$
3 is multiplicative inverse of 2 modulo 5.
Good when modulus is prime!!

Two points make a line.

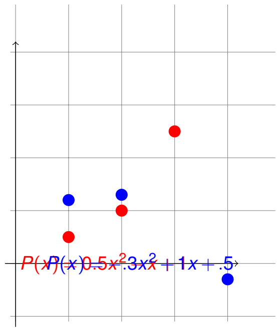
Fact: Exactly 1 degree $\leq d$ polynomial contains $d+1$ points. [2]

Two points specify a line. Three points specify a parabola.

Modular Arithmetic Fact: Exactly 1 degree $\leq d$ polynomial with arithmetic modulo prime p contains $d+1$ pts.

[2] Points with different x values.

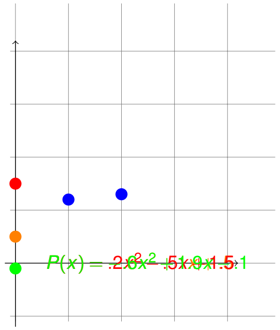
3 points determine a parabola.



Fact: Exactly 1 degree $\leq d$ polynomial contains $d+1$ points. [3]

[3] Points with different x values.

2 points not enough.



There is $P(x)$ contains blue points and *any* $(0,y)$!

Modular Arithmetic Fact and Secrets

Modular Arithmetic Fact: Exactly 1 degree $\leq d$ polynomial with arithmetic modulo prime p contains $d+1$ pts.

Shamir's k out of n Scheme:
Secret $s \in \{0,\ldots,p-1\}$

1. Choose $a_0 = s$, and random $a_1,\ldots,a_{k-1}$.
2. Let $P(x) = a_{k-1}x^{k-1} + a_{k-2}x^{k-2} + \cdots a_0$ with $a_0 = s$.
3. Share i is point $(i, P(i) \mod p)$.

Roubustness: Any k shares gives secret.
Knowing k pts $\implies$ only one $P(x)$ $\implies$ evaluate $P(0)$.
Secrecy: Any $k-1$ shares give nothing.
Knowing $\leq k-1$ pts $\implies$ any $P(0)$ is possible.

From $d+1$ points to degree d polynomial?

For a line, $a_1x + a_0 = mx + b$ contains points $(1,3)$ and $(2,4)$.

$$\begin{aligned} P(1) = m(1)+b &\equiv m+b \equiv 3 \pmod 5 \\ P(2) = m(2)+b &\equiv 2m+b \equiv 4 \pmod 5 \end{aligned}$$

Subtract first from second..

$$\begin{aligned} m+b &\equiv 3 \pmod 5 \\ m &\equiv 1 \pmod 5 \end{aligned}$$

Backsolve: $b \equiv 2 \pmod 5$. Secret is 2.

And the line is...

$$x + 2 \quad \text{mod } 5.$$

Quadratic

For a quadratic polynomial, $a_2x^2 + a_1x + a_0$ hits $(1,2);(2,4);(3,0)$.
Plug in points to find equations.

$$\begin{aligned} P(1) = a_2 + a_1 + a_0 &\equiv 2 \pmod 5 \\ P(2) = 4a_2 + 2a_1 + a_0 &\equiv 4 \pmod 5 \\ P(3) = 4a_2 + 3a_1 + a_0 &\equiv 0 \pmod 5 \end{aligned}$$

$$\begin{aligned} a_2 + a_1 + a_0 &\equiv 2 \pmod 5 \\ 3a_1 + 2a_0 &\equiv 1 \pmod 5 \\ 4a_1 + 2a_0 &\equiv 2 \pmod 5 \end{aligned}$$

Subtracting 2nd from 3rd yields: $a_1 = 1$.
$a_0 = (2 - 4(a_1))2^{-1} = (-2)(2^{-1}) = (3)(3) = 9 \equiv 4 \pmod 5$
$a_2 = 2 - 1 - 4 \equiv 2 \pmod 5$.

So polynomial is $2x^2 + 1x + 4 \pmod 5$

In general..

Given points: $(x_1,y_1);(x_2,y_2)\cdots(x_k,y_k)$.

