Erasure Errors

- use polynomial interpolation to get missing point.
- Sending n packets, guard against k errors, send $n+k$ packets.
- GF(q) - each packet can be encoded mod q, so q > largest number in data & send packets, ensure $n+k \leq q$.
- use delta reconstruction $\Delta_3(x) = \frac{(x-a_1)(x-a_2)(x-a_4)}{(a_3-a_1)(a_3-a_2)(a_3-a_4)}$
- Add all up: $y_1\Delta_1 + y_2\Delta_2$... to get original polynomial.

General Errors

- n length message, k errors, send $n+2k$ message
- $Q(x) = P(x)E(x) \rightarrow Q(x)/E(x) = P(x)$
- Sending the message:
  1. get n points, fine deg(n-1) polynomial.
  2. evaluate $2k$ more points.
  3. send $P(i)$ for $i \in 0, 1, ...(n+2k)$
- Decoding the message:
  1. Note - rememeber $GF(p)$, so mod stuff!
  2. get $n+2k$ points
  3. $deg(E(x)) = k$, $(E(x) = x^k + b_{k-1}x^{k-1} + \cdots + b_1x + b_0)$

     If the $e_1, e_2, \ldots, e_k$ packets are corrupted so that the received points are $r_1, r_2, \ldots, r_k$ we can define $E(x) = (x-e_1)(x-e_2)\ldots(x-e_k)$
  4. $Q(x)$ is degree $n+k-1$, $(Q(x) = a_{n+k-1}x^{n+k-1} + \cdots + a_1x + a_0)$
  5. for each point x_i, subsitute in to $Q(i) = r_iE(x)$
  6. Solve system for $a_1, a_2 ... b_1, b_2 ...$
  7. These are coefficents of $Q(x)$ and $E(x)$. $\frac{Q(x)}{E(x)} = P(x)$

Eulerian Walks/Tours, Dis 6

- The necessary and sufficient conditions for an undirected graph to have an Eulerian walk.
  - If an undirected graph G has an Eulerian walk W, the graph can have at most two odd degree vertices.
  - If a connected graph has at most two odd degree vertices, it has an Eulerian walk.
- If G has an Eulerian tour, its edge set can be decomposed into cycles. Proved using induction on the number of edges. $\sum_{v=i} deg(v_i) = 2 \mid E \mid$

Counting
There are $n!$ ways to order n objects.
First Rule of Counting:
Order matters, w/o replacement:
$n*(n-1)*\cdots*(n-(k-1))$ Example: 52 cards, draw 5 52 51 50 49 48 n!(nk)!
Order matters, w/ replacement: n^k
Example 2^n ways of flipping a {H, T} coin n times
Second Rule of Counting:
Order does not matter, w/o replacement: $\binom{n}{k}$
Example Example: 52 cards, Queen of Hearts, King of Spades, Jack of Diamonds, Ace of Clubs, 10 of diamonds
$\binom{n}{k} = \frac{n!}{(n-k)!k!}$
Order doesn't matter, w/ replacement: Choose multisets of size k from set S with $\binom{n+k-1}{k}$
Example 3 types of veggies, pick 5 from an unlimited number
Think: Balls and Bins
Balls: number of servings we want to make (n balls)
Bins: different types of veggies we have (k bins)
$\binom{n-1+k}{n}$ ** Also equivalent to: $\binom{n-1+k}{k-1}$

Graphs
A directed graph $G(V, E)$ consists of a finite set of vertices V and a set of edges E. An edge (v, w) in a directed graph is usually indicated by drawing a line between v and w, with an arrow pointing towards w. Undirected graphs may be regarded as special kinds of directed graphs, in which $(u, v) \in E$ if and only if $(v, u) \in E$.

- A *path* in a directed graph $G = (V, E)$ is a sequence of neighboring edges.
- A *cycle* is a path that begins and ends at the same vertex. A graph is said to be connected if there is a path between any two distinct vertices.
- An *Eulerian tour* or *Eulerian cycle* is a cycle that uses each edge exactly once.

Eulers Theorem: An undirected graph G=(V, E) has an Eulerian tour if and only if the graph is connected (except possibly for isolated vertices) and even degree.
A **Hamiltonian path** of a graph is a sequence of vertices $v_0, v_1, \ldots, v_k$ such that:

- Each vertex appears exactly once in the sequence.
- Each pair of consecutive vertices is connected by an edge.
- v_0 and v_k are connected by an edge.

Hypercubes
The vertex set of an n-dimensional hypercube is $0, 1^n$ (i.e., there are exactly 2^n vertices, each labeled with a distinct n-bit string), and with an edge between vertices x and y iff x and y differ in exactly one bit position.
Another recursive definition of the hypercube: The n-dimensional hypercube consists of two copies of the $n-1$-dimensional hypercube (the 0-subcube and the 1-subcube), and with edges between corresponding vertices in the two subcubes. i.e., there is an edge between vertex x in the 0-subcube (also denoted as vertex $0x$) and vertex x in the 1-subcube (denoted $1x$).
Theorem $|ES| \geq |S|$.
Claim Total number of edges in n-dimensional hypercube is $n2^{n-1}$.
Proof: Each vertex has n edges incident to it, since there are exactly n bit positions that can be toggled to get an edge. Since each edge is counted twice, once from each endpoint, this yields a grand total of $\frac{n2^n}{2}$.

