# Math Notes

**John L. Manferdelli**

These notes were written for my personal use, partly to learn tex.
They are clearly not written for third parties (second parties either)
and may be incomplete, inaccurate or even incoherent.
However, you are welcome to use them at your own risk.
I disclaim any and all liability for inaccuracy, infringement of any kind, or anything else.

Please send corrections to:
JohnManferdelli@hotmail.com, jlmUCB@yahoo.com,
jmanfer@microsoft.com, jlm@cs.washington.edu.

Last modified: 24 January 2009 14:43

# Chapter 1

# Math

## 1.1 Number Theory, Inequalities and Combinatorics

### 1.1.1 Basic Number Theory

**$\pi$ is irrational:** Lemma: Define $f_n(x) = \frac{x^n(1-x)^n}{n!}$ then (i) $f_n(x) = \frac{1}{n!}\sum_{i=n}^{2n} c_i x^i, c_i \in \mathbb{Z}$, (ii) For $0 < x < 1$, $0 < f(x) < \frac{1}{n!}$, (iii) the derivatives $f_n^{(k)}(0), f_n^{(k)}(1) \in \mathbb{Z}, k \geq 0$. Now, assume $r = \pi^2 = \frac{a}{b}$ is rational. Let $F_n(x) = b^n(r^n f_n(x) - r^{n-1} f_n^{(2)}(x) + \ldots + (-1)^n f_n^{(2n+2)}(x))$. For $0 \leq k \leq n, b^n r^{n-k} = a^{n-k}b^k$ so $F_n(0), F_n(1) \in \mathbb{Z}$. $f_n^{(2n+2)}(x) = 0$ so $F_n''(x) + rF_n(x) = ra^n f_n(x)$ and $\frac{d}{dx}(F'(x)sin(\pi x) - \pi F(x)cos(\pi x)) = ra^n f_n(x)sin(\pi x)$. Use this to show: $\pi a^n \int_0^1 f_n(x)sin(\pi x)dx = F_n(1) + F_n(0)$. Thus for all integers $n \geq 1$, $0 < \pi a^n \int_0^1 f_n(x)sin(\pi x)dx < \frac{\pi a^n}{n!}$. Picking $n$ large enough to $\frac{\pi a^n}{n!} < 1$, we get a contradiction.

**Transcendence of $e$:** If $f(x)$ is a polynomial of degree $r$, set $F(x) = f(x) + f'(x) \ldots + f^{(r)}(x)$. Then $F(i) - e^i F(0) = -ie^{i(1-\theta_i)} f(i\theta_i) = \epsilon_i$. Suppose $e$ satisfies $g(e) = c_n e^n + \ldots + c_0 = 0$. Then $c_n F(n) + \ldots + c_0 F(0) = c_1\epsilon_1 + c_2\epsilon_2 + \ldots + c_n\epsilon_n$. Put $f(x) = \frac{1}{(p-1)!}x^{p-1}(1-x)^p(2-x)^p \ldots (n-x)^p$. $p \mid F(i), i > 0$ but $p \nmid F(0)$. So, $c_n F(n) + \ldots + c_0 F(0)$ is an integer not divisible by $p$ but $c_n F(n) + \ldots + c_0 F(0) = c_1\epsilon_1 + c_2\epsilon_2 + \ldots + c_n\epsilon_n$. Now, let $p \rightarrow \infty$.

**Wilson:** $(p-1)! = (-1) \pmod{p}$. Proof: There are only 2 solutions to $x^2 = 1 \pmod{p}, namely, \pm 1$ all other multiplicative elements can be paired with their inverses and cancel.

$x^2 = -1 \pmod{p}$ iff $p = 2$ or $p = 1 \pmod{4}$. Proof: $(p-1)! = -1 = (-1)^{\frac{p-1}{2}} \prod_{j \in \{1,2,\ldots,\frac{p-1}{2}\}} j^2 \pmod{p}$, if $p = 1 \pmod{4}$, first factor is 1 and thus $(\prod_{j \in \{1,2,\ldots,\frac{p-1}{2}\}} j)^2 = -1 \pmod{p}$.

If $p = 1 \pmod{4} : \exists a, b : a^2 + b^2 = p$. Proof: $\exists x : x^2 + 1 = rp$. Set $k = \lfloor \sqrt{p} \rfloor, k < \sqrt{p} < k+1$. Set $f(u,v) = ux + v$; consider $S = \{(u,v) : 0 \leq u \leq k, 0 \leq v \leq k\}$. $|S| = (k+1)^2 > p$, so $\exists u_1, u_2, v_1, v_2 : f(u_1, v_1) = f(u_2, v_2)$ and $a = u_1 - u_2, b = v_1 - v_2$ then $a + bx = 0 \pmod{p}$. Now $a^2 + b^2 = a^2 + a^2x^2 = 0 \pmod{p}$. $|a| < \sqrt{p}$ and $|b| < \sqrt{p}$ so $0 < a^2 + b^2 < 2p$ and $a^2 + b^2 = p$.

If $q \mid (a^2 + b^2)$ and $q = 3 \pmod{4}$ then $q \mid a$ and $q \mid b$. Proof suppose $(a, q) = 1$, pick $\overline{a} : a\overline{a} = 1 \pmod{q}$. $a^2 = -b^2 \pmod{q}$ so $-1 = (b\overline{a})^2 \pmod{q}$.

If $n = 2^\alpha \prod_{p=1 \pmod{4}} p^\beta \prod_{q=3 \pmod{4}} q^\gamma$ then $n = a^2 + b^2$ iff all $\gamma$ are even. Proof: Use $(a^2+b^2)(c^2+d^2) = (ac - db)^2 + (ad - bc)^2$.

**Chinese Remainder Theorem**: If $(m_1, m_2) = 1$, for any $a, b$, there is an $n$ such that $n = a \pmod{m_1}$ and $n = b \pmod{m_2}$. Further, if $n'$ is another such number, $n = n' \pmod{m_1 m_2}$.

**Solving Linear Equations over $\mathbb{Z}$:** $ax = b \pmod{m}$ has a solution iff $(a, m) \mid b$. If so there are $\frac{m}{(a,m)}$ solutions. If $(m_1, m_2) = 1$ then $\phi(m_1 m_2) = \phi(m_1)\phi(m_2)$. If $N_f(m)$ is the number of solutions of $f(x) = 0 \pmod{m}$ and $(m_1, m_2) = 1$ then $N(m_1 m_2) = N(m_1)N(m_2)$.

If $R = R_1 \times R_2 \times \ldots \times R_n$, $U(R) = U(R_1) \times U(R_2) \times \ldots \times U(R_n)$; consequence: if $(m_i, m_j) = 1$ and $m = m_1 m_2 \ldots m_n$ then $\mathbb{Z}/(m) = \mathbb{Z}/(m_1) \times \mathbb{Z}/(m_2) \times \ldots \times \mathbb{Z}/(m_n)$. Applying this to $n = 2^{e_0} {p_1}^{e_1} {p_2}^{e_2} \ldots {p_n}^{e_n}$ we find $n$ has a primitive root iff $n = 2, 4, p^e$. $p$ has $\phi(p-1)$ primitive roots. Lucas: If $(a, m) = 1$ and $a^{p-1} = 1 \pmod{m}$ and $p-1$ is the smallest such exponent then $m$ is prime.

Note that solutions of $f(x) = 0 \pmod{p}$ are solutions of $(f(x), x^p - x)$. **Chevalley:** Suppose $f, g \in \mathbb{Z}_p[x_1, \ldots, x_n]$ of degree $r, r < n$ then (1) if $f(x) = 0 \pmod{p}$ has a solution, it has at least two; and (2) if $g$ is homogeneous, it has at least one non-trivial solution.

**Hensel:** Suppose $f(x) \in \mathbb{Z}[x]$. If $f(a) = 0 \pmod{p^j}$ and $f'(a) \neq 0 \pmod{p^j}$ there is a unique $t : f(a + tp^j) = 0 \pmod{p^{j+1}}$. If $deg(f(x)) = n$ with leading coefficient 1 then $f(x)$ has $n$ solutions iff $f(x) \mid (x^p - x)$. If $d \mid (p-1)$ then $x^d = 1 \pmod{p}$ has $d$ solutions.

$\sum_{d|n} \phi(d) = n$. $\phi(n) = n \prod_{p|n} (1 - \frac{1}{p})$. $\phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}$. $0 = \sum_{d|n} \mu(d)$. **Moebius:** $F(n) = \sum_{d|n} f(d) \rightarrow f(n) = \sum_{d|n} \mu(d) F(\frac{n}{d})$. Proof: $\sum_{d|n} \mu(d) F(\frac{n}{d}) \sum_{d|n} \mu(d) \sum_{\delta | \frac{n}{d}} f(\delta) = \sum_{\delta|n} \sum_{d|\frac{n}{\delta}} \mu(d) f(\delta) = \sum_{\delta|n} f(\delta) \sum_{d|\frac{n}{\delta}} \mu(\delta) = f(n)$. If $(x, n) = 1$ then $x^{\phi(n)} = 1 \pmod{n}$. Counterexample to converse: (First **Carmichael Number**): 561.

$(\frac{a}{p}) = a^{\frac{p-1}{2}}$. **Quadratic Reciprocity:** If $p, q$ are odd primes, $(\frac{p}{q})(\frac{q}{p}) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$, $(\frac{2}{p}) = (-1)^{\frac{p^2-1}{8}}$.

Gauss's first proof of the law of quadratic reciprocity: $(\frac{p}{q})(\frac{q}{p}) = (-1)^{\frac{(p-1)(q-1)}{4}}$. Proof: Let $Dx = g_x p + r_x$. Set $\rho_x = r_x$, if $r_x < \frac{p}{2}$, $\rho_x = p - r_x$, if $r_x > \frac{p}{2}$. Let $n$ be the number of $\rho_x$ that are less than 0. Multiply $D, 2D, 3D \ldots \frac{p-1}{2} D$ together, this yields: $D^{\frac{p-1}{2}} \frac{p-1}{2}! = (-1)^n \frac{p-1}{2}!$ or $D^{\frac{p-1}{2}} = (\frac{D}{p}) = (-1)^n$. Let $D = q \neq p$ then either $x = \rho_x + g_x \pmod{2}$ or $x = \rho_x + g_x + 1 \pmod{2}$. $\sum_x^{\frac{p-1}{2}} x = n + \sum_x^{\frac{p-1}{2}} \rho_x + \sum_x^{\frac{p-1}{2}} g_x \pmod{2}$. So $n = \sum_x^{\frac{p-1}{2}} g_x \pmod{2}$. But $g_x = \lfloor \frac{qx}{p} \rfloor$. So $(\frac{q}{p}) = (-1)^{\sum g_x} = (-1)^{\sum \lfloor \frac{qx}{p} \rfloor}$ and $(\frac{p}{q})(\frac{q}{p}) = (-1)^{\sum \lfloor \frac{qx}{p} \rfloor + \sum \lfloor \frac{px}{q} \rfloor}$. Now use the fact that $\sum_{i=1}^{\frac{p-1}{2}} \lfloor \frac{iq}{p} \rfloor + \sum_{i=1}^{\frac{q-1}{2}} \lfloor \frac{ip}{q} \rfloor = \frac{(p-1)(q-1)}{2}$. This can be derived by looking at the number of lattice points in a $\frac{p-1}{2} \times \frac{q-1}{2}$ rectangle with vertex at $(0, 0)$.

Another Proof of QR using **Gauss Sums:** $g_a(\zeta) = \sum_{t=0}^{p-1} \zeta(t) \varsigma^{at}$. Set $g(x) = g_1(x)$. Number of solutions to $x^2 = t \pmod{p}$ is $1 + (\frac{t}{p})$. $g_a(\zeta) = \zeta(a^{-1}) g(\zeta)$ if $a \neq 0 \pmod{p}$ otherwise it's 0. $\sum (\frac{t}{p}) \varsigma^{at} = (\frac{a}{p}) \sum (\frac{t}{p}) \varsigma^t$. If $\zeta$ is the principal character, $g(\zeta) = \sqrt{p}$. If $\zeta$ is real and $g^k(\zeta) = (g(\zeta))^k$ then $g^2(\zeta) = (-1)^{\frac{p-1}{2}} p$.
Look at $|g(\zeta)|^2 = T = \sum_a g_a(\zeta) \overline{g}_a(\zeta)$. On one hand, it's $\sum_t (\frac{t}{p})(\frac{-t}{p}) g^2 = (\frac{-1}{p})(p-1) g^2$. On the other, it's $\sum_x \sum_y \sum_a g_a(\zeta(x)) g_{-a}(\zeta(y)) = \sum_a \sum_x \sum_y (\zeta(xy)) \varsigma^{(x-y)a} = (p-1)p$.
Proof or QR: Set $p^* = (-1)^{\frac{p-1}{2}} p$. $g^{q-1} = (g^2)^{\frac{q-1}{2}} = (\frac{p^*}{q})$. So $g^q = (\frac{p^*}{q}) g$. On the other hand, $g^q = (\sum_t (\frac{t}{p}) \varsigma^t)^q \pmod{q} = (\sum_t (\frac{t}{p})^q \varsigma^{qt}) \pmod{q} = (\frac{q}{p}) g$. So $(\frac{p^*}{q}) = (\frac{q}{p})$.

$b^n + 1$ is prime only if $n$ is a power of 2. If $M_p = 2^p - 1$ is prime, $\Delta_M = \frac{1}{2} M(M+1)$ is perfect. **Beatty:** If $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ and $A = \{\lfloor m\alpha \rfloor\}$, $B = \{\lfloor m\beta \rfloor\}$ then $A \cup B = Z$ and $A \cap B = \emptyset$.

There are no solutions to $x^2 + y^2 = n$ if $n = 3 \pmod{4}$. There are solutions to $x^2 + y^2 = n$ if $n = 1 \pmod{4}$. If $a$ has $A$ divisors $a_1, \ldots, a_A$ with $a_i = 1 \pmod{4}$ and $B$ divisors $b_1, \ldots, b_B$ with $b_i = 3 \pmod{4}$ then $x^2 + y^2 = n$ has $4(A - B)$ solutions in the integers.

**Pell's Equation:** $x^2 - dy^2 = 1$ is solvable (if $d$ is not a perfect square) using continued fractions. Let $\frac{p}{q} < \frac{r}{s}$ be two rationals such that $ps - rq = -1$ then $\forall \lambda, \mu$, $\frac{p}{q} \leq \frac{\lambda p + \mu r}{\lambda q + \mu s} \leq \frac{r}{s}$. Let $\frac{p}{q} \leq \frac{a}{b} \leq \frac{r}{s}$ with $ps - rq = -1$ then $a = \lambda p + \mu r$ and $b = \lambda q + \mu s$.

There are infinitely many primes of the form $4n + 3$. Largest power of $p$ dividing $n!$ is $\sum_{l \geq 0} \lfloor \frac{n}{p^l} \rfloor$.