Solve...

$$\begin{aligned} a_{k-1}x_1^{k-1} + \cdots + a_0 &\equiv y_1 \pmod p \\ a_{k-1}x_2^{k-1} + \cdots + a_0 &\equiv y_2 \pmod p \\ &\cdot \\ &\cdot \\ a_{k-1}x_k^{k-1} + \cdots + a_0 &\equiv y_k \pmod p \end{aligned}$$

Will this always work?

As long as solution **exists** and it is **unique!** And...

Modular Arithmetic Fact: Exactly 1 degree $\leq d$ polynomial with arithmetic modulo prime p contains $d+1$ pts.

Another Construction: Interpolation!

For a quadratic, $a_2x^2 + a_1x + a_0$ hits $(1,3);(2,4);(3,0)$.

Find $\Delta_1(x)$ polynomial contains $(1,1);(2,0);(3,0)$.

Try $(x-2)(x-3) \pmod 5$.

Value is 0 at 2 and 3. Value is 2 at 1. Not 1! Doh!!
So "Divide by 2" or multiply by 3.
$\Delta_1(x) = (x-2)(x-3)(3) \pmod 5$ contains $(1,1);(2,0);(3,0)$.

$\Delta_2(x) = (x-1)(x-3)(4) \pmod 5$ contains (1,0);(2,1);(3,0).

$\Delta_3(x) = (x-1)(x-2)(3) \pmod 5$ contains (1,0);(2,0);(3,1).

But wanted to hit $(1,3);(2,4);(3,0)$!

$P(x) = 3\Delta_1(x) + 4\Delta_2(x) + 0\Delta_3(x)$ works.

Same as before?

...after a lot of calculations... $P(x) = 2x^2 + 1x + 4 \mod 5$.

The same as before!

Fields.. .

Flowers, and grass, oh so nice.

Set and two commutative operations: addition and multiplication with additive/multiplicative identities and inverses expect for additive identity has no mulitplicative inverse.

E.g., Reals, rationals, complex numbers.
Not E.g., the integers, matrices.

We will work with polynomials with arithmetic modulo p.

Addition is cool. Inherited from integers and integer division (remainders).
Multiplicative inverses due to $gcd(x,p) = 1$, forall $x \in \{1,\ldots,p-1\}$

Delta Polynomials: Concept.

For set of x-values, $x_1,\ldots,x_{d+1}$.

$$\Delta_i(x) = \begin{cases} 1, & \text{if } x = x_i. \\ 0, & \text{if } x = x_j \text{ for } j \neq i. \\ ?, & \text{otherwise.} \end{cases} \tag{1}$$

Given $d+1$ points, use Δ_i functions to go through points?
$(x_1,y_1),\ldots,(x_{d+1},y_{d+1})$.

Will $y_1\Delta_1(x)$ contain (x_1,y_1)?

Will $y_2\Delta_2(x)$ contain (x_2,y_2)?

Does $y_1\Delta_1(x) + y_2\Delta_2(x)$ contain
(x_1,y_1)? and (x_2,y_2)?

See the idea? Function that contains all points?
$P(x) = y_1\Delta_1(x) + y_2\Delta_2(x)\ldots + y_{d+1}\Delta_{d+1}(x)$.

There exists a polynomial...

Modular Arithmetic Fact: Exactly 1 degree $\leq d$ polynomial with arithmetic modulo prime p contains $d+1$ pts.

Proof of at least one polynomial:
Given points: $(x_1,y_1);(x_2,y_2)\cdots(x_{d+1},y_{d+1})$.

$$\Delta_i(x) = \frac{\prod_{j\neq i}(x-x_j)}{\prod_{j\neq i}(x_i-x_j)} = \prod_{j\neq i}(x-x_j)\prod_{j\neq i}(x_i-x_j)^{-1}$$

Numerator is 0 at $x_j \neq x_i$.

"Denominator" makes it 1 at x_i.

And..

$$P(x) = y_1\Delta_1(x) + y_2\Delta_2(x) + \cdots + y_{d+1}\Delta_{d+1}(x).$$

hits points $(x_1,y_1);(x_2,y_2)\cdots(x_{d+1},y_{d+1})$. Degree d polynomial!