HW6 Problem 5 (Touring the hypercube)
Let G be a hypercube of dimension n, i.e.
The vertices of G are the binary strings of length n.
u and v are connected by an edge if they differ in exactly one location.
Claim The hypercube has an Eulerian tour iff n is even.
Proof In the n-dimensional hypercube, every vertex has degree n. If n is odd, then from lecture there can be no Eulerian tour. On the other hand, the hypercube is connected: we can get from any one bit-string x to any other y by flipping the bits they differ in one at a time. Therefore, when n is even, since every vertex has even degree and the graph is connected, there is an Euler tour.
Claim The hypercube has a Hamiltonian tour.
Proof (By induction on n) When $n = 1$, there are two vertices connected by an edge; we can form a Hamiltonian tour by walking from one to the other and then back. Let $n \geq 1$ and suppose the n-dimensional hypercube has a Hamiltonian tour. Let H be the $n+1$-dimensional hypercube, and let H_0 be the n-dimensional subcube consisting of those strings with final bit b. By the inductive hypothesis, there is some hamiltonian tour T on the n-dimensional hypercube. Now consider the following tour in H. Start at an arbitrary vertex x_0 in H_0, and follow the tour T except for the very last step to vertex y_0 (so that the next step would bring us back to x_0). Next take the edge from y_0 to y_1 to enter cube H_1. Next, follow the tour T in H_1 backwards from y_1, except the very last step, to arrive at x_1. Finally, take the step from x_1 to x_0 to complete the tour. By assumption, the tour T visits each vertx in each subcube exactly once, so our complete tour visits each vertex in the whole cube exactly once.

Combinatorial Proofs
$\binom{n}{k+1} = \binom{n-1}{k} + \binom{n-2}{k} + \cdots + \binom{k}{k}$
$\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n$
Discrete Probability
Random Experiment: A probabilistic experiment consists of drawing a sample of k elements from a set S of cardinality n. The outcome of the random experiment is called a *sample point*. The *sample space* is the set of all possible outcomes.
HW7: Proof of Fermats Little Theorem
$\binom{b}{k} = 0 \bmod p$ for $0 < k < p$
$(x+y)^p \equiv \sum_{k=0}^{p} \binom{p}{k} x^k y^{p-k} \equiv x^p + y^p (\mathrm{mod} p)$
Probability Spaces
A probability space is a sample space Ω, together with a probability $Pr[\omega]$ for each sample point ω, such that

1. $0 \leq Pr[\omega] \leq 1$ for all $\omega \in \Omega$.
2. $\sum_{\omega \in \Omega} Pr[\omega] = 1$, i.e., the sum of the probabilities of all outcomes is 1.

For any event $A \subseteq \Omega$, we define the probability of A to be $Pr[A] = \sum Pr[\omega]$.
$Pr[A] = \frac{\#ofsamplepointsinA}{\#ofsamplepointsin\Omega} = \frac{|A|}{|\Omega|}$
Conditional Probability
Definition (conditional probability): For events A, B in the same probability space, such that $Pr[B] > 0$, the conditional probability of A given B is $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

Random Message, HW9 #5
Let p be a large prime, and let $P(x)$ be a polynomial of degree (at most) 2 over $GF(p)$. Suppose Alice is trying to reconstruct the message $(P(1), P(2))$. Assume that Alice has no prior information about the message, so that every pair (i, j) has probability $1/p^2$.

1. Suppose Alice learns that $P(5) = 3$. What is the probability that the message $(P(1), P(2)) = (1, 1)$?

   Solution To construct a polynomial of degree at most 2, we need 3 points. One of the points is fixed and we have p^2 possible pairs for $(P(1), P(2))$. Since $(1, 1)$ is one of the possible pairs,

   $Pr[(P(1), P(2)) = (1, 1)] = \frac{1}{p^2}$

2. Now suppose Alice learns that $P(4) = P(5) = 1$. What is the probability that the message $(P(1), P(2)) = (1, 1)$?

   Solution Now we have 2 of our 3 points fixed. We have $Pr[P(1) = 1] = \frac{1}{p}$ and $Pr[P(2) = 1] = 1$ which we can verify using the polynomial.

   $Pr[P(1) = 1 \cap P(2) = 1 \mid P(4) = 1 \cap P(5) = 1]$

   $Pr[P(1) \mid P(4) = 1 \cap P(5) = 1] * Pr[P(2) = 1 \mid P(2) = 1 \cap P(4) = 1 \cap P(5) = 1]$

   $= \frac{1}{p}$

Formulas/Definitions
disjoint: outcomes do not overlap
independent: outcome of one event does not affect probability of other event
Bayesian Inference is a way to *update knowledge* after making an observation.