**Erdos' proof of Bertrand:** (1) prove for $n < 4000$, (2) $\Pi_{p \leq x} p \leq 4^{x-1}$, (3) $\binom{2n}{n}$ contains $p$ $\sum_{k \geq 1} \lfloor \frac{2n}{p^k} \rfloor - 2 \lfloor \frac{n}{p^k} \rfloor$, (4) $\frac{4^n}{2n} \leq \binom{2n}{n} \leq (\Pi_{p \leq \sqrt{2n}} p)(\Pi_{\sqrt{2n} < p \leq \frac{2n}{3}} p)(\Pi_{\frac{2n}{3} < p \leq 2n} p)$, (5) $4^n \leq (2n)^{1 + \sqrt{2n}} 4^{\frac{2}{3} n}$, so $n < 4000$.

**Dirichlet:** If $a > 0$ and $(a, n) = 1$, then there are infinitely many primes $p$, such that $p = a \pmod{n}$.

The following moduli have primitive roots for $p > 2$, $2, 4, p^k, 2p^k$. Fact for **Miller-Rabin:** $n - 1 = 2^s r$, $r \neq 0 \pmod{2}$, $(a, n) = 1$. If $n$ is prime, either $a^r = 1 \pmod{n}$ or $a^{2^j r} = -1 \pmod{n}$ for some $j : 0 \leq j \leq (s-1)$.

$\zeta(s) = \sum \frac{1}{n^s}$ which converges for $Re(s) > 1$. Note: $\zeta(2) = \frac{\pi^2}{6}$. **Riemann Hypothesis:** If $s = a + bi$, all the zeros of $\zeta(s)$ have $a = \frac{1}{2}$.

**Prime Number Theorem:** Let $\Pi(x)$ be the number of primes $\leq x$. $\Pi(x) \approx (\frac{x}{ln(x)})$.

**Euler:** $\sum_{y<n\leq x} f(n) = \int_y^x f(t)dt + \int_y^x (t - \lfloor t \rfloor) f'(t)dt + (x - \lfloor x \rfloor) f(x) - (y - \lfloor y \rfloor) f(y)$.

$\sum_{n \leq x} \frac{1}{n} = ln(x) + C + O(\frac{1}{x})$. $\sum_n \frac{\mu(n)}{n^2} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2}$.

**Dirichlet:** Let $\alpha$ be a real number and $Q$ a positive integer. There is a rational number $\frac{p}{q}$ with $1 \leq q \leq Q$ such that $|\alpha - \frac{p}{q}| \leq \frac{1}{qQ}$. Proof: Let $B_q = \{\frac{q-1}{Q} \leq x < \frac{q}{Q}\}$. Let $c_q = q\alpha - \lfloor q\alpha \rfloor$. By the pigeon hole principle, at least 2 $c_q$'s must lie in a single $B_k$. This completes the proof. It's easy to extend this to show that if $\alpha$ is irrational, there are infinitely many rational numbers $\frac{p}{q}$ such that $|\alpha - \frac{p}{q}| \leq \frac{1}{q^2}$, which was sharpened by Hurwitz. **Hurwitz:** If $\alpha$ is irrational, there are infinitely many rational numbers $\frac{p}{q}$ such that $|\alpha - \frac{p}{q}| \leq \frac{1}{\sqrt{5}q^2}$. **Liouville:** Let $\alpha$ be an algebraic number of degree $d \geq 2$. There is a constant $c(\alpha) > 0$ such that for all $\frac{p}{q}$, $|\alpha - \frac{p}{q}| > \frac{c(\alpha)}{q^d}$. **Roth:** Let $\alpha$ be an algebraic number of degree $d \geq 2$ and $\epsilon > 0$. There is a constant $c(\alpha, \epsilon) > 0$ such that for all $\frac{p}{q}$, $|\alpha - \frac{p}{q}| > \frac{c(\alpha,\epsilon)}{q^{2+\epsilon}}$. Consequence: $z = \sum_i^\infty 10^{-i!}$ is transcendental.

$(2^a - 1, 2^b - 1) = 2^{(a,b)} - 1$. Proof: Let $a = bq + r$, $x^a - 1 = (x^b - 1)(x^{a-b} + x^{a-2b} + \ldots + x^{a-qb}) + x^r - 1$. This parallels the construction of $(a, b)$ in the Euclidean algorithm.

***p*-adic valuation:** $x = p^k \frac{a}{b}, (a, b) = (a, p) = (b, p) = 1$ then $\nu_p(x) = k$. If $f(x, y, z)$ over $\mathbb{Z}$ is quadratic, then $f$ has a solution over $\mathbb{Z}$ iff it has a solution in the $p$-adics over for all $p$. Counterexample for higher order equations: $3x^3 + 4y^3 + 5z^3 = 0 \pmod{p}$ is solvable for $p$ but $3x^3 + 4y^3 + 5z^3 = 0$ has no solutions.

### 1.1.2 Inequalities

**Arithmetic-Geometric:** $\frac{1}{n}\sum_n a_i \geq (\prod_n a_i)^{\frac{1}{n}}$. Proof: $A_n = \sum_n a_i$, $G_n = (a_1 a_2 \ldots a_n)^{\frac{1}{n}}$. Put (1) $A = \frac{x_{n+1} + (k-1)A_{n+1}}{n}$ and (2) $A_n + A = 2A_{n+1}$. Now apply induction to equation (1) and (2).

**Triangle:** $|x| + |y| \geq |x + y|$. **Simple Cauchy:** $(a^2 + b^2)(c^2 + d^2) \geq (ac + bd)^2$ with equality iff $bc - ad = 0$. **Cauchy-Schwartz:** $|u \cdot v| \leq ||u||||v||$. Proof: Look at $\sum(a_i x + b_i)^2$. Get $(\sum {a_i}^2)x^2 + 2(\sum a_i b_i)x + \sum {b_i}^2$. Complete square. Constant is always $\geq 0$.

**Holder:** If $\frac{1}{p} + \frac{1}{q} = 1$ then $\frac{a^p}{p} + \frac{b^q}{q} \geq ab$ and $(\sum_i {a_i}^p)^{\frac{1}{p}} \cdot (\sum_i {b_i}^q)^{\frac{1}{q}} \geq \sum_i a_i b_i$. Proof: If $f$ is monotonically increasing, $f(0) = 0$, then $\int_0^a f + \int_0^b f^{-1} \geq ab$. Another proof: You can prove first part using Arithmetic-Geometric inequality. Apply this inequality repeatedly with $a = \frac{a_i}{(\sum_{i=1}^n {a_i}^p)^{\frac{1}{p}}}$ and $b = \frac{b_i}{(\sum_{i=1}^n {b_i}^q)^{\frac{1}{q}}}$. Adding these we get $(\sum_{i=1}^n {a_i}^p)^{\frac{1}{p}} (\sum_{i=1}^n {b_i}^q)^{\frac{1}{q}} \geq \sum_{i=1}^n a_i b_i$.

**Minkowski:** $(\sum {a_i}^p)^{\frac{1}{p}} + (\sum {b_i}^p)^{\frac{1}{p}} \geq (\sum (a_i + b_i)^p)^{\frac{1}{p}}$. Proof: Write $(x_1 + x_2)^p + (y_1 + y_2)^p = [(x_1 + x_2)^{p-1} x_1 + (y_1 + y_2)^{p-1} y_1] + [(x_1 + x_2)^{p-1} x_2 + (y_1 + y_2)^{p-1} y_2]$. Apply Holder to each term to get $(x_1^p + y_1^p)^{\frac{1}{p}} [(x_1 + x_2)^{(p-1)q} + (y_1 + y_2)^{(p-1)q}]^{\frac{1}{q}} \geq x_1(x_1 + x_2)^{p-1} + y_1(y_1 + y_2)^{p-1}$ and $(x_2^p + y_2^p)^{\frac{1}{p}} [(x_1 + x_2)^{(p-1)q} + (y_1 + y_2)^{(p-1)q}]^{\frac{1}{q}} \geq x_2(x_1 + x_2)^{p-1} + y_2(y_1 + y_2)^{p-1}$. Since $\frac{1}{p} + \frac{1}{q} = 1$, $(p-1)q = p$. Adding the two inequalities and dividing by $[(x_1^p + x_2^p) + (y_1^p + y_2^p)]^{\frac{1}{q}}$ while noting that $1 - \frac{1}{q} = \frac{1}{p}$, we get Minkowski.

**Chebyshev:** If $a_1 \leq a_2 \ldots a_n$, $b_1 \leq b_2 \ldots b_n$. $(\frac{1}{n}\sum a_i)(\frac{1}{n}\sum b_i) \leq (\frac{1}{n}\sum a_i b_i)$. Proof: $\sum_{i,j}(a_i b_i - a_i b_j) = n\sum_i a_i b_i - (\sum_i a_i)(\sum_i b_i)$ $\sum_{i,j}(a_j b_j - a_j b_i) = n\sum_i a_i b_i - (\sum_i a_i)(\sum_i b_i)$ so $n\sum_i a_i b_i - (\sum_i a_i)(\sum_i b_i) = \frac{1}{2}\sum(a_j - a_i)(b_i - b_j) \leq 0$.

$\sum_i a_i b_i$ is max when $a_i$ and $b_i$ are in order, $a_, b_i \geq 0$. $min(a,b) \leq \frac{2ab}{a+b} \leq \sqrt{a}b \leq \frac{a+b}{2} \leq \sqrt{\frac{a^2+b^2}{2}} \leq max(a,b)$. **Concave (convex cap):** $f(tx+(1-t)y) \geq tf(x)+(1-t)f(y)$. $f$ is concave if $f''(x) < 0$. $log$ is concave. **Convex (convex cup):** $f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$. $f$ is convex if $f''(x) > 0$. $x^2, x > 0$ is concave. **Concave Jensen:** $E(f(X)) \leq f(E(X))$. **Convex Jensen:** $E(f(X)) \geq f(E(X))$.

$log(x) \leq (x-1)$, equality iff $x = 1$. **Hadamard inequality:** $|D(a_1, a_2, a_3, \ldots, a_n)| \leq ||a_1|| \cdot ||a_2|| \ldots ||a_n||$. $a^2+b^2+c^2 \geq ab+ac+bc$ and $\frac{b}{a+c}+\frac{a}{b+c}+\frac{c}{b+c} \geq \frac{3}{2}$. **Weighted AM-GM:** If $\lambda_1, \ldots, \lambda_n > 0$ and $\sum_{i=1}^n \lambda_i = 1$, then $\sum_{i=1}^n \lambda_i x_i \geq \prod_{i=1}^n x_i^{\lambda_i}$.

### 1.1.3 Combinatorics and Sets

Let $f(x) = c_k x^k + \ldots + c_0$ be a polynomial with $c_0 c_k \neq 0$ which factors as $f(x) = c_k(x-r_1)^{m_1} \ldots (x-r_l)^{m_l}$, then a sequence $\{a_n\}$ satisfies a **linear recurrence** with characteristic polynomial $f(x)$ iff $\exists : g_1(x), \ldots, g_l(x)$ such that $a_n = g_1(n){r_1}^n + \ldots + g_l(n){r_l}^n$ where $deg(g_i) < m_i$.

**Linear congruential generator:** $x_{n+1} = (ax_n + c) \pmod{m}$ has period $n$ if $(c,m) = 1$. $b = a-1$, $b = 0(p)$ if $p|m, b = 0(4)$ if $m = 0(4)$.

**Burnside:** Let a permutation group $G$ act on $A$ inducing an equivalence relation $S$. Let $n$ be the number of equivalence classes. $n = \frac{1}{|G|}\sum_{g \in G} |A_g|$. Let $D$ be a set of elements operated on by $G$ and $R$ be a set of colors. A coloring is a map $f: D \rightarrow R$. Two colorings, $f_1, f_2$, are equivalent if $f_1(d) = f_2(d^g), \forall d$. Let $cyc(\pi)$ be the number of cycles in $\pi$ and $c \in C(D,R)$ be a coloring. To use Burnside to count colorings, show that $|{D^R}_\pi| = |R|^{cyc(\pi)}$. **Polya:** $P_G(x_1, x_2, \ldots x_n) = \frac{1}{|G|}\sum_{g \in G} {x_1}^{\pi_1(g)} {x_2}^{\pi_2(g)} \ldots {x_n}^{\pi_n(g)}$. Example (Vertices on a cube): $P_G = \frac{1}{24}(x_1^8 + 9x_2^4 + 6x_4^2 + 8x_1^2x_3^2)$. Example (Faces on cube): $P_G = \frac{1}{24}(x_1^6 + 6x_1^2x_4 + 3x_1^2x_2^2 + 6x_2^3 + 8x_3^2)$. For $f \in R^D$, store: $\sum w(r)$, inventory: $W(f) = \prod_d f(d)$, pattern inventory of $R^D = \sum_f W(f)$. Polya: pattern inventory $= P_G(\sum w(r), \sum w(r)^2, \ldots \sum w(r)^n)$. Number of equivalence classes$=$ $P_G(|R|, |R|, \ldots |R|)$.

$(\mathbf{v}, \mathbf{k}, \mathbf{t}, \lambda)$ **design:** $|X| = v$, $B$ is a set of $k$ subsets of $X$ is a design if each $t$ subset $T$ of $X$, the number of blocks containing $T$ is $\lambda$ and $|B| = b$. $r$, the incidence number, is the number of blocks incident with one point. These designs are denoted $t-(v,k,\lambda)$ or $S_\lambda(t,k,v)$. $b_i = \lambda\frac{\binom{v-i}{t-i}}{\binom{k-i}{t-i}}$, $b_0 = b$, $b_1 = r$. $\frac{(vr)}{k} \leq \binom{v}{k}$.

**Hall's Theorem:** $J(A) = \{y \in Y, (x,y) \in E, x \in A\}$ and $|J(A)| \geq |A|$ if and only if there is a complete matching.

**Inclusion-Exclusion:** Let $A_1, A_2, \ldots, A_n$ be a family of subsets of $X$. The elements of X that are not in $\bigcup_i^n A_i$ is $\sum_{I \subseteq [n]}(-1)^{|I|}|A_I|$ where $A_I = \bigcap_{i \in I} A_i$. (Note: $A_\phi = X$.) For classical statement, let $A_i = \{x : c_i(x)\ is\ true\}$.

**Ramsay:** Let $P_r(S)$ be the $r$-subsets of $S$. Let $P_r(S) = A_1 \cup \ldots \cup A_t$ and $1 \leq r \leq q_1, \ldots q_t$. $\exists N(r, q_1, \ldots q_t)$ such that for $n \geq N$, S contains a $(q_i, A_i)$. $R(m,n) \leq R(m-1,n) + R(m,n-1)$ and $R(s,t) \leq \binom{(s+t-2)}{(s-1)}$.

**Generating Functions:** Let 12 objects be distributed to $A, B, C$ subject to: $A$ gets at least 4, $B$ and $C$ get at least 2 and $C$ gets no more that 5. The coefficient of $x^{12}$ in $(x^4 + \ldots x^8)(x^2 + \ldots x^8)(x^2 + \ldots x^5)$ is the number of ways this can happen. For selections with repetitions note that: $(\frac{1}{1-x})^n = \sum_i \binom{n+i-1}{i} x^i$. For partitions, examine $\frac{1}{1-x}(\frac{1}{1-x})^2 \ldots$. Exponential generating functions: $f(x) = a_0 + a_1 x + \frac{1}{2!}a_2 x^2 + \ldots \frac{1}{k!}a_k x^k + \ldots$. Difference calculus: $\sum_i i^n = (1+\Delta)^n u_0$.

**Dearrangements:** $n!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} \ldots + (-1)^n \frac{1}{n!})$. Menages ($i$ is not in $i+1 \pmod{n}$): $\sum_{r=0}^n (-1)^r (n-r)! \binom{(2n-r)}{r} \frac{2n}{2n-r}$. Number of solutions of $n_1 + n_2 + \ldots + n_r = r$ is $\binom{(n+r-1)}{r}$. Restricted permutation positions: $N(a_1', a_2', \ldots, a_{n-1}') = n! - \binom{n-1}{2}(n-2)! + \binom{n-1}{3}(n-3)! - \ldots + (-1)^n \binom{n-1}{n-1}(n-1)!$. For permutations of a, b, c, d, e, f which don't contain ace or fd: $N(a_1', a_2') = 6! - 4! - 5! + 3!$. Rook polynomials: $R(x,C) = xR(x,C_i) + R(x,C_e)$. Forbidden positions: $N(a_i', a_2', \ldots, a_n') = e_0 = n! - r_1(n-1)! + r_2(n-2)! - \ldots = \sum (-1)^j r_j (n-j)!$. Exactly $m$ with property: $e_m = \sum_{j=0}^n (-1)^j \binom{m+j}{j} s_{m+j}$.

Number of surjective maps from $[n] \rightarrow [k]$ is $\sum_{i=0}^k (-1)^i \binom{k}{i}(k-i)^n$. $n! = \sum_{i=0}^n (-1)^i \binom{n}{i}(n-i)^n$.

Multinomial coefficients: $\binom{a+b+c}{a,b,c}$ and $(x+y+z)^{a+b+c}$. $\binom{ne}{k} \leq (\frac{ne}{k})^k$, $\binom{n}{k} \geq (\frac{n}{k})^k$. Identities: $\binom{r}{k} = \frac{r}{k}\binom{r-1}{k-1}, \binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}, \binom{r}{k} = (-1)^k\binom{k-r-1}{k}, \binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}, \sum_{k=0}^{n}\binom{r+k}{k} = \binom{r+n+1}{n}, \sum_{k=0}^{n}\binom{k}{m} = \binom{n+1}{m+1}, \sum_{k=0}^{n}\binom{r}{k}\binom{s}{n-k} = \binom{r+s}{n}, \sum_{k=a}^{b-1} f(k) = \int_{k=a}^{b-1} f(x)dx + \sum_{k=1}^{m}\frac{B_k}{m!}f^{(k-1)}(x)_a^b + R_m, a_nT_n = b_nT_{n-1} + c_n \rightarrow s_na_nT_n = s_nb_nT_{n-1} + s_nc_n, s_nb_n = s_{n-1}a_{n-1}, R_n = s_na_nT_n, R_n = R_{n-1} + s_nc_n, \binom{-n}{r} = (-1)^r\binom{n+r-1}{r}, (1+x)^{-n} = 1 + \binom{-n}{1}x^{-1} + \ldots + \binom{-n}{n}x^{-n}$. $S(n,k)$ (Stirling numbers of the first kind) is the number permutations in $S_n$ with exactly $k$-cycles. $T(n,k)$ (Stirling numbers of the second kind) is the number of ways of grouping $n$ objects into $k$ groups. "Bell" numbers,($B_n$: the number of ways to divide $n$ things into groups. $B_{n+1} = \sum_{k=0}^{n}\binom{n}{k}B_n$. $\sum_{k=0}^{n} S(n,k) = b_n, S(n+1,k) = kS(n,k) + S(n,k-1)$ $\sum_{k=0}^{n} T(n,k) = n!, T(n+1,k) = nT(n,k) + T(n,k-1)$. Let $B_n$ denote the $n$th Bernoulli number. $\sum_{j=0}^{m}\binom{m+1}{j}B_j = 0$ and $B_0 = 1$. $\frac{x}{e^x-1} = \sum_{n=0}^{\infty} B_n\frac{x^n}{n!}$. Catalan numbers: $c_n = \frac{1}{n+1}\binom{2n}{n}, c_n = \sum_{k=0}^{n-1} c_kc_{n-k-1}$.

Let $p(n)$ be the number of partitions of $n$. Then, $p(n) \approx \frac{1}{4n\sqrt{3}}e^{\sqrt{\frac{2n}{3}}}$. The number of partitions of $n$ into $k$ things is the number of partitions of $n$ with largest partition $k$.

A sequence of $(n-1)(m-1)+1$ different numbers has either increasing sub-sequence of length $n$ or a decreasing sub-sequence of length $m$. Proof: Let $x \in B_r$ if the longest increasing sequence beginning with $x$ has length $n$. If any $B_r$, with $r \geq n$ is non empty, we're done. Otherwise, there must be a $B_k$ with $k < n$ containing at least $m$ elements. These $m$ elements form a decreasing sequence.

Similarly, if $1 \leq a_1, \ldots, a_n \leq m$ and $1 \leq b_1, \ldots, b_n \leq m$ , $\exists p, q, r, s$ with $a_{p+1} + \ldots + a_{p+q} = b_{r+1} + \ldots + b_{r+s}$. Proof: Let $j = j(k)$ be the smallest integer with $a_1 + \ldots + a_j \geq b_1 + \ldots + b_k$. Let $c_k = \sum_{i=1}^{j(k)} a_i - \sum_{i=1}^{k} b_i$. At least two $c_l$'s (say $c_u$ and $c_v$, $u > v$) are equal. $c_u - c_v$ provides the right sequence.

In permutation, $i < j$ and $a_i > a_j$ is **inversion**. Inversion table is $(b_j)$ where $b_j$ = number of elements left of $j$ that are $> j$. For 5 9 1 8 2 6 4 7 3, it's 2 3 6 4 0 2 2 1 0. Inversion table uniquely determines permutation. Inverse has same number of inversions.

Generating permutations of $\{1, 2, 3, \ldots, n\}$:

1. Set $\pi = 123\ldots n$. Output $\pi$.
2. If $\pi_i > \pi_{i+1}$, $\forall i$, stop.
3. Get largest $i$: $\pi_i < \pi_{i+1}$.
4. Find smallest $j$: $i < j$ such that $\pi_i < \pi_j$.
5. $\pi_i \leftrightarrow \pi_j$.
6. Reverse the order of the numbers following, $\pi_j$, denote this by $\pi$. Output $\pi$. Go to 2.

Another algorithm: Steinhaus weaving generator (by recursion).

**Permanent:** $per(a_{ij})$, $m \times n$ matrix, is $\sum_{\sigma} a_{1i_1}a_{2i_2}\ldots a_{mi_m}$ where $\sigma$ runs through $m$ permutations of $[n]$. $n! = per(J) = \sum_{r=0}^{n-1}\binom{n}{r}(-1)^r(n-r)^n$.

Let $A_r$ be the matrix obtained by replacing $r$ specified columns of $A$ by 0. Let $S(A_r)$ be the product of row sums of $A_r$. Let $\sum_r S(A_r)$ over all choices of r: $per(A) = \sum S(A_{n-m}) - \binom{n-m+1}{1}S(A_{n-m+1}) + \ldots(-1)^{m-1}\binom{n-1}{m-1}S(A_{m-1})$.

$\mathcal{G}(V,E)$ a graph with vertex set $V$ and edge set $E$. $g(\mathcal{G})$ - girth - length of minimum cycle. $\omega(\mathcal{G})$- clique number. $\alpha(\mathcal{G}) = \omega(\overline{\mathcal{G}})$- independence number. $\chi(\mathcal{G})$ - chromatic number. $\delta(\mathcal{G})$ - minimum degree. $\Delta(\mathcal{G})$ - maximum degree. $d(x,y)$ = number of edges between x and y. $D_{\mathcal{G}}(x,y) = max_{x,y}d(x,y)$.

A graph is **bipartite** iff it contains no cycles of odd length. Theorem: $\alpha(\mathcal{G})\chi(\mathcal{G}) \geq n$. Cayley graph. Strongly regular graphs. Expander graphs and short paths (Todo).

There are $n^{n-2}$ labeled trees with $n$ nodes. Proof: Use **Prufer Code** for tree $T$: remove leaf with smallest label, add the label of the vertex it's connected to at end of sequence.

$G(n,M), N = \binom{n}{2}$. Random graph selecting $M$ of the $N$ edges. $Pr[G = H] = p^{e(H)}q^{N-e(H)}$. $X_s(G) =$ number of complete graphs of order $s$. $E(X_s) = \sum_{\alpha \in S} E(Y_\alpha(G)$, where $Y_\alpha(G) = 1$, if $G[\alpha] = K_\alpha$, 0 otherwise. $E_M(Y_\alpha) = P_M(G_p[\alpha] = K_\alpha) = p^S = \binom{N-S}{M-S}\binom{N}{M}^{-1}$. $E_p(X_s) = \binom{n}{s}p^s$. If $a$ is the order of the automorphism group of $F$ then $K_k$ has $\frac{k!}{a}$ subgraphs isomorphic to $F$. $N_F = \binom{n}{k}\frac{k!}{a} = \frac{(n)_k}{a}$. For cycles, $a = 2k$.

**Erdos:** There is a graph, G, with $g(G) \geq n$ and $\chi(G) \geq n$. Another formulation: Given natural numbers $g \geq 3$, $k \geq 2$, $\exists G$, with $|G|k^{3g}$, $g(G) \geq g$ and $\chi(G) \geq k$.

Fact 1: If $G \in G(n,p)$, $q = 1-p$ then $Pr[\alpha(G) \geq k] \leq \binom{n}{k}q^{\binom{k}{2}}$ Fact 2: Markov's inequality. Fact 3: Let X be a r.v. representing the number of $k$-cycles. $E(X) = \frac{(n)_k}{2k}p^k$. Fact 4: If $k > 3$ and $p(n)$ is a function with $p(n) \geq \frac{6kln(n)}{n}$ then $lim_{n\to\infty}Pr[\alpha \geq \frac{n}{2k} = 0$: $\binom{n}{r}q^{\binom{r}{2}} \leq n^n q^{\binom{r}{2}} \leq (ne^{-p\frac{r-1}{2}})^r$ inside expression is $\leq \sqrt{\frac{e}{n}} \to 0$. Argument: Fix $0 < \epsilon < \frac{1}{k}$, $p = n^{1-\epsilon}$, $X(G)$ is the number of cycles $\leq k$. $E(X) \leq \sum \frac{(n)_i}{2i}p^i \leq \frac{1}{2}(k-2)(np)^k$. $Pr[X \geq \frac{n}{2}] = \frac{E(X)}{\frac{n}{2}} \leq (k-2)n^{k\epsilon-1}$. Pick $n$ big enough so that $Pr[X \geq \frac{n}{2}] > \frac{1}{2}$ and $Pr[\alpha \geq \frac{n}{2k}] < \frac{1}{2}$. So $\exists G$ with $< \frac{n}{2}$ short cycles and $\alpha(G) < \frac{n}{2k}$ delete up to $\frac{n}{2}$ points to eliminate the short cycles producing a graph $H \subseteq G$. $\chi(H) \geq \frac{H}{\alpha(H)} \geq \frac{\frac{n}{2}}{\alpha(G)} > k$.

$\epsilon$-regular: $(A,B)$ with $X \subseteq A$ and $Y \subseteq B$ such that $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$ satisfy $|d(X,Y) - d(A,B)| \leq \epsilon$. $\epsilon$ regular partition: (1) $|V_0| < \epsilon|V|$, (2) $|V_i| = |V_1|$, for $i \geq 1$, (3) all but $\epsilon k^2$ of the pairs $(V_i, V_j)$ are $\epsilon$ regular. Szemeredi Regularity Lemma: For every $\epsilon > 0$ and every $m \geq 0$, $\exists M$ such that every graph of order at least m admits an $\epsilon$ regular partition $\{V_0, V_1, \ldots, V_k\}$ with $m \leq k \leq M$.

**Giant component** in $G(n,p)$ when $p = \frac{1+\epsilon}{n}$. Sunflower Lemma: Let $T = \{S_1, S_2, \ldots, S_k\}$ be a system over a set $U$, such that (1) $|S_i| \leq l$ and (2) $k > (p-1)^l l!$. Then $\exists F \subseteq T$, $F = \{S_{i_1}, S_{i_2}, \ldots, S_{i_p}\}$ such that $\forall A, B \in F, A \cap B = F$.

**Random function statistics:** Tail, cycle, predecessor length: $\sqrt{\frac{\pi n}{8}}$, Tree Size: $\frac{n}{3}$, Number of components: $\frac{lg(n)}{2}$, Component Size: $\frac{2n}{3}$.

**Sperner:** A collection $F$ of non-empty subsets of a set $X$ is called an antichain if no set in $F$ is properly contained in another set of $F$. If $|X| = n$, $|F| \leq \binom{n}{n'}$, where $n' = \lfloor \frac{n+1}{2} \rfloor$. If $|X|$ is even there are exactly 2 maximal antichains, the collection of $\lfloor \frac{n-1}{2} \rfloor$ subsets of $X$ and the collection of $\lfloor \frac{n+1}{2} \rfloor$ subsets of $X$. If $n$ is even, there is exactly one maximal antichain, namely, the collection of $\lfloor \frac{n}{2} \rfloor$ subsets of $X$.

Posets, chains (totally ordered subset) and antichains (set in which all subsets are incomparable). **Dilworth:** The cardinality of a maximal antichain is equal to the minimum number of disjoint chains into which a poset can be partitioned. In a chain of $mn+1$ elements there is a chain of $m+1$ elements or there are $n+1$ incomparable elements. **Rubik group:** $|G_R| = 2^{27}3^{14}5^3 7^2 11$.

If $f(x) \in \mathbb{Z} \to x \in \mathbb{Z}$ then $\lfloor f(x) \rfloor = \lfloor f(\lfloor x \rfloor) \rfloor$. $\lfloor \frac{x+m}{n} \rfloor = \lfloor \frac{\lfloor x \rfloor + m}{n} \rfloor$. $\sum_{i=0}^{m-1} \lceil \frac{n-i}{m} \rceil = n$. $\sum_{k=0}^{m-1} \lfloor \frac{nk+x}{m} \rfloor = \sum_{k=0}^{n-1} \lfloor \frac{mk+x}{n} \rfloor$.

The following are equivalent: (1) [**Axiom of choice**] If $I \neq \emptyset$ and $\forall i \in I, A_i \neq \emptyset$ then $\prod_{i \in I} A_i \neq \emptyset$; (2) [**Zorn's Lemma**] If $A \neq \emptyset$ is partially ordered and if every chain (including infinite chains!) has an upper bound in $A$ then $A$ contains a maximal element; (3) [**Well ordering**] If $A \neq \emptyset$ has a linear order, $\leq$, then $(A, \leq)$ is has a least element. Transfinite Induction: if $B \subseteq A$ and $A$ is well ordered under $\leq$ and if $\{c \in A : c < a\} \subseteq B \to a \in B$ then $A = B$.

$|P(A)| > |A|$. Proof: $f : a \mapsto \{a\}$ shows $|P(A)| \geq |A|$. Suppose $|P(A)| = |A|$, then there is a bijection $f$ between $P(A)$ and $A$. Let $B = \{a : a \notin f(a)\}$. If $b \in B$ and $b \mapsto f(b)$ then $b \notin B$, this is a contradiction.

**Schroeder-Bernstein:** If $A, B$ are two sets and there are injections $f : A \rightarrow B$ and $g : B \rightarrow A$ then there is a bijection $h : A \rightarrow B$. Lemma: If there is a subset $A' \subseteq A$ satisfying the hypothesis of the theorem with $A' = B$ then there is a bijection $h : A \rightarrow A'$. The Lemma implies the theorem: Let $A' = g(f(A))$ then by the lemma, $\exists h : A \rightarrow A'$ and $g^{-1} \circ h$ is the desired bijection. Proof of Lemma: Set $X = \bigcap_{n \geq 0} f^{(n)}(A \setminus A')$ and define $h(x) = f(x), x \in X, h(x) = x, x \notin X$; this is a bijection. First note $f(X) \subseteq X$. If $x, y \in X$ or $x, y \notin X$ it is clear that $h(x) = h(y) \rightarrow x = y$ and by construction, there is no $x \in X, y \notin X$ with $h(x) = h(y)$. If $y \in A'$ and $y \in X$, then $y \in f^{(n)}(A \setminus A')$ for some $n$ in which case $\exists x \in X : h(x) = y$ otherwise $y \notin X$ and $h(y) = y$.

## 1.2 Algebra

### 1.2.1 General Algebra

For $x^2+px+q=0$, $(x_1-x_2)=\sqrt{D}, D=p^2-4q, x_1+x_2=p$. For $ax^3+bx^2+cx+d=f(x)$, substitute $y=x+\frac{b}{3a}$ and divide by $a$ to get $x^3+nx+p$. Put $y=(u+v)$, get $p=3uv$, $q=u^3+v^3$. Note $S_3 \supseteq A_1 \supseteq 1$, $(x_1-x_2)(x_1-x_3)(x_2-x_3)=\sqrt{D}, D=-4p^3-27q^2$. Solution is: $y=(-\frac{p}{2}+-\sqrt{\frac{p^2}{4}+\frac{n^3}{27}})^{\frac{1}{3}}$. For $ax^4+bx^3+cx^2+dx+e=f(x)$, substitute $y=x+\frac{b}{4a}$ and divide by $a$ to get $x^4+px^2+qx+r$. Note $S_4 \supseteq A_4 \supseteq C_4 \supseteq Z_1 \supseteq 1$ and $\theta_1=(x_1+x_2)(x_3+x_4)$ is fixed by $C_4$ but not $A_4$. The $\theta_i$ are solutions of $\Theta^3-b_1\Theta^2+b_2\Theta_3-b_3$ with $b_1=2p, b_2=p^2-4r, b_3=-q^2$ and $D=16p^4r-4p^3q^2-128p^2r^2+144pq^2r-27q^4+256r^3$. Look at $(y^2+p)^2=py^2-qy-r$ and pick $z$ to make RHS $(y^2+p+z)^2=(p+2z)y^2-qy+(p^2-r+2pz+z^2)$ a perfect square.

**Fundamental Theorem of Algebra:** Let $f(z)=z^n+a_{n-1}x^{n-1}+...+a_0$ and $\mu=inf(|f(z)|)$. If $\mu=0$, we're done (min must occur in bounded ball). So assume $\mu \neq 0$. Let the minimum occur at $z_0$ and put $f(z_0)=w_0$, $w=f(z_0+\zeta)$. $\frac{w}{w_0}=1+q\zeta^\nu(1+\zeta\xi)=1-h\rho^\nu(1+\zeta\xi)$ where $\zeta=\rho(cos(\theta)+isin(\theta))$ and $q=h(cos(\lambda)+isin(\lambda))$. So we can find a point with smaller modulus than $w_0$. This contradicts the assumes minimality at $z_0$.

**Roots of Unity:** Consider $f(x)=x^h-1$ over $F$ where $(char(F),h)=1$ or $char(F)=0$. The roots of $f$ form an abelian group, $G$. $x \in G \rightarrow |x| \mid |G|$. Since $(f,f')=1$ there are $h$ distinct roots, set $h=\prod_{i=1}^m q_i^{v_i}$. $\{x : x^{h/q_i}=1\}$ is a group of order $h/q_i$ so $\forall i, \exists x_i \in G : x^{h/q_i} \neq 1$. Setting $b_i=x_i^{h/{q_i}^{v_i}}$, then $\chi=\prod b_i$ has order exactly $h$ and is a primitive $h$th root of unity. Let the number of such roots be $\varphi(h)$; if $(r,s)=1, \varphi(rs)=\varphi(r)\varphi(s)$ so $\varphi(\prod_i {q_i}^{v_i})=\prod_i \varphi({q_i}^{v_i})=\prod_i({q_i}^{v_i}-{q_i}^{v_i-1})=h\prod_i(1-\frac{1}{q_i})$. Set $n=\varphi(h)$ and $\Phi_n(x)=\prod_i(x-\psi_i)$ where $\psi_i$ are the primitive roots. $x^h-1=\prod_{d|h}\Phi_d(x)$ and by Moebius inversion, $\Phi_h(x)=\prod_{d|h}(x^d-1)^{\mu(\frac{h}{d})}$. $\Phi_h(x)$ is irreducible of degree $\varphi(h)$. Proof: Let $\zeta \in \mathbb{C}$ be a primitive root of $\Phi_h(x)$ with minimal polynomial $f(x)$ and $(p,h)=1$. Let $g(x)$ be the minimal polynomial for $\zeta^p$ so $g(\zeta^p)=0$. $x^h-1=f(x)g(x)h(x)$ and $g(x^p)=f(x)k(x)$. $g(x^p)=g(x)^p \pmod{p}$. If $\phi(x) \mid f(x)$ then $\phi(x) \mid g(x)^p \pmod{p}$. So $\phi(x)^2 \mid x^h-1$ but this contradicts the fact that $x^h-1$ does not have roots of multiplicity 2. It follows the if $(p_i,h)=1$, $\zeta^{p_1p_2\cdots p_k}$ is a primitive root and the degree of $f(x)$ is $\varphi(h)$. Note this shows that $Aut(\mathbb{Q}[\zeta]) \cong \mathbb{Z}_h^*$. It also allows us to calculate the Galois group if $h=q=p^n$ (its cyclic) and the subfields correspond to the cyclic subgroups of $\mathbb{Z}_q^*$. The $q$th roots of 1 are expressible as radicals if $char(F)=0$ or $char(F)>q$. If $N_p(d)=$ number of irreducible monic polynomials of degree $d$ in $GF(p)[x]$ then $p^n=\sum_{d|n} dN_p(d)$ and $N_p(d)=\frac{1}{n}\sum_{d|n}\mu(\frac{n}{d})p^d$. $x^{p^n}-x=\prod_{f irred,monic,deg(f)|n} f$.

**Eisenstein:** If $f(x)=\sum_{i=0}^n a_nx^n$, $a_n \neq 0 \pmod{p}$, $a_i=0 \pmod{p}, i<n$ and $a_0 \neq 0 \pmod{p^2}$ then $f$ is irreducible. **Factoring in finite number of steps:** Let $g(x) \in \mathbb{Z}[x]$ if $f(x) \mid g(x)$ then $f(n) \mid g(n)$ for all $n$. $deg(f)=s \leq \lfloor \frac{deg(g)}{2} \rfloor$. Pick $s$ integers $i_j$ and use the integer factors of $g(i_j)$ to get possible $g(i_j)$; there are a finite number of ways to pick the factors. For each possibility, we can solve for the $s$ coefficients of $f$.

### 1.2.2 Free Groups, Rings and Modules

Every group is the homomorphic image of a free group. If $G$ is a free abelian group generated by $n$ elements and $H$ is a subgroup of $G$ then $H$ is generated by $m \leq n$ elements.

Let $F_m$ be a free abelian group generated by $a_1, a_2, \ldots, a_m$ and define $E_i=r_{i1}a_1+r_{i2}a_2+\ldots+r_{im}a_m$ where $r_{ij} \in \mathbb{Z}$ and $1 \leq i \leq n$; further, put $b_i=E_i$ and let $K=<b_i>$. Suppose $G$ is the free abelian group generated by $a_i$ subject to $E_i=0$. Then $G \cong F_m/K$. Let $R$ represent the matrix $(r_{ij})$ then (1) if the matrix $S=(s_{ij})$ is obtained from $R$ by elementary row operations then $c_i=s_{i1}a_1+\ldots+s_{im}a_m \in K$; and, (2) if the matrix $S=(s_{ij})$ is obtained from $R$ by elementary column operations then $\exists a_i' \in F_m : b_i=s_{i1}a_1'+\ldots+s_{im}a_m'$ (so the $a_i'$ generate $K$). By applying elementary row and column operations we can transform $R$ into the diagonal matrix $D=diag(d_1,d_2,\ldots,d_r,0,\ldots,0)$ where $d_i \mid d_{i+1}$ and $G \cong \mathbb{Z}/(d_1)\times\mathbb{Z}/(d_2)\times\ldots\times\mathbb{Z}/(d_r)\times\mathbb{Z}\times\ldots\times\mathbb{Z}$ where there are $m-r$ copies of $\mathbb{Z}$ in the product.

Let $D$ be a UFD and $f(x)=a_0+a_1x+\ldots+a_nx^n, a_i \in D$. Let $K$ be the field of fractions. If $f_1(x)$

and $f_2(x)$ are primitive in $R[x]$ and are associates in $K[x]$ then they are associates in $D[x]$. Define $cont(f) = gcd(a_0, a_1, \ldots, a_n)$. **Gauss' Lemma**: If $D$ is a UFD and $f, g \in D[x]$ then $cont(f(x)g(x)) = cont(f(x))cont(g(x))$. If $f(x) \in R[x], deg(f) > 0$ and $f(x)$ is irreducible in $R[x]$ then it is irreducible in $K[x]$. Theorem: If $D$ is a UFD then $D[x]$ is a UFD. (Embed $D$ in its field of quotients and apply Gauss). Euclidean domains are principal ideal domains (PID) and all PIDs are UFDs.

**Ring theoretic CRT:** If $I_j, j = 1, 2, \ldots, n$ are ideals of $R$ and $I_j + I_k = R$ for $j \neq k$, then $\forall x_1, x_2, \ldots, x_n \in R, \exists x \in R$ such that $x = x_j \pmod{I_j}$. Corollary: Under the same assumptions, $\psi : R \rightarrow R/I_1 \times R/I_2 \times \ldots \times R/I_n$ given by $x \mapsto x \pmod{I_1} \times \ldots \times x \pmod{I_n}$ is surjective and $R/(\bigcap_{j=1}^n I_j) \cong R/I_1 \times R/I_2 \times \ldots \times R/I_n$. $\mathbb{Z}/(m\mathbb{Z}) \cong \prod_i \mathbb{Z}/({p_i}^{r_i}\mathbb{Z})$ and $\psi(m) = \prod_i \psi({p_i}^{r_i})$. If $R$ is cyclic of order $n$ then $End(R) \cong \mathbb{Z}/(n\mathbb{Z})$ and $(\mathbb{Z}/(n\mathbb{Z}))^* \cong Aut(R)$.

Groups with operators ($M$) and invariant subgroups. Projection commutes with all inner automorphisms; such an endomorphism is called normal. An $M-$group $G$ is decomposable iff there are projections. Any $M-$group satisfying DCC is a direct product of a finite number of indecomposable $M-$groups. If $\eta \in End(G)$ then $\sqrt{\eta} = \{z \in G : z\eta^s = 1\}$. **Fitting:** Let $G$ be an $M-$group that satisfies ACC and DCC and $\eta$ is a normal endomorphism of $G$ then $G = \sqrt{\eta} \times H$ and $H\eta = H$. If $G$ is an indecomposable $M-$group satisfying ACC and DCC then any normal $M-$endomorphism of $G$ is either **nilpotent** or an automorphism. Suppose $\eta_1, \eta_2$ are normal nilpotent $M-$endomorphisms, if $\eta_1 + \eta_2$ is an endomorphism it is nilpotent. **Krull-Schmidt** follows from this. Unitary: $RM = M$. **Hilbert Basis Theorem:** If $R$ is a ring with identity such that every ideal is finitely generated then $R[x]$ has the same property.

If $A, B$ are ideals, we say $A \mid B$ if $B \subseteq A$. $Q$ is **primary** iff $ab = 0 \pmod{I} \rightarrow a = 0 \pmod{Q}$ or $b \in \sqrt{I}$. If $Q$ is primary then $\sqrt{Q}$ is prime. Every irreducible ideal in a Noetherian ring is primary. Every ideal in a Noetherian ring is the finite intersection of primary ideals. If $Q_1, Q_2$ are primary and $\sqrt{Q_1} = \sqrt{Q_2}$ then $Q_1 \cap Q_2$ is primary. If $Q_1 \cap Q_2 \cap \ldots \cap Q_r = Q_1' \cap Q_2' \cap \ldots \cap Q_s'$ are two irredundant representations into primary ideals whose associated primes are distinct, then $r = s$ and the set of associated primes is identical. If $R^2 = R$ is a commutative ring then every maximal ideal is prime. Let $P$ be a prime ideal of $R$ ($1 \in R$) then (1) There is a 1-1 correspondence between the set of prime ideals of $R$ contained in $P$ and (2) the prime ideals of $R_P$ given by $\mathbb{Q} \mapsto \mathbb{Q}_P$. A local ring is a commutative ring with identity containing a unique maximal ideal. If $R$ is a commutative ring with identity, the following are equivalent: (1) $R$ is a local ring; (2) all non-units of $R$ are contained in an ideal $M \neq R$; (3) the non-units form an ideal. Substitution from a the polynomial ring to the ring of coefficients is a homomorphism.

If $R$ is Noetherian and $a \in M$ is $R-$**integral** iff $\exists$ a finitely generated submodule of $M$ that contains all powers of $a$. The totality $G$ of elements of $M$ that are $R-$integral is a subring of $M$ containing $R$. The ring $G$ or $R-$integral elements is integrally closed in $R$.

If $A, B, C, A', B', C'$ are modules over a ring $R$ with identity and we have the diagrams $0 \rightarrow A \xrightarrow{\mathrm{f}} B \xrightarrow{\mathrm{g}} C \rightarrow 0$ and $0 \rightarrow A' \xrightarrow{\mathrm{f}'} B' \xrightarrow{\mathrm{g}'} C' \rightarrow 0$ with $A \xrightarrow{\alpha} A'$, $B \xrightarrow{\beta} B'$, and $C \xrightarrow{\gamma} C'$, then (1) $\beta$ is a monomorphism if $\alpha$ and $\beta$ are and (2) $\beta$ is a epimorphism if $\alpha$ and $\beta$ are. $P$ is projective if given $A, B, g, f$ and morphism diagrams: $A \xrightarrow{\mathrm{g}} B \rightarrow 0$ and $P \xrightarrow{\mathrm{f}} B$, $\exists h, P \xrightarrow{\mathrm{h}} B$ which makes the diagram commute. $J$ is injective if given $A, B, g, f$ and morphism diagrams: $A \xrightarrow{\mathrm{g}} B \rightarrow 0$ and $A \xrightarrow{\mathrm{f}} J$, $\exists h, B \xrightarrow{\mathrm{h}} J$ which makes the diagram commute. Every free module $F$ over $R$ with identity is projective. If $R$ is a ring with identity, TFAE: (1) $P$ is projective, (2) every short exact sequence $0 \rightarrow A \xrightarrow{\mathrm{f}} B \xrightarrow{\mathrm{g}} P \rightarrow 0$ splits so $B = P \oplus A$ and (3) $\exists F$, free such that $F = K \oplus P$. If $R$ is a ring with identity, TFAE: (1) $J$ is injective, (2) every short exact sequence $0 \rightarrow A \xrightarrow{\mathrm{f}} B \xrightarrow{\mathrm{g}} C \rightarrow 0$ splits so $B = J \oplus C$ and (3) $J$ is a direct summand. $0 \rightarrow A \xrightarrow{\psi} B \xrightarrow{\phi} C$ is exact if: $0 \rightarrow Hom(D, A) \xrightarrow{\psi} Hom(D, B) \xrightarrow{\phi} Hom(D, C)$ is. $A \xrightarrow{\theta} B \xrightarrow{\zeta} C \rightarrow 0$ is exact if: $0 \rightarrow Hom(A, D) \xrightarrow{\psi} Hom(B, D) \xrightarrow{\phi} Hom(C, D)$ is. The full short exact sequence is split exact iff the corresponding dual (Hom) sequence is.

If $A$ us a unique factorization domain, $A$ is integrally closed. The integral closure in a number field $K$ is called the ring of algebraic integers. Algebraic integers form a free $\mathbb{Z}$-module of rank $[K : \mathbb{Q}]$.

Let $[E : F] = n$ and $[F(x) : F] = d$ and $x_1, x_2, \ldots, x_d$ be the roots of $min_F(x)$ then $N_{E/F}(x) = (\prod_{i=1}^d x_i)^{\frac{n}{d}}$ and $Tr_{E/F}(x) = \frac{n}{d}(\sum_{i=1}^d x_i)$. If $E/F$ is separable then $N_{E/F}(x) = \prod_{i=1}^n \sigma_i(x)$ and $Tr_{E/F}(x) =$

$\sum i = 1^n \sigma_i(x)$. If $F \subseteq E \subseteq K$ then $N_{E/F}(N_{K/E}(x)) = N_{K/F}(x)$ and $Tr_{E/F}(N_{K/E}(x)) = N_{K/F}(x)$. If $E/F$ is a finite separable extension, $\exists x \in E : Tr_{E/F}(x) = 0$ and $(x, y) \to Tr_{E/F}(xy)$ is bilinear.

For this paragraph, $L$ be a separable extension of $K$, $A \subseteq K$ be a ring of integers and $B \subseteq L$ be a ring of algebraic integers. $\vec{x}$ is a basis for $L/K$ iff $\Delta(\vec{x}) \neq 0$. If $L = K(x)$ and $f$ is a minimal polynomial of $x$ over $K$ then $\Delta(1, x, x^2, \ldots, x^{n-1}) = disc(f) = \prod_{i<j}(x_i - x_j) = (-1)^{\binom{n}{2}} N_{L/K}(f'(x))$. There is a basis for $L/K$ consisting of elements of $B$. If $A$ is a PID then $B$ is a free $A$-module of rank $[L : K]$. If $a_i \in A$ then $(x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$ is a maximal ideal.

Let $M$ be an $R$ module. The following are equivalent (1) $M$ satisfies **ACC (Noetherian)**, (2) Any non empty collection of submodules of $M$ has a maximal element. The following are equivalent (1) $M$ satisfies **DCC (Artinian)**, (2) Any non empty collection of submodules of $M$ has a minimal element. $M$ is Noetherian iff every submodule is finitely generated. $M$ is Artinian iff every submodule is finitely co-generated. Noetherian: PIDs, $F[x]$. $F[x_1, x_2, \ldots]$ is neither Noetherian nor Artinian. $IfN \subseteq M$ then $M$ is Noetherian iff $N$ and $M/N$ are. $M$ has a composition series iff $M$ is Noetherian and Artinian. $L$ be a separable extension of $K$, $A \subseteq K$ be a ring of integers $B \subseteq L$ be a ring of algebraic integers, if $A$ is integrally closed in $K$ and $A$ is Noetherian, so is $B$. Let $P$ be a prime ideal of $R$ and $P \supseteq I_1 I_2 \ldots I_n$ then $\exists k : P \supseteq I_k$. Let $I$ be a non-zero ideal of a noetherian integral domain $R$ then $I \supseteq P_1 P_2 \ldots P_n$ for $P_i$ prime. Let $R$ be a non-zero ideal of a noetherian integral domain and $K$ its field of quotients, $I$ is a fractional ideal if $I$ is an $R$-module and $\exists r \in R : rI \subseteq R$. If $I$ is a finitely generated $R$ submodule of $K$ then $I$ is a fractional ideal. If $R$ is Noetherian and $I$ is a fractional ideal of $R$ then $I$ is a finitely generated $R$ submodule of $K$.

A **Dedekind Domain ("DD")** is an integral domain, $R$, such that (1) $R$ is Noetherian, (2) $R$ is integrally closed, and, (3) Every non-zero prime ideal of $R$ is maximal. PIDs are DDs. Algebraic integers in a number field is a DD. If $P$ is a non-zero prime ideal in a DD, $R$ and $J = \{x \in K : xI \subseteq R\}$ then (1) $R \subseteq J$ and (2) $J$ is a fractional ideal and $PJ = R$. If $I$ is a fractional ideal in a DD, $R$ then $I = \prod_{i=1}^{N} {P_i}^{n_i}$ ($n_i \in \mathbb{Z}$ not just $\mathbb{Z}^{\geq 0}$), $n_P(I) = n_i$. The fractional ideals form a group. A non-zero fractional ideal is integral iff all $n_i$ in the forgoing representation are $\geq 0$. $I_1 \supset I_2$ iff $\forall P, n_P(I_1) \leq n_P(I_2)$. If $I_1, I_2$ are integral ideals then $I_1 \mid I_2$ if $I_2 = JI_1$. $I_1 \mid I_2$ iff $I_1 \supseteq I_2$. $L$ be a separable extension of $K$, $A \subseteq K$ be a ring of integers, if $A$ is a DD, $B$ is a DD.

If $0 \to A \xrightarrow{f} B \xrightarrow{g} P \to 0$, $B$ satisfies ACC (resp DCC) iff $A$ and $C$ do. $A$ satisfies ACC on submodules iff each submodule is finitely generated (same for rings). Jordan-Holder for modules (composition series have unique refinements). $A$ has a composition series iff $A$ satisfies ACC and DCC. If $D$ is a division ring then $Mat_{n \times n}(D)$ is both Noetherian and Artinian. An ideal $P (\neq R)$ in a commutative ring $R$ is prime iff $R - P$ is a multiplicative set. If $S$ is multiplicative and $S \cap I \neq \emptyset$, $\exists P$, prime that is maximal with respect to the disjoint property. $Rad(I) = \{r \in R : r^n \in I\}$.

Every transcendental extension has a **transcendence basis.** If $< x_1, x_2, \ldots, x_n >$ spans $E$ algebraically and $S \subseteq E$ is algebraically independent then $|S| \leq n$. (Use Steinmetz replacement.)

**Noetherian Normalization Lemma:** Let $R$ be an integral domain which is a finitely generated extension of $K$ and suppose $r$ is the transcendence degree over $K$ of the quotient field of $R$, then $\exists t_1, \ldots, t_r$ algebraically independent elements such that $R$ is integral over $K[t_1, \ldots, t_r]$.

**Localization:** Let $S$ be a multiplicative subset of $R$ and $h : a \mapsto a/1$ be the natural map. If $J$ is an ideal in $S^{-1}R$ then $S^{-1}J = I$ is an ideal of $R$ and $I \subseteq h^{-1}(S^{-1}(I)$ with equality if $I \cap S = \emptyset$. If $I$ is a prime ideal of $R$ and $I \cap S = \emptyset$ then $S^{-1}R$ is a prime ideal of $S^{-1}R$. If $P$ is a prime ideal of $R$ and $S = R - P$ is a multiplicative set, denote $S^{-1}R$ as $R_P$. $R_P$ has a unique maximal ideal consisting of non-units of $R_P$. $\sqrt{I} = P_1 \cap P_2 \cap \ldots \cap P_k$ for some prime ideals $P_i$.

### 1.2.3 Polynomials

**Basic Symmetric polynomials:** $\sigma_1 = \sum x_i$, $\sigma_2 = \sum x_i x_j$, etc. Every symmetric function $f(x_1, \ldots, x_n) = (z - x_1) \ldots (z - x_n)$ can be written as a polynomial with coefficients in the basic symmetric polynomials. Proof 1: Let $a{x_1}^{a_1}{x_2}^{a_2} \ldots {x_n}^{a_n}$ be the leading coefficient of a symmetric form in lexicographic order, subtracting $a{\sigma_1}^{a_1 - a_2}{\sigma_2}^{a_2 - a_3} \ldots {\sigma_n}^{a_n}$ leaves a symmetric form with leading coefficient smaller in lexicographic order. Proof 2: By induction on the weight. True for 1. If $f(x_1, \ldots, x_n)$ is symmetric, so is $\frac{f(x_1,\ldots,x_{n-1},0)}{z}$. So $\frac{f(x_1,\ldots,x_{n-1},0)}{z} = \phi((\sigma_1)_0, \ldots, (\sigma_{n-1})_0)$ Set $f_1(x_1, \ldots, x_n) = f(x_1, \ldots, x_n) - \phi((\sigma_1)_0, \ldots, (\sigma_{n-1})_0)$. $f_1(x_1, \ldots, x_{n-1}, 0) = 0$ so $x_n$ and hence $\sigma_n$ divides $f_1$ thus $f_1 = \sigma_n g$ and $g$ is writable as a polynomial in the basic symmetric functions by induction so $f(x_1, \ldots, x_n) = \sigma_n \psi(\sigma_1, \ldots, \sigma_n) + \phi(\sigma_1, \ldots, \sigma_{n-1})$. Further, the representation is essentially unique which you can show by proving $\phi(y_1, \ldots, y_n) \neq 0 \rightarrow \phi(\sigma_1, \ldots, \sigma_n) \neq 0$ (Prove).

**Resultant:** If $f_v(x) = v_n x^n + \ldots + v_0$ and $g_w(x) = w_m x^m + \ldots + w_0$, $\exists \phi_{v,w}(x), \psi_{v,w}(x) : \phi_{v,w}(x) f_v(x) + \psi_{v,w}(x) g_u(x) = R(v, w) = v_m^n w_n^m \prod_{i<j}(t_i - u_j)$, where $t_i, u_j$ are roots of $f, g$ respectively. Resultant is 0 iff equations have common solution. Consider the equations written in matrix notation:

$$\begin{pmatrix} x^{m-1} f_v(x) \\ x^{m-2} f_v(x) \\ \ldots \\ f_v(x) \\ x^{n-1} g_w(x) \\ x^{n-2} g_w(x) \\ \ldots \\ g_w(x) \end{pmatrix} = \begin{pmatrix} v_n & v_{n-1} & \ldots & v_0 & 0 & 0 & \ldots & 0 \\ 0 & v_n & v_{n-1} & \ldots & v_0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & v_n & v_{n-1} & \ldots & v_0 \\ w_m & w_{m-1} & \ldots & w_0 & 0 & 0 & \ldots & 0 \\ 0 & w_m & w_{m-1} & \ldots & w_0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & w_m & w_{m-1} & \ldots & w_0 \end{pmatrix} \begin{pmatrix} x^{n+m-1} \\ x^{n+m-2} \\ \ldots \\ \ldots \\ \ldots \\ \ldots \\ x \\ 1 \end{pmatrix}$$

Proof: Let the column vectors be $C_{m+n-1} \ldots C_0$. $C = (x^{m-1} f_v(x), \ldots, g_w(x))^T$. $C = C_{m+n-1} \cdot x_{m+n-1} + \ldots + 1 \cdot C_0$. Now solve for 1. $1 = \frac{det(C_{m+n-1} \ldots C_1, C)}{det(C_{m+n-1} \ldots C_1, C_0)}$. Get $\phi_{v,w}(x) f_v(x) + \psi_{v,w}(x) g_w(x) = R(v, w)$.

**Theorem**: Let $f_1, \ldots, f_s$ be polynomials of one variable with indeterminate coefficients. $\exists d_1, d_2, \ldots, d_h$ of integral polynomials in the coefficients of $f_i$ such that if the coefficients are assigned values (**"specialized"**) from $k$, $d_i = 0$ iff either the $f_i = 0$ have a common solution or the leading coefficients vanish. Proof: Set $f_u = u_1 f_1 + \ldots + u_s f_s$, $f_v = v_1 f_1 + \ldots + v_s f_s$. $(f_u, f_v) = 1$ iff $(f_1, f_2 \ldots, f_s) = 1$. $R(f_u, f_v) = 0$ iff $f_u$ and $f_v$ have a non-trivial common factor. But $R(f_u, f_v)$ is a polynomial in $u_i, v_j$ with coefficients which are integral in the coefficients of $f_i$. Arrange these in the order of powers of $u_i v_j$. These are the $d_i$. The proof also shows that $d_i = 0 \pmod{(f_1, f_2, \ldots f_r)}$ and $(d_1, d_2 \ldots d_l) = 0 \pmod{(f_1, f_2, \ldots f_r)}$.

**Theorem**: If $f_1, \ldots, f_r \in F[x_1, \ldots, x_n]$ has no common zeros, $\exists A_1, ...A_r$ such that $\sum_i A_i f_i = 1$. Proof by the induction on number of variables. True for $n = 1$ by usual theory of polynomials over fields. Assume it's true for $n - 1$. Let $\overline{f}_i(x) = f_i(x, x_2, \ldots, x_n) = \sum_{j-0}^{n_i} g_{ij}(x_2, \ldots, x_n) x^j$. The $\overline{f}_i$ have no common solution or the $f_i$ would; thus by the previous result, regarding the coefficients of $x^j$ as indeterminants, $\exists d_{lk}$ which are not simultaneously 0 [or again, the $f_i$ would have a common solution], such that $\sum_{lk} B_{lk} d_{lk} = 1$. After substitution, $\sum_{ij} C_{ij} g_{ij} = 1$. Further, $g_{ij} = \sum_j A_j f_j$, again by the previous result. After substituting again, we get $\sum_j D_j f_j(x, x_2, \ldots, x_n)$ which is what we want.

**Nullstellensatz**: If $f(x_1, \ldots, x_n) \in F$ vanishes at all the common zeros of $f_1(x_1, \ldots, x_n), \ldots, f_r(x_1, \ldots, x_n)$ in every extension of $F$, then $f^k(x_1, \ldots, x_n) \in (f_1(x_1, \ldots, x_n), \ldots, f_r(x_1, \ldots, x_n))$ for some $k$. Look at $f_1, \ldots, f_r, 1 - zf$, put $z = \frac{1}{f}$ and clear denominators. Note that if $h_1, \ldots, h_m$ are zero for all common zeros of the $f_i$, $(h_1, \ldots, h_m)^\rho = 0(f_1, f_2, \ldots, f_r)$.

Note that an algebraic condition for solvability is not always possible: Consider $a_1 x_1 + a_2 x_2 + a_3 = 0$, $b_1 x_1 + b_2 x_2 + b_3 = 0$; they have a solution in general if $a_1 b_2 - b_1 a_2 \neq 0$ and the $d_i$ (the resultant system) would have to vanish for indeterminant $a, b$ and the equation would always have a solution but it doesn't. However, this does work for homogeneous equations (forms).

General idea of **elimination** for forms relies on three lemmas: Lemma 1: We can assume $x_1$ appears with non-zero constant coefficient. Proof: if not, substitute $x_1 = u_1 x_1'$, $x_2 = x_2' + u_2 x_1'$, ..., $x_n = x_n' + u_n x_1'$. Lemma 2: If $\mathcal{F}$ has a non-trivial common solution, the $d_i$ do too. Proof: If the coefficients do not vanish, the $d_i$ give rise to a solution $(\xi_2, \ldots, \xi_n)$ in $(x_2, \ldots, x_n)$ which can be extended to $x_1$; if not, the $d_i$ vanish identically and have a solution, say $(1, 1, \ldots, 1)$ and the $f_i$ have a solution $(1, 0, \ldots, 0)$ with the coefficients of the $x_1$ terms 0. Lemma 3: The system $\mathcal{F}$ has a resultant system of integral polynomials $b_j$ in the coefficients of the $f_i$ such that for a specialization of the coefficients of the $f_i$, $\mathcal{F}$ has a non-trivial common solution iff the $b_j = 0$; further, the $b_j$ are homogeneous in the coefficients of the forms. Elimination procedure: Successively eliminate $x_1, x_2, \ldots, x_n$. At each step, the $d_i$ obtained by eliminating previous $x_i$ are forms, we can continue the elimination procedure until only $x_n$ remains and the resultant system becomes: ${x_n}^{s_1}b_1, {x_n}^{s_2}b_2, \ldots, {x_n}^{s_k}b_k$ and by the above ${x_n}^{s_j}b_j = 0 \pmod{(f_1, \ldots, f_n)}$. If elimination results in a non-zero constant, there is no common solution and we get $1 = 0 \pmod{(f_1, f_2, \ldots, f_r)}$.

Observe that not all solutions can be obtained by specialization. Consider $f_1 = x_1^2 + x_1x_2$, $f_2 = x_1x_2 + x_2^2 + x_1 + x_2$, $(x_1 + x_2)$ is a common factor so the resultant vanishes. $\xi_1 = -\xi_2$ is a solution; however, if $\xi_2 = -1, \xi_1 = 0$ is also solution which does not fit the specialization solution.

For the next few paragraphs, the system $\mathcal{F}$ consists of $r$ forms, $f_1, ..., f_r$ in $n$ variables with indeterminant coefficients. The indeterminants in $f_1$ are $a_1, \ldots, a_\omega$, the indeterminants in $f_2$ are $b_1, \ldots, b_\omega$ and the indeterminants in $f_r$ are $e_1, \ldots, e_\omega$. When $r = n$ the resultant system is generated by a single polynomial, $R$, called the resultant.

Let $\mathcal{F}$ be a system of forms as above with $deg(f_i) = l_1$ and $l_1 = \alpha, l_2 = \beta, \ldots, l_r = \epsilon$. By the above, $\exists T \in \mathbb{Z}[a_1, \ldots, e_\omega]$ such that ${x_i}^\tau T = 0 \pmod{(f_1, \ldots, f_n)}$. $T$ is called an inertial form. Set $f_1 = f_1^* + a_\omega {x_n}^\alpha$, $f_2 = f_2^* + b_\omega {x_n}^\beta$, ... , $f_n = f_n^* + e_\omega {x_n}^\epsilon$, substituting $a_\omega = -\frac{f_1^*}{x_n^\alpha}$, ..., $e_\omega = -\frac{f_r^*}{x_n^\alpha}$, we get $T(a_1, \ldots, -\frac{f_1^*}{x_n^\alpha}, \ldots, -\frac{f_r^*}{x_n^\alpha}) = 0$ (Condition "A") and this actually holds for all $i$ if it holds for any $x_i$. Conversely, if Condition "A" is satisfied, ${x_n}^\tau T = 0 \pmod{(f_1, \ldots, f_r)}$. Proof: We can use Condition A to rearrange $T$ in powers of $a_\omega + \frac{f_1^*}{x_n^\alpha}, \ldots, e_\omega + \frac{f_r^*}{x_n^\epsilon}$ and the term independent of the powers vanishes so $T = 0 \pmod{(a_\omega + \frac{f_1^*}{x_n^\alpha}, \ldots, e_\omega + \frac{f_r^*}{x_n^\epsilon})}$; multiplying through by the largest power of $x_n$ in the denominators, we get $x_n^\tau T = 0 \pmod{(f_1, f_2, \ldots, f_r)}$. The inertial forms form and ideal $\mathcal{I}$ which is prime and a basis for $\mathcal{I}$ thus forms a resultant system.

**Theorem**: If the number of forms, $f_i$, is less than the number of variables, $n$, then there is no inertial form distinct from 0; if $r = n$, there is no inertial form independent of $e_\omega$ and distinct from 0. The proof uses the following Lemma: When a sequence of polynomials $f_1, \ldots, f_s$ in indeterminants $a_1, a_2, \ldots, a_p, x_1, x_2, \ldots, x_q$ is algebraically dependent in $k[a_1, \ldots, a_p]$, this dependence is valid for every specialization $a_p = \alpha$. Proof of Lemma: Since $F(a_1, \ldots, a_p, f_1, \ldots, f_s) = 0$ and $F(a_1, \ldots, a_p, z_1, \ldots, z_s) \neq 0$, $F(a, z)$ is not divisible by $(a_p - \alpha)$ or we could reduce the relations. So $F(a_1, \ldots, a_{p-1}, \alpha, f_1, \ldots, f_s) \neq 0$. Proof of theorem: If $r < n$, by Condition "A", $-\frac{f_1^*}{x_n^\alpha}, \ldots, -\frac{f_r^*}{x_n^\alpha}$ would be algebraically dependent relative to $k[a_1, \ldots, a_{\omega-1}, e_1, \ldots, e_{\omega-1}]$ and this continues to be true if $x_n = 1$. If $r = n$ and the hypothesis is false, $-\frac{f_1^*}{x_n^\alpha}, \ldots, -\frac{f_{n-1}^*}{x_n^\delta}$ would be algebraically dependent relative and we can set $x_n = 1$. In both cases, the lemma applies and we can specialize over any of the indeterminantes without losing dependency. Choose a specialization so $f_1, \ldots, f_s^\delta \rightarrow x_1^\alpha, \ldots, x_s^\delta$. This is a contradiction since these terms are algebraically independent.

**Theorem**: If $r = n$, there is a non-vanishing inertial form $D_e$, homogeneous in the indeterminantes and of degree $L_n = l_1 l_2 \ldots l_{n-1}$ in the $e_j$. Proof: Put $l = 1 + \sum_i^n (l_i - 1)$ and consider, $\mathcal{P}$, the monomials of degree $l$ in the $x_i$. $\mathcal{P}$ is a disjoint union of the following sets: monomials of degree $l$ containing ${x_1}^{l_1}$, monomials of degree $l$ containing ${x_2}^{l_2}$ but not ${x_1}^{l_1}$, ..., monomials of degree $l$ containing ${x_n}^{l_n}$ but not ${x_1}^{l_1}$, ${x_2}^{l_2}$, etc. Suppose $H_{l-l_1}^{(m)}$ are the complementary monomials of the elements of the disjoint sets, i.e. - ${x_1}^{l_1} H_{l-l_1}^{(m)}$ are in the disjoint sets. $H_{l-l_n}^{(m)}$ has $l_1 l_2 \ldots l_{n-1}$ power products ($x_1^k, 0 \leq k < l_1$, etc). Now form $H_{l-l_i}^{(m)} f_i$. Since there are as many of these as power products, the matrix is square. Denote its determinant as $D_e$ which has the value 1 under the specialization $f_i = {x_i}^{l_i}$. Multiplying the equations $H_{l-l_i}^{(j)} f_i = \sum a_{mk} H_l^{(k)}$ by the subdeterminants of a column of $D_e$ and adding, the left hand side becomes linear in the $f_i$ and the right hand side, $D_e H_l^{(k)}$. Letting $H_l^{(k)} = x_i^l$, we get $D_r x_i^l = 0 \pmod{(f_1, f_2, \ldots, f_r)}$ and $D_e$ is homogeneous in each form, $f_i$ and has degree $L_n$ in the coefficients of $f_n$.

Now, let $f_1, f_2, \ldots, f_n$ be forms in $x_1, x_2, \ldots, x_n$ with indeterminate coefficients and $\mathcal{I}$ the ideal generated by the inertial forms. **Theorem:** If $R$ is a polynomial of minimal degree in $e_\omega$, every element of $\mathcal{I}$ is divisible by $R$. $R$ is the resultant. Proof: Arrange $R$ in powers of $e_\omega, R = S{e_\omega}^\lambda + \ldots$. If $T$ is in $\mathcal{I}$, we can get a polynomial, $T' = S^j T - QR$ of lower degree which is also in $\mathcal{I}$ but then $T' = 0$ and $R \mid T$. Note if $R$ vanishes for a specialization, every element of $\mathcal{I}$ does also and the $f_i$ have a common 0; conversely, if the $f_i$ have a common zero, since ${x_i}^\tau R = A_1 f_1 + \ldots + A_n f_n$, substitution makes the right side of the equation 0 but at least one $x_i \neq 0$ so $R = 0$. We have: **Theorem**: $R(gh, f_2, \ldots, f_n) = R(g, f_2, \ldots, f_n)R(h, f_2, \ldots, f_n)$, $R$ is homogeneous of degree $L_1$ in the coefficients of $F_1$, homogeneous of degree $L_2$ in the coefficients of $F_2$, ..., and homogeneous of degree $L_n$ in the coefficients of $F_n$, $R = (D_a, D_b, \ldots, D_e)$ is a principal ideal and the resultant contains a principal term ${a_1}^{L_1} \ldots {e_\omega}^{L_n}$.

**Bezout**: Suppose the system $\mathcal{F}, r = n$ has a finite number of non-trivial solutions $(\xi_1^{(\alpha)}, \ldots, \xi_n^{(\alpha)})$, $\alpha = 1, 2, \ldots, q$. Add the form $l = u_1x_1 + \ldots + u_nx_n$ and form the resultant system $b_1(u), b_2(u), \ldots, b_t(u)$. The resultant system has a solution iff $l_\alpha = u_1{\xi_1}^{(\alpha)} + \ldots + u_n{\xi_n}^{(\alpha)} = 0$. By the Nullstellensatz: $(b_i(u))^{\tau_i} = 0 \pmod{(\prod_\alpha l_\alpha)}$ and $(\prod_\alpha l_\alpha)^\tau = 0 \pmod{D(u)}$ where $D(u) = (b_1(u), \ldots, b_t(u))$. So $D(u) = \prod_\alpha {l_\alpha}^{\rho_\alpha}$ (the $\rho_{alpha}$'s are the multiplicities). If we consider $n - 1$ forms $f_i$ and add the form $l = u_1x_1 + \ldots + u_nx_n$, we get Bezout's theorem, namely: If $n - 1$ homogeneous equations have a finite number of solutions then sum of the multiplicities (defined above) equals the product of the degrees of the equations.

**Berlekamp polynomial factorization:** $f(x)$ square free. Compute $x^{iq} \pmod{f(x)} = \sum q_{ij}x^j$. Find null space of $Q - I$ with basis $v_1(x), \ldots, v_t(x)$. Compute $(f(x), v_k(x) - \alpha)$.

$$f_n(x) = \frac{(x^n - 1)}{\prod_{d|n, d<n} f_d(x)}$$

**Submodules of finitely generated free modules over a PID:** Let $D$ be a PID and $D^{(n)}$ be a free module of rank $n$ over $D$. Then any submodule, $K$ of $D^{(n)}$ is free with base $m \leq n$ elements. Proof: By induction. For $n = 1$, submodule is isomorphic to a principal ideal. Inductive step: examine $\overline{D}^{(n)} = \frac{D^{(n)}}{D^{(n-1)}}$.

Fix a monomial order $(\leq)$ for terms in $x_1, x_2, \ldots x_n$. Denote leading term of $f$ under this order as $in_{\leq}(f)$. The division algorithm for $f$ with respect to the monomial order produces $f(x) = a_1(x)f_1(x) + \ldots + a_m(x)f_m(x) + r(x)$ where $r = 0$ or $r$ is a linear combination of monomials none of which are divisible by $in_{\leq}(f_i)$. This is written as $r = f^F$. Procedure for **multi-variable division algorithm:** Set $r \leftarrow f(x), a_i(x) \leftarrow 0$. Pick ordering of $f_1(x), f_2(x), \ldots, f_m(x)$. If $in_{\leq}(f_j) | in_{\leq}(r)$ for any $j$, pick first such $j$, set $t \leftarrow \frac{in_{\leq}(r)}{in_{\leq}(f_j)}$, $s \leftarrow s - tf_j(x)$, $a_j(x) \leftarrow a_j(x) + t$; repeat this step until if condition fails. $r \leftarrow s$. In general, the result depends on the ordering of the $f_j(x)$.

**Grobner Basis**: A finite subset $G = \{g_1, g_2, ..., g_s\}$ is a Grobner basis for an ideal $I$ with respect to the monomial order $\leq$ if $< in_{\leq}(g_1), in_{\leq}(g_2), ..., in_{\leq}(g_s) > = < in_{\leq}(I) >$. Equivalently, if $f \in I$, $in_{\leq}(g_i) | in_{\leq}(f)$ for some $i$. If $G$ is a Grobner basis $f^G$ is independent of the order of the $f_i(x)$. If $G$ is a Grobner basis and $I = < G >$, $f \in I$ iff $f^G = 0$.

**Dickson's Lemma:** If $S \subseteq N^n$ then $\exists v_1, v_2, \ldots v_m$ such that $S \subseteq (v_1 + N^n) \cup (v_2 + N^n) \cup \ldots \cup (v_m + N^n)$. Consequence: Every ideal has a Grobner basis. Proof: Let $S = \{v : x^v = in_{\leq}(f), f \in I\}$. By Dickson, $S \subseteq \bigcup_i (v_i + N^n), i = 1, 2, \ldots m$. If $f(x) \in I$, $ax^w = in_{\leq}(f), w = v_i + v$ for some $i, v$ then $x^w = x^{v_i}x^v$ hence $in_{\leq}(f_i) | in_{\leq}(f)$.

**Buchberger reduction:** $f \in R$ reduces to 0 with respect to $f = < f_1, f_2, \ldots, f_m > \subseteq R - \{0\}$ iff $\exists a_1, a_2, \ldots, a_m \in R$: $f = a_1f_1 + a_2f_2 + \ldots + a_mf_m$ and $in_{\leq}(a_if_i) \leq in_{\leq}(f)$ if $a_if_i \neq 0$. This is denoted by $f \rightarrow_F 0$. Let $G = (g_1, g_2, \ldots, g_m)$, $I = < G >$. If $f \rightarrow_G 0$ for all $f \in I$ then $G$ is a Grobner basis. If $G$ is a Grobner basis for $I$, $f^G = 0$ iff $f \rightarrow_G 0, \forall f \in I$. $S(f, g) = \frac{x^\gamma}{in_{\leq}(f)}f - \frac{x^\gamma}{in_{\leq}(g)}g$, where $x^\gamma = LCM(in_{\leq}(f), in_{\leq}(g))$. If $S(f_i, f_j) \rightarrow_F 0, \forall i, j$ then $f \rightarrow_F 0, \forall f \in I$. $F$ is a Grobner basis iff $S(f_i, f_j) \rightarrow_F 0, \forall i, j$ iff $S(f_i, f_j)^F = 0, \forall i, j$.

**Buchberger Algorithm:** Test $S(f_i, f_j)^F \neq 0$, $F = F \cup \{S(f_i, f_j)\}$. Do this until all $S(f_i, f_j)^F = 0$. This procedure terminates.

**Minimal Grobner:** $in_{\leq}(f_i)$ does not divide $in_{\leq}(f_j)$ and coefficients are 1. Reduced Grobner: Minimal Grobner where $in_{\leq}(f_i)$ does not divide any term of $in_{\leq}(f_j)$. Example: $F = (x^2 + y, x^2y + 1)$. $S(x^2 + y, x^2y + 1) = y^2 - 1$, $(x^2 + y, x^2y + 1, y^2 - 1)$ is a Grobner basis. **Elimination ideals:** $I_l = I \cap k[x_{l+1}, ..., x_n]$.

**More on Resultants.** Condition 1: $F_0(x_0, x_1, \ldots, x_n) = F_1(x_0, x_1, \ldots, x_n) = \ldots = F_n(x_0, x_1, \ldots, x_n) = 0$, with each $F_i$ homogeneous of degree $d_i$ in the $x_i$. Let $F_i(x_0, x_1, \ldots, x_n) = \sum_{|\alpha|=d_i} u_{i,\alpha} x^\alpha$. **Theorem 1:** Fix $d_0, d_1, \ldots, d_n$, there is a unique polynomial $Res \in \mathbb{Z}[u_{i,\alpha}]$ such that if $u_{i,\alpha}$ are replaced by the corresponding $c_{i,\alpha} \in \mathbb{C}$ and $F_i$ is homogeneous of degree $d_i$ then (a) the equations of condition 1 have a non-trivial solution in $\mathbb{C}$ iff $Res(F_0, F_1, \ldots, F_n) = 0$, (b) $Res(x_0^{d_0}, x_1^{d_1}, \ldots, x_n^{d_n}) = 1$, (c) $Res$ is irreducible. Sometimes we write $Res_{d_0,d_1,\ldots,d_n}$ to emphasize degrees. Note that $Res_{1,1,\ldots,1}$ is just the determinant. **Theorem 2:** For fixed $j, 0 \leq j \leq n$, $Res$ is homogeneous in $u_{j,\alpha}$ of degree $d_0 \cdot d_1 \cdot d_{j-1} \cdot d_{j+1} \cdot d_n$; further, $Res(F_0, \ldots, F_{j-1}, \lambda F_j, F_{j+1}, \ldots, F_n) \lambda^{d_0 \cdot d_1 \cdot d_{j-1} \cdot d_{j+1} \cdot d_n} Res(F_0, F_1, \ldots, F_n)$ and the total degree of $Res$ is $\sum_{j=0}^n d_0 \cdot d_1 \cdot d_{j-1} \cdot d_{j+1} \cdot d_n$. $Res$ is alternating in the $F_i$ and $Res(gh, F_2, \ldots, F_n) = Res(g, F_2, \ldots, F_n) Res(h, F_2, \ldots, F_n)$. Example: $Res_{2,2,2}(F_0, F_1, F_2)$ has 18 variables of total degree 12 and 21,894 terms. If $f(x) = a_l x^l + \ldots + a_0$ and $g(x) = b_m x^m + \ldots + b_0$ then $Res(f, g, x) = a_l^m b_m^l \prod_{i=1}^l \prod_{j=1}^m (\xi - \eta_i) = a_l^m \prod_{i=1}^l g(\xi_i) = b_m^l \prod_{i=1}^m f(\eta_i)$. Put $A_f = k[x]/(f(x))$ and let $[h]_f$ be the natural map from $k[x] \rightarrow A_f$, further, let $m_g : [h]_f \mapsto [gh]_f$ then $m_g$ is a linear map and $Res(f, g, x) = det(m_g)$.

### 1.2.4 Linear Algebra

**Homomorphisms on modules:** Left module $M$ over $R$ with $RM \subset M$, $1m = m, (r+s)m = rm + sm$, etc. Notation: $End_R(X) = Hom_R(X, X)$. $Hom_R(U, V) = \{f, f : U \rightarrow V, f(r_1 u + r_2 v) = r_1 f(u) + r_2 f(v)\}$ where $r_i \in R$.

If $V$ is a vector space (or module) then $V^*$, the set of linear functions over $V$, is the dual space. If $V$ is finite dimensional, $dim(V) = dim(V^*)$. $V \approx V^{**}$. Solution space as kernel of linear map, $L$. $colRank + dim(ker(L)) = n$. $rowRank + dim(ker(L)) = n$.

**Theorem:** The row rank equals column rank. Proof: Let $A = (a_{ij})$ be and $m \times n$ matrix. $R_i = (r_{ij})$ are rows, $C_j$ are columns. Let row rank be $r$ and $S_1, \ldots, S_r$ be a basis. Put $S_i = (s_{ji})$. $R_i = \sum_j k_{ij} S_j$. So $r_{ji} = \sum_i k_{ji} s_{ji}$. So the column vectors $(k_{1i}, k_{2i}, \ldots, k_{mi})^T$ span the column space. Thus the row rank $\leq$ column rank. The same holds for $A^T$.

**Artin's proof that row rank equals column rank.** Lemma: If $W \subset V$ are vector spaces over $k$ and $W^\perp \subset V^*$ then $dim(W) + dim(W^\perp) = dim(V)$. Proof of result: Let $T : k^n \rightarrow k^m$ be the linear transformation represented by the matrix $M$ with rows $r_1, r_2, \ldots, r_m$ and columns $c_1, \ldots, c_n$ and let the row space of $M$ be $R$ and the column space, $C$; finally, let $r = dim(R)$, $c = dim(C)$ and $W = ker(T)$. Since $dim(Im(T)) + dim(W) = n = dim(V)$, $r = n - dim(W)$ and $dim(W) + dim(W^\perp) = n$, it suffices to show $dim(W^\perp) = r$. Note that $r_i \cdot w = 0$ for $w \in W$ so, if $\lambda_i$ is the usual dual basis of $V^*$ with respect to $< e_1, e_2, \ldots, e_n >$ where $< e_1, e_2, \ldots, e_k >= W$. Let $\lambda_j$ be the natural dual basis and note that $R \subseteq < e_{k+1}, e_{k+2}, \ldots, e_n >$ since $r_i \cdot \lambda_j = 0$ for $j \leq k$. Now let $b_{k+1}\lambda_{k+1} + \ldots + b_n \lambda_n = \lambda \in W^\perp$. Consider $\varphi : \lambda \mapsto b_{k+1} e_{k+1} + \ldots + b_n e_n$. If $\varphi(\lambda) = 0$, $\lambda = 0$ so $dim(R) = dim(W^\perp)$ and the result holds.

**Change of basis for matrix:** Let $[e] = \{e_1, \ldots, e_n\}$ be a basis for $V_n$ and let $L$ be a linear transformation on $V_n$. Let $v_{[e]} = [c_1, c_2, ..., c_n]^T$ denote the coordinates of $v$ with respect to $[e]$: $v_{[e]} = c_1 e_1 + ... + c_n e_n$. Let $L_{[e]}$ denote the matrix for $L$ with respect to $[e]$: $L_{[e]} = \sum_j a_{ji} e_j$. Then $L_{[e]} v_{[e]} = (Lv)_{[e]}$. If $f_i = \sum_j b_{ji} e_j$ is another basis, $P = (b_{ij})$ is called the transition matrix from $[f]$ to $[e]$ and $P^{-1}$ is the transition matrix from $[e]$ to $[f]$ (note the sum over the first index). $P v_{[f]} = v_{[e]}$ and $v_{[f]} = P^{-1} v_{[e]}$. Finally, $L_{[f]} = P^{-1} L_{[e]} P$. The same holds over free modules. Alternate notation: $L : V \rightarrow W$, $V$ has basis $\mathcal{B}$ and $W$ has basis $\mathcal{B}'$ with $L(w_i) = \sum_j a_{ij} v_j$ then $M_{\mathcal{B}'}^{\mathcal{B}}(F) = A^T$. If $\mathcal{B}$ and $\mathcal{B}'$ are over the same space, $M_{\mathcal{B}'}^{\mathcal{B}'}(F) = N^{-1} M_{\mathcal{B}}^{\mathcal{B}}(F) N$ where $N = M_{\mathcal{B}}^{\mathcal{B}'}(id)$.

The group of affine transformations is isomorphic to the subgroup of the matrices with last column $(0, 0, \ldots, 0, 1)$. The translations form a normal subgroup.

**Cayley-Hamilton:** Any matrix over an algebraically closed field is similar to a triangular one. The minimum polynomial divides the characteristic polynomial.

Let $A^*$ denote the **adjoint** (conjugate transpose). $(Ax, y) = (x, A^* y)$. **Hermitian:** Self adjoint over complex numbers. **Symmetric:** self adjoint over reals. **Unitary:** $AA^* = I$; equivalently: $A$ is length preserving: $(Ax, Ay) = (x, y)$. If $A$ is symmetric and $X$ is orthogonal then $XAX^{-1}$ is symmetric. If $A$ is symmetric there is a $P$ such that $P^*AP$ is diagonal. All the eigenvalues are real.

Suppose $V$ is a vector space with a non-degenerate bilinear form and $T$ a linear transformation on $V$, if $W$ is $T$-invariant, so is $W^\perp$ and $V = W \oplus W^\perp$. **Witt's Theorem:** Let Q be a non-degenerate quadratic form on $V$ over $F$ of $char(F) \neq 2$, $U_1$ and $U_2$ non-degenerate sub-spaces which are isometric. Then $U_1^\perp$ and $U_2^\perp$ are isometric. Isometries over subspaces can be extended to the whole space. If $V$ is a vector space with over $\mathbb{R}$ with a positive definite form (resp. $\mathbb{C}$ with a hermitian form) and $W$ is a subspace of $V$ then $V = W \oplus W^\perp$. $V^* \otimes V \to \mathcal{L}(V, V)$ via $L_{\phi \otimes v}(w) = \phi(w)v$.

Extreme point in **convex** set: $P$ with no $Q_1$, $Q_2$ such that $P = tQ_1 + (1-t)Q_2$. **Krien Millman:** If $S$ is a closed, bounded convex set, then $S$ is the convex closure of its extreme points.

$A$ is **orthogonal** iff it takes orthonormal basis into orthonormal basis which happens iff $AA^T = I$. Every real quadratic form is equivalent to a diagonal one with a signature of positive and negative coefficients. Two forms are equivalent iff they have the same rank and signature.

**Principal Axis Theorem:** Any real quadratic form is equivalent to one with $Q(\eta) = \lambda_1 {x_1}^2 + \ldots + \lambda_n {x_n}^2$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. Proof: Find eigenvector $v$, $V =< v > \oplus < v >^\perp$.

If $T$ is any linear transform on $V_n$, $\exists M_0, M_1, \ldots, M_n$: (i) $AM_k \subseteq M_k$, (ii) $dim(M_j) = j$, (iii) $\{0\} = M_0 \subseteq M_1 \subseteq \ldots \subseteq M_n = V_n$.

**Nilpotent:** $\exists q$: $A^q = 0$, smallest $q$ is degree of nilpotence. If $A$ is nilpotent of degree $q$, $\exists x$: $A^{q-1}x \neq 0$ and $x, Ax, A^2x, \ldots A^{q-1}x$ are linearly independent. Every linear transform is the direct sum of a nilpotent and an invertible transform.

If $A$ is a linear transform on $V_n$ with proper values $\lambda_1, \lambda_2, \ldots, \lambda_p$ having multiplicity $m_1, m_2, \ldots, m_p$ then $V_n = M_1 \oplus \ldots \oplus M_p$ with $AM_j \subseteq M_j$, $dim(M_j) = m_j$ and $A - \lambda_j I$ is nilpotent on $M_j$.

An $n \times n$ matrix is **diagonalizable** iff it has $n$ linearly independent eigenvectors. A matrix is diagonalizable iff its minimal polynomial is a product of different linear factors. Two matrices are simultaneously diagonalizable iff they are diagonalizable and commute.

**Spectral Theorem:** If $T$ is normal ($TT^* = T^*T$), $\exists E_1, \ldots E_r$ such that $T = \sum_i^r \lambda_i E_i$ with $T = \sum_i^r E_i = I$, $E_i E_j = 0$ and transforming matrix, A, unitary ($\overline{A}^t = A^{-1}$).

Let $f : A \to A'$ be surjective. $A$, $A'$ abelian, $A'$ free. $\exists C \subseteq A$ such that $A = ker(f) \oplus C$.

**Structure Theorem for Finitely Generated Modules over Principal Ideal Domains**: If $M (\neq 0)$ is a finitely generated module over a PID, $D$, $M = Dz_1 \otimes Dz_2 \otimes \ldots \otimes Dz_s$ such that: $(z_1) \supset z_2 \supset \ldots (z_s)$, $z_k \neq D$. Proof: $\eta : D^{(n)} \to M$ canonically (base of $D$ is $e_i$) by $\sum_i a_i e_i \to \sum a_i x_i$. $M \cong D^n/K$. $K$ has base $f_i, i = 1, 2, ..., m$ and $f_j = \sum_i a_{ji} e_i$. Let $e' = Pe$, $f' = Qf$. Relations matrix is $QAP^{-1} = \{d_1, d_2, ..., d_r, 0, 0, ...0\}$ and $d_i \mid d_{i+1}$ $f'_j = d_i e'_i$. Then $y = Px$ is another set of generators and the $y_i$ are linearly independent over $D$. $ann(y_i) = (d_i)$ if $d_i$ is a unit, drop it from the list of generators. If the first $t$ are units, put $z_1 = y_{t+1}$.... $s = n - t$ in the statement of the theorem.

Application to an endomorphism, $Tu_i = \sum_j a_{ij} e_j$. $M \cong D^{(n)}/K$, with $D = F[\lambda]$. $f_i = \lambda e_i - \sum_j a_{ij} e_j$ are generators of $K$. After diagonalization by elementary row and column operations, $P(\lambda I - A)Q = diag(1, \ldots, 1, d_1(\lambda), \ldots, d_s(\lambda))$. $K$ is generated by $f'_i = d_i e'_i$. If $Q^{-1} = (q^*_{ij})$, $v_i = \sum_j q^*_{ij} u_j$, $z_i = v_{n-s+i}$ and $V = F[\lambda] z_1 \oplus \ldots \oplus F[\lambda] z_s$. Example:

$$A = \begin{pmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 2-\lambda & -3 \end{pmatrix}, C = \begin{pmatrix} 1 & 3 & \lambda - 3 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

$$B(\lambda I - A)C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda - 1) & 0 \\ 0 & 0 & (\lambda - 1)^2 \end{pmatrix}, C^{-1} = \begin{pmatrix} 1 & \lambda & -3 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

$v_1 = u_1 + \lambda u_2 - 3u_3$, $v_2 = -u_2 + u_3$, $v_3 = -u_2$. $z_1 = v_2 = -u_2 + u_3$, $z_2 = v_3 = -u_2$, $z_3 = \lambda v_3 = u_1 - 3u_3$. So,

$$\begin{pmatrix} 0 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix} \begin{pmatrix} -1 & -2 & 6 \\ -1 & 0 & 3 \\ -1 & -1 & 4 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix}$$

Let $O_{ij} = (\delta_{il}\delta_{jk})_{1 \le l \le n, 1 \le k \le n}$ and $J_{ij}(\alpha) = I + \alpha O_{ij}$. $J_{ij}(\alpha)A$: add $\alpha$ times row $j$ to row $i$. $AJ_{ij}(\alpha)A$: add $\alpha$ times column $i$ to column $j$.

**Rational Canonical Form** and **Jordan Canonical Form** are the same over an algebraically closed field. A finite group of transformations over $\mathbb{R}^3$ has fixed points. $|G| = v_p n_p$, $2(|G|-1) = \sum_p (v_p - 1)$.

$P_A = A(A^TA)^{-1}A^T$ where the rank of $A$ is the number of columns, is the symmetric projector; $P_{A^\perp} = I - P_A$. $P_A^2 = P_A$, $P_{A^\perp}^2 = P_{A^\perp}$, $P_A^T = P_A$, $P_{A^\perp}^T = P_{A^\perp}$. $S = AA^T$ is invertible. $P_{\vec{a}}(\vec{w})$ is the projection of $\vec{w}$ along $\vec{a}$. The linear system $A\vec{f} = P_A\vec{b}$ has solution $\vec{f} = A^{-1}P_A\vec{b}$ the least squares approximation of data points $(x_i, y_i)$ can be calculated from this too. Example, fit $f(x) = f_0 + xf_1$ to the data $(-1,1), (0,0), (1,2)$ by solving $\begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} (f_0, f_1)^T = (1,0,2)^T$. In general, the least squares approximation arises from the symmetric projection in the sample space $\mathbb{R}^s$ where $s$ is the number of data points. $f(A)\vec{v} = (f_0 + f_1A + \ldots + f_nA^n)\vec{v} = (\vec{v}, A\vec{v}, \ldots, A^n\vec{v})$. Vandermonde determinant and Fourier $V(x_0, x_1, \ldots, x_n)$ where the $x_i$ are roots of $x^{n+1} - 1 = 0$. $Z(x) = (x - x_j)Z_j(x)$ solves for coefficients $f_0, f_1, \ldots, f_n$ using Lagrange interpolants $\Lambda_j(x) = \frac{Z_j(x)}{Z_j(x_j)}$. For PCA, $\mu_A(x) = (x - \lambda_1)^{m_1}(x - \lambda_2)^{m_2} \ldots (x - \lambda_t)^{m_t}$. There are polynomials in $A$, denoted $A_{\lambda_1}A_{\lambda_2} \ldots A_{\lambda_t}$ such that $(A - \lambda_i I)^m A_{\lambda_i} = 0$ and $A = A_{\lambda_1}A_{\lambda_2} \ldots A_{\lambda_t}$. The $A_{\lambda_i}$ are called components. The list of basic eigenvectors of $A$ form the columns of the diagonalizing matrix, $P$ and $AP = PD$; $A$ is diagonalizable when $P$ is invertible.

Approximating a rank $r$ $n \times n$ matrix requires $2nr$ terms. Mean clustering: replace $M$ with $D = diag(\alpha_1, \ldots, \alpha_i)$ where $\alpha_i = \sqrt{\frac{1}{(MM^T)_{ii}}}$. How closely can a scatterplot be approximated by a line $A$ with direction $\vec{a}$? Find the vector $\vec{a}$ that maximizes $|P_{\vec{a}}(\vec{m_1})|^2 + |P_{\vec{a}}(\vec{m_2})|^2 + |P_{\vec{a}}(\vec{m_s})|^2 = (\vec{a}^T\vec{m_1})^2 + (\vec{a}^T\vec{m_2})^2 + \ldots (\vec{a}^T\vec{m_t})^2$. Maximize $\vec{a}^TMM^T\vec{a}, \forall\vec{a} \in \mathbb{R}^s, |\vec{a}| = 1$. $C = MM^T$ is a correlation matrix with $c_{ij}$ is the correlation of $i, j$; if $u_i \perp u_j$ they are uncorrelated. $\exists P : CP = PD$, $MM^T = C = PDP^{-1}$ and maximize $\vec{u}^TD\vec{u}, |\vec{u}| = 1$, $\vec{u} = P^T\vec{a} \in \mathbb{R}^s$. $C$ is diagonalized by $P : PP^T = I$.

### 1.2.5 Bilinear Forms and Classical Groups

A **pairing**, $(W, V) \to k$ is a bilinear map. If $V_0 \subset V$, $V_0^* = \{\vec{w} \in W : (\vec{w}, \vec{v_0}) = 0, \forall\vec{v_0} \in V_0\}$, $v_0 \subset (V_0^*)^*$. $V^*$ is called the left kernel. Same holds *mutatis mutandis* for $W_0 \subseteq W$ provided $W^* = 0$ is the right kernel. If $(W, V) \to k$ is a pairing with left kernel 0 and $\vec{w} \in W$, define $\varphi_{\vec{w}}(\vec{v}) = (\vec{w}, \vec{v})$. $\varphi_{\vec{w}} \in \hat{V}$ and the map $\vec{w} \mapsto \varphi_{\vec{w}}$ is an injection from $W \to \hat{V}$. Similarly, if the right kernel is 0, there is an injection $V \to \hat{W}$.

If $W_0 \subseteq W$, $codim_W(W_0) = dim(W) - dim(W_0)$. If $W_0 \subset W$, $V_0 \subset V$ and $V^* = 0$, there are natural injective morphisms $V/W_0^* \to \hat{W_0}$ and $V_0^* \to \hat{V/V_0}$. Thus, $dim(V/W_0^*) \le dim(\hat{W_0}) = dim(W_0)$ and $dim(W_0^{**}) \le codim(W_0^*) \le dim(W_0)$. If $W = \hat{V}$, both kernels are 0. If $(W, V)$ is a pairing, (a) $dim(W/V^*) = dim(V/W^*)$, (b) if $V^* = 0$, $dim(W_0^{**}) = codim(W_0^*) = dim(W_0)$ and if $W_0$ is finite dimensional, $W_0^{**} = W_0$ and $W_0$ and $V/W_0^*$ are naturally dual, (c) If $V^* = 0$ and $W^* = 0$, and $V$ and $W$ are finite dimensional, $V$ and $W$ are naturally dual and there is a 1-1, inclusion reversing correspondence of subgroups of $V$ and $W$ under the $*$ operator: $W_0 \leftrightarrow W_0^*$.

Let $A = (a_{ij})$ be an $m \times n$ matrix with entries in $k$ and $\vec{x} = (x_1, \ldots, x_n)^T$. Let $\vec{b} = (b_1, \ldots, b_m)^T$ then $A\vec{x} = \vec{b}$ is a system of linear equations. Set $x = E_1x_1 + E_2x_2 + \ldots + E_nx_n, E_i \in V = k^n$. Suppose $\hat{V}$ is dual to $V$ with basis $\varphi_1, \ldots, \varphi_n$: $\varphi_jE_k = \delta_{jk}$. Let $\psi_i(x) = (a_{i1}\varphi_1 + \ldots + a_{in}\varphi_n)(E_1x_1 + \ldots + E_nx_n) = a_{i1}x_1 + \ldots + a_{in}x_n$. $W \subset \hat{V}$, $W =< \psi_j(x) >$ and $dim(W) =$ row rank. $S_m$ is the $m-$tuple column vectors with entries in $k$. Note that if $A_i$ are column vectors forming $A$, they are in the column space of $A$ as is $\vec{b}$ and $A_1x_1 + \ldots + A_nx_n = \vec{b}, \vec{b} = (\psi_1(x), \ldots, \psi_m(x))$. If $f : V \to S_m, f(x) = (\psi_1x, \ldots, \psi_nx)$, $ker(f) = W^*$. If $Im(f) = U$, $U \cong V/W^*$ and $dim(U) = codim(W^*) = dim(W)$. This shows the row rank equals the column rank.

Let $B_{ij}(\lambda) = I + \lambda(\delta_{il}\delta_{jk})_{lk}$. If $A \in GL_n(k), A = BD(\lambda)$ where $B \in SL_n(k)$ and $D(\lambda)$ is the same as the identity except for $\lambda$ in the lower rightmost position. Put $Z = Z(k), S =< x^2, x \in k >$ (as an additive group). If $x^2 \in Z, \forall x$ then $k$ is commutative; further, unless $k$ is commutative and $char(k) = 2$, $S = k$.

$\tau \in GL_n(k)$ is a **transvection** if $\exists H = \{h : \varphi(h) = 0, \varphi \in \hat{V}\}$ with $\tau(h) = h, h \in H$ and $\tau(x) - x \in H, \forall x \in V$. If $\tau$ is a transvection with hyperplane $H$, pick $\vec{b} : \varphi(\vec{b}) = a \neq 0$, set $t(x) = x - \vec{b}a^{-1}\varphi(x)$ then $\tau(t(x)) = t(x)$, thus $\tau(x) = x + \vec{a}\varphi(x)$ with $\vec{a} = \tau(\vec{b}a^{-1}) - \vec{b}a^{-1}$. So all transvections are of this form. $B_{ij}(\lambda)$ is a transvection. If $\vec{a}, \vec{b} \in H$ then $\tau_{\vec{a}}(\tau_{\vec{b}}(x)) = \tau_{\vec{a}+\vec{b}}(x)$. If $\sigma \in GL_n(k)$ and $\tau$ is a transvection, so is $\tau' = \sigma\tau\sigma^{-1}$ and $\tau'(x) = x + (\sigma(A))\varphi(\sigma^{-1}(x))$; conversely, if $\tau''(x) = x + \vec{a'}\psi(x)$ is another transvection with hyperplane $H'$, we show $\exists\sigma$: $\sigma(H) = H'$ and $\sigma(\vec{a}) = a'$ and thus that all transvections are conjugate and hence have the same determinant. Proof: Pick $\vec{b}, \vec{b'}$ with $\varphi(\vec{b}) = \psi(\vec{b'}) = 1$. $\exists\sigma : \sigma(\vec{a}) = \vec{a'}, \sigma(H) = H', \sigma(\vec{b} = \vec{b'}$. Then $\tau''(x) = x + \vec{a'}\varphi(\sigma^{-1}(x))$, $\exists c : \phi(x) = \varphi(\sigma^{-1}(x)$, setting $x = \vec{b'}, \sigma^{-1}(x) = \vec{b}$ we get $c = 1$ and $\tau'' = \tau'$. If $H$ has at least three vectors then $\exists\vec{a}, \vec{b}, \vec{c}$ with $\vec{c} = \vec{a} + \vec{b}$ and $\tau_{\vec{a}}(\tau_{\vec{b}}(x)) = \tau_{\vec{c}}(x)$ and since they all have the same determinant, it must be 1. In that case, $f : GL_n(k) \to GL_n(k)/GL_n(k)'$, $f(\sigma\tau\sigma^{-1}) = f(\tau)$ so all transvections have the same image under $f$ and $\tau \in GL_n(k)' = SL_n(k)$. If $n \geq 3$ $H$ and $H'$ have independent vectors and we can choose $\sigma : det(\sigma) = 1$ so the transvections are conjugate in $SL_n(k)$. Finally, the center of $SL_n(k)$ consists of the matrices $\alpha I$ with $\alpha^n = 1$. We can conclude: If $G$ is a normal subgroup of $GL_n(k)$ containing a transvection and $n \geq 3$ or $n = 2$ and $|k| \geq 4$ then $SL_n(k) \subseteq G$ if $G > Z(GL_n(k))$.

Pairings and isometries: Let $V \times V \to k$ be a pairing with trivial left and right kernels. $\sigma$ is an **isometry** if $(x, y) = (\sigma x, \sigma y), \forall x, y \in V$. $det(\sigma)^2 = 1$ for all isometries; if $det(\sigma) = 1$, $\sigma$ is a rotation, if $det(\sigma) = -1$, $\sigma$ is a reflection. A quadratic map, $Q$ satisfies $Q(ax) = a^2Q(x)$ and $(x, y) = Q(x+y) - Q(x) - Q(y) = (y, x)$ is a pairing. If $char(F) \neq 2$, $Q(x) = \frac{1}{2}(x, x)$. Pairings arising from quadratic maps are symmetric. $\vec{a} \perp \vec{b} \leftrightarrow (\vec{a}, \vec{b}) = 0$. If $< v_1, v_2, \ldots, v_n >$ span $V$ and $(\vec{v_i}, \vec{v_j}) = g_{ij}$ and if $< u_1, u_2, \ldots, u_n >$ is another basis related to the original by $u_i = \sum_j a_{ji}v_j$ then $\overline{g_{ij}} = A^TGA$, where $G = (g_{ij})$. The form is symmetric if $a_{ij} = a_{ji}$, antisymmetric if $a_{ij} = -a_{ji}$.

Let $V^* = rad(V) = V \cap V^\perp$ and $V = rad(V) \oplus U$, $U \cong V/rad(V)$. Suppose $V$ is non-singular and $U \subset V$ then $U^{**} = U, dim(U) + dim(U^*) = dim(V)$ and $rad(U) = rad(U^*) = U \cap U^*$. The subspace $U$ is non-singular iff $U^*$ is non-singular and then $V = U \perp U^*$. A vector $\vec{v}$ is isotropic if $(\vec{v}, \vec{v}) = 0$. $U$ is isotropic if $(u_1, u_2) = 0, \forall u_1, u_2 \in U$. There are two geometries for symmetric metric spaces: (1) **Symplectic** if $(\vec{v}, \vec{v}) = 0, \forall\vec{v} \in V$ and $(x, y) = -(y, x)$; (2) **Orthogonal** if $(x, y) = (y, x), \forall x, y \in V$. If $V$ is orthogonal and every vector is isotropic then $V$ is isotropic.

Suppose $dim(V) = 2$ and $V$ is non-singular but has an isotropic vector, $\vec{n}$ then $\exists\vec{m} : \vec{n}^2 = \vec{m}^2 = 0, \vec{n}\vec{m} = 1, V =< \vec{n}, \vec{m} >$. ( $V =< \vec{n}, \vec{a} >$ for some $\vec{a}$. Set $\vec{m} = x\vec{n} + y\vec{a}$; if $\vec{n}\vec{a} = 0$, $V$ is singular so we can find $y : y\vec{n}\vec{a} = 1$. Can also find $x : \vec{m}^2 = 0$.) $< \vec{n}, \vec{m} >$ is a hyperbolic plane. A non-singular space, $V$, with orthogonal geometry is an orthogonal sum of lines. A non-singular space, $V$, with symplectic geometry is an orthogonal sum of hyperbolic planes. **Witt's Theorem:** Let $V$ and $W$ be isometric via $\rho$. Let $\sigma