Construction proves the existence of a polynomial!

Example.

$\Delta_i(x) = \frac{\prod_{j\neq i}(x-x_j)}{\prod_{j\neq i}(x_i-x_j)}$.

Degree 1 polynomial, $P(x)$, that contains $(1,3)$ and $(3,4)$?

Work modulo 5.

$\Delta_1(x)$ contains $(1,1)$ and $(3,0)$.

$\Delta_1(x) = \frac{(x-3)}{1-3} = \frac{x-3}{-2}$
$= 2(x-3) = 2x-6 = 2x+4 \pmod 5$.

For a quadratic, $a_2x^2 + a_1x + a_0$ hits $(1,3);(2,4);(3,0)$.

Work modulo 5.

Find $\Delta_1(x)$ polynomial contains $(1,1);(2,0);(3,0)$.

$\Delta_1(x) = \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{(x-2)(x-3)}{2} = (2)^{-1}(x-2)(x-3) = 3(x-2)(x-3)$
$= 3x^2 + 3 \pmod 5$

Put the delta functions together.

In general.

Given points: $(x_1,y_1);(x_2,y_2)\cdots(x_k,y_k)$.

$$\Delta_i(x) = \frac{\prod_{j\neq i}(x-x_j)}{\prod_{j\neq i}(x_i-x_j)} = \prod_{j\neq i}(x-x_j)\prod_{j\neq i}(x_i-x_j)^{-1}$$

Numerator is 0 at $x_j \neq x_i$.

Denominator makes it 1 at x_i.

And..

$$P(x) = y_1\Delta_1(x) + y_2\Delta_2(x) + \cdots + y_k\Delta_k(x).$$

hits points $(x_1,y_1);(x_2,y_2)\cdots(x_k,y_k)$.

Construction proves the existence of the polynomial!

Uniqueness.

Uniqueness Fact. At most one degree d polynomial hits $d+1$ points.

Roots fact: Any nontrivial degree d polynomial has at most d roots.

A line, a degree 1 polynomial, can intersect $y=0$ at most one time or be $y=0$.

A parabola (degree 2), can intersect $y=0$ at most twice or be $y=0$.

Proof:
Assume two different polynomials $Q(x)$ and $P(x)$ hit the points.

$R(x) = Q(x) - P(x)$ has $d+1$ roots and is degree d.
Contradiction.

$\square$

Must prove **Roots fact.**

Polynomial Division.

Divide $4x^2 - 3x + 2$ by $(x-3)$ modulo 5.

```
                 4 x + 4 r 4
         -----------------
x - 3 )  4x^2 - 3 x + 2
         4x^2 - 2x
         ----------
                4x + 2
                4x - 2
                -------
                     4
```

$4x^2 - 3x + 2 \equiv (x-3)(4x+4) + 4 \pmod 5$

In general, divide $P(x)$ by $(x-a)$ gives $Q(x)$ and remainder r.

That is, $P(x) = (x-a)Q(x) + r$

Only d roots.

Lemma 1: $P(x)$ has root a iff $P(x)/(x-a)$ has remainder 0:
$P(x) = (x-a)Q(x)$.
Proof: $P(x) = (x-a)Q(x) + r$.
Plugin a: $P(a) = r$.
It is a root if and only if $r = 0$.

$\square$

Lemma 2: $P(x)$ has d roots; $r_1, \dots, r_d$ then
$P(x) = c(x-r_1)(x-r_2)\cdots(x-r_d)$.
Proof Sketch: By induction.

Induction Step: $P(x) = (x-r_1)Q(x)$ by Lemma 1. $Q(x)$ has smaller degree so use the induction hypothesis. $\square$

$d+1$ roots implies degree is at least $d+1$.

Roots fact: Any degree d polynomial has at most d roots.

Finite Fields

Proof works for reals, rationals, and complex numbers.

..but not for integers, since no multiplicative inverses.

Arithmetic modulo a prime p has multiplicative inverses..

..and has only a finite number of elements.

Good for computer science.

Arithmetic modulo a prime m is a **finite field** denoted by F_m or $GF(m)$.

Intuitively, a field is a set with operations corresponding to addition, multiplication, and division.

Secret Sharing

Modular Arithmetic Fact: Exactly one polynomial degree $\leq d$ over $GF(p)$, $P(x)$, that hits $d+1$ points.

Shamir's k out of n Scheme:
Secret $s \in \{0, \dots, p-1\}$

1. Choose $a_0 = s$, and randomly $a_1, \dots, a_{k-1}$.
2. Let $P(x) = a_{k-1}x^{k-1} + a_{k-2}x^{k-2} + \cdots a_0$ with $a_0 = s$.
3. Share i is point $(i, P(i) \mod p)$.

Roubustness: Any k knows secret.
Knowing k pts, only one $P(x)$, evaluate $P(0)$.
Secrecy: Any $k-1$ knows nothing.
Knowing $\leq k-1$ pts, any $P(0)$ is possible.

Minimality.

Need $p > n$ to hand out n shares: $P(1) \dots P(n)$.

For b-bit secret, must choose a prime $p > 2^b$.

Theorem: There is always a prime between n and $2n$.
Chebyshev said it,
And I say it again,
There is always a prime
Between n and $2n$.

Working over numbers within 1 bit of secret size. **Minimality.**

With k shares, reconstruct polynomial, $P(x)$.

With $k-1$ shares, any of p values possible for $P(0)$!

(Almost) any b-bit string possible!

(Almost) the same as what is missing: one $P(i)$.

Runtime.

Runtime: polynomial in k, n, and $\log p$.

1. Evaluate degree $k-1$ polynomial n times using $\log p$-bit numbers.
2. Reconstruct secret by solving system of k equations using $\log p$-bit arithmetic.

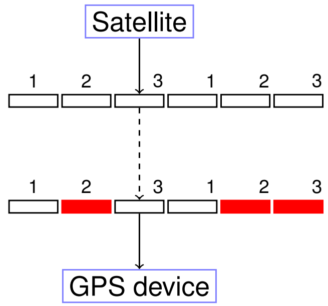
A bit more counting.

What is the number of degree d polynomials over $GF(m)$?

- m^{d+1}: $d+1$ coefficients from $\{0,\ldots,m-1\}$.
- m^{d+1}: $d+1$ points with y-values from $\{0,\ldots,m-1\}$

Infinite number for reals, rationals, complex numbers!

Erasure Codes.



3 packet message. So send 6!

Lose 3 out 6 packets.

Gets packets 1,1,and 3.

Solution Idea.

n packet message, channel that loses k packets.

Must send $n+k$ packets!

Any n packets should allow reconstruction of n packet message.

Any n point values allow reconstruction of degree $n-1$ polynomial.

Alright!!!!!!

Use polynomials.

The Scheme

Problem: Want to send a message with n packets.

Channel: Lossy channel: loses k packets.

Question: Can you send $n+k$ packets and recover message?

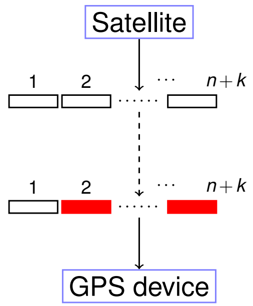
A degree $n-1$ polynomial determined by any n points!

Erasure Coding Scheme: message = $m_0, m_1 \ldots, m_{n-1}$.

1. Choose prime $p \approx 2^b$ for packet size b.
2. $P(x) = m_{n-1}x^{n-1} + \cdots m_0 \pmod p$.
3. Send $P(1), \ldots, P(n+k)$.

Any n of the $n+k$ packets gives polynomial ...and message!

Erasure Codes.



n packet message. So send $n+k$!

Lose k packets.

Any n packets is enough!

n packet message.

Optimal.

Information Theory.

Size: Can choose a prime between 2^{b-1} and 2^b.
(Lose at most 1 bit per packet.)

But: packets need label for x value.

There are Galois Fields $GF(2^n)$ where one loses nothing.

– Can also run the Fast Fourier Transform.

In practice, $O(n)$ operations with almost the same redundancy.

Comparison with Secret Sharing: information content.

Secret Sharing: each share is size of whole secret.
Coding: Each packet has size $1/n$ of the whole message.

Erasure Code: Example.

Send message of 1,4, and 4.

Make polynomial with $P(1) = 1$, $P(2) = 4$, $P(3) = 4$.

How?

Lagrange Interpolation.
Linear System.

Work modulo 5.

$P(x) = x^2 \pmod 5$
$P(1) = 1, P(2) = 4, P(3) = 9 = 4 \pmod 5$

Send $(0, P(0)) \ldots (5, P(5))$.

6 points. Better work modulo 7 at least!

Why? $(0, P(0)) = (5, P(5)) \pmod 5$

Example

Make polynomial with $P(1) = 1$, $P(2) = 4$, $P(3) = 4$.

Modulo 7 to accommodate at least 6 packets.

Linear equations:

$$\begin{aligned} P(1) = a_2 + a_1 + a_0 &\equiv 1 \pmod 7 \\ P(2) = 4a_2 + 2a_1 + a_0 &\equiv 4 \pmod 7 \\ P(3) = 2a_2 + 3a_1 + a_0 &\equiv 4 \pmod 7 \end{aligned}$$

$6a_1 + 3a_0 = 2 \pmod 7$, $5a_1 + 4a_0 = 0 \pmod 7$

$a_1 = 2a_0$. $a_0 = 2 \pmod 7$ $a_1 = 4 \pmod 7$ $a_2 = 2 \pmod 7$

$P(x) = 2x^2 + 4x + 2$

$P(1) = 1$, $P(2) = 4$, and $P(3) = 4$

Send
Packets: $(1,1), (2,4), (3,4), (4,7), (5,2), (6,0)$

Notice that packets contain “x-values”.

Bad reception!

Send: $(1,1), (2,4), (3,4), (4,7), (5,2), (6,0)$

Recieve: (1,1) (2,4), (6,0)
Reconstruct?

Format: $(i, R(i))$.

Lagrange or linear equations.

$$\begin{aligned} P(1) = a_2 + a_1 + a_0 &\equiv 1 \pmod 7 \\ P(2) = 4a_2 + 2a_1 + a_0 &\equiv 4 \pmod 7 \\ P(6) = 2a_2 + 3a_1 + a_0 &\equiv 0 \pmod 7 \end{aligned}$$

Channeling Sahai ...

$P(x) = 2x^2 + 4x + 2$

Message? $P(1) = 1, P(2) = 4, P(3) = 4$.

Questions for Review

You want to encode a secret consisting of 1,4,4.

How big should modulus be?
Larger than 144 and prime!

Remember the secret, $s = 144$, must be one of the possible values.

You want to send a message consisting of packets 1,4,2,3,0

through a noisy channel that loses 3 packets.

How big should modulus be?
Larger than 8 and prime!

The other constraint: arithmetic system can represent $0, 1, 2, 3, 4$.

Send n packets b-bit packets, with k errors.
Modulus should be larger than $n + k$ and also larger than 2^b.

Polynomials.

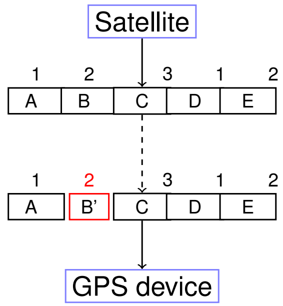
- ..give Secret Sharing.
- ..give Erasure Codes.

Error Correction:

Noisy Channel: corrupts k packets. (rather than loss.)

Additional Challenge: Finding which packets are corrupt.

Error Correction



3 packet message. Send 5.

Corrupts 1 packets.

The Scheme.

Problem: Communicate n packets $m_1, \ldots, m_n$ on noisy channel that corrupts $\leq k$ packets.

Reed-Solomon Code:

1. Make a polynomial, $P(x)$ of degree $n-1$, that encodes message.
   - $P(1) = m_1, \ldots, P(n) = m_n$.
   - Comment: could encode with packets as coefficients.
2. Send $P(1), \ldots, P(n+2k)$.

After noisy channel: Recieve values $R(1), \ldots, R(n+2k)$.

Properties:
(1) $P(i) = R(i)$ for at least $n+k$ points i,
(2) $P(x)$ is unique degree $n-1$ polynomial that contains $\geq n+k$ received points.

Properties: proof.

$P(x)$: degree $n-1$ polynomial.
Send $P(1), \ldots, P(n+2k)$
Receive $R(1), \ldots, R(n+2k)$
At most k i's where $P(i) \neq R(i)$.

Properties:
(1) $P(i) = R(i)$ for at least $n+k$ points i,
(2) $P(x)$ is unique degree $n-1$ polynomial that contains $\geq n+k$ received points.

Proof:
(1) Sure. Only k corruptions.
(2) Degree $n-1$ polynomial $Q(x)$ consistent with $n+k$ points.
$Q(x)$ agrees with $R(i)$, $n+k$ times.
$P(x)$ agrees with $R(i)$, $n+k$ times.

Total points contained by both:	$2n+2k$.	P	Pigeons.
Total points to choose from	: $n+2k$.	H	Holes.
Points contained by both	: $\geq n$.	$\geq P-H$	Collisions.

$\implies Q(i) = P(i)$ at n points.
$\implies Q(x) = P(x)$. □

Example.

Message: $3, 0, 6$.

Reed Solomon Code: $P(x) = x^2 + x + 1 \pmod 7$ has $P(1) = 3, P(2) = 0, P(3) = 6$ modulo 7.

Send: $P(1) = 3, P(2) = 0, P(3) = 6, P(4) = 0, P(5) = 3$.

(Aside: Message in plain text!)

Receive $R(1) = 3, R(2) = 1, R(3) = 6, R(4) = 0, R(5) = 3$.

$P(i) = R(i)$ for $n+k = 3+1 = 4$ points.

Slow solution.

Brute Force:
For each subset of $n+k$ points
Fit degree $n-1$ polynomial, $Q(x)$, to n of them.
Check if consistent with $n+k$ of the total points.
If yes, output $Q(x)$.

- For subset of $n+k$ pts where $R(i) = P(i)$, method will reconstruct $P(x)$!
- For any subset of $n+k$ pts,
  1. there is unique degree $n-1$ polynomial $Q(x)$ that fits n of them
  2. and where $Q(x)$ is consistent with $n+k$ points $\implies P(x) = Q(x)$.

Reconstructs $P(x)$ and only $P(x)$!!

Example.

Received $R(1) = 3, R(2) = 1, R(3) = 6, R(4) = 0, R(5) = 3$

Find $P(x) = p_2 x^2 + p_1 x + p_0$ that contains $n+k = 3+1$ points.

All equations..

$$\begin{aligned}
p_2 + p_1 + p_0 &\equiv 3 \pmod 7 \\
4p_2 + 2p_1 + p_0 &\equiv 1 \pmod 7 \\
2p_2 + 3p_1 + p_0 &\equiv 6 \pmod 7 \\
2p_2 + 4p_1 + p_0 &\equiv 0 \pmod 7 \\
1p_2 + 5p_1 + p_0 &\equiv 3 \pmod 7
\end{aligned}$$

Assume point 1 is wrong and solve..no consistent solution!
Assume point 2 is wrong and solve...consistent solution!

In general..

$P(x) = p_{n-1}x^{n-1} + \cdots p_0$ and receive $R(1), \ldots R(m = n+2k)$.

$$\begin{aligned}
p_{n-1} + \cdots p_0 &\equiv R(1) \pmod{p} \\
p_{n-1}2^{n-1} + \cdots p_0 &\equiv R(2) \pmod{p} \\
&\cdot \\
p_{n-1}i^{n-1} + \cdots p_0 &\equiv R(i) \pmod{p} \\
&\cdot \\
p_{n-1}(m)^{n-1} + \cdots p_0 &\equiv R(m) \pmod{p}
\end{aligned}$$

Error!! .... Where???
Could be anywhere!!! ...so try everywhere.
Runtime: $\binom{n+2k}{k}$ possibilitities.

Something like $(n/k)^k$ ...Exponential in k!.

How do we find where the bad packets are efficiently?!?!?!

Ditty...

Where oh where can my bad packets be ...

On Thursday.