- **Bayes Rule:** $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]} = \frac{Pr[B|A]Pr[A]}{Pr[B]}$
- Total Probability Rule:
  $\Pr[A] = \Pr[A|B]\Pr[B] + \Pr[A|\overline{B}]\Pr[\overline{B}]$
  $Pr[A] = Pr[A \cap \overline{B}] + Pr[A \cap B]$
  $Pr[\overline{A}] = Pr[\overline{A} \cap \overline{B}] + Pr[\overline{A} \cap B]$
- $\Pr[A \cap B] = \Pr[A] * \Pr[B]$, intersection, AND. (assume independent)
- $\Pr[A \cup B] = \Pr[A] + \Pr[B]$, union, OR (independent)
- $\Pr[A \cap B] = \Pr[A]\Pr[B|A]$, intersection(dependent)
- $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cup B]$, union(dependent)
- Conditional probability - A given B: $\Pr[A|B] = \frac{P[A \cap B]}{\Pr[B]}$
- event C we get exactly r results of probability p given n trials $= P[C] = \binom{n}{r} p^r (1-p)^{n-r}$
- For events $A_1, \ldots, A_n$ in some probability space, we have $Pr[\cup_{i=1}^{n} Pr[A_i] = \sum Pr[A_i] - \sum Pr[A_i \cap A_j] + \sum Pr[A_i \cap A_j \cap A_k] - \cdots \pm \cup_{i=1}^{n} Pr[A_i]$. (basically count all individual events, subtract intersections of pairs, add back intersectons of triples, repeat alternating.)

Balls and Bins

- $\Pr[\text{bin 1 is empty}] = (\frac{n-1}{n})^m = (1 - \frac{1}{n})^m$
- $\Pr[\text{first k out of n bins empty}] = (1 - \frac{k}{n})^m$
- Given k out of n bins empty, Pr[(k+1)th bin empty] $= \frac{(1-\frac{k+1}{n})^n}{(1-\frac{k}{n})^n} = (\frac{m-k-1}{m-k})^n$
- Birthday paradox. Probability NOT same birthday is: $\frac{365*364*\ldots*(365-n+1)}{365^n}$, so with 1 - *this* we get 50% with 23 people.

Fermat's Little Theorem: For any prime p and any $a \in \{1, 2, ..., p-1\}$, we have $a^{p-1} = 1 \bmod p$.
LaGrange:
Given these three points find the polynomial: (1,0) (2, 1), (3,1)
$\Delta x_1 = ((x-2)(x-3))/((1-2)(1-3))$
$\Delta x_2 = ((x-1)(x-3))/((2-1)(2-3))$
$\Delta x_3 = ((x-1)(x-2))/((3-1)(3-2))$
$P(x) = y_1 \Delta x_1 + y_2 \Delta x_2 + y_3 \Delta x_3$

Definition 10.1 (independence):
Two events A, B in the same probability space are independent if $Pr[A \cap B] = Pr[A] * Pr[B]$ or $Pr[A \mid B] = Pr[A]$ or $Pr[B \mid A] = Pr[B]$.
Definition 10.2 (mutual independence): Events $A_1, \ldots, A_n$ are mutually independent if for every subset $I \subseteq \{1, \ldots, n\}$, $Pr[\cap_{i \in I} A_i] = \prod_{i \in I} Pr[A_i]$.
Theorem 10.1: [Product Rule] For any events A, B, we have $Pr[A \cap B] = Pr[A]Pr[B|A]$.
Theorem 10.2: [Inclusion/Exclusion] For events $A_1, \ldots, A_n$ in some probability space, we have $Pr[\cup_{i=1}^{n} Pr[A_i] = \sum Pr[A_i] - \sum Pr[A_i \cap A_j] + \sum Pr[A_i \cap A_j \cap A_k] - \cdots \pm \cup_{i=1}^{n} Pr[A_i]$.

Lines, F12 MT2 # 3
Assume that you are working modulo p, where p is a prime greater than 10. Select a random line (a polynomial $A(x)$ of degree at most 1).

1. What is the chance that it goes through a particular point, (x, y), for example if $(x, y) = (0, 5)$, the question asks what is the probability that $A(0) = 5$?

   Solution The total number of polynomials of the form $ax + b$ is p^2, since we can independently choose a and b. By Lagrange interpolation, every distinct value of $y^{‘}$ in {0, 1,. . . ,p-1}, there is a distinct line connecting (x, y) and (x+1, $y^{‘}$); moreover, every line passing through (x, y) must be one of those p lines. Thus, there are exactly p such lines, and the probability is $\frac{p}{p^2} = \frac{1}{p}$.

2. What is the chance that it goes through two particular points $(x_1, y_1), (x_2, y_2)$, where $x_1 \neq x_2$?

   Solution Again by Lagrange interpolation, of the p^2 lines, there is exactly 1 connecting those 2 points. Thus the probability is $\frac{1}{p^2}$.

3. What about 3 particular points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$, where x_1, x_2, x_3 are distinct?

   Solution There are two distinct cases. Use Lagrange interpolation to obtain the line connecting (x_1, y_1) and (x_2, y_2). If (x_3, y_3) lies on this line (i.e., the 3 points are collinear), there is exactly 1 such line, so the probability is $\frac{1}{p^2}$. If it does not, there is no such line and the probability is 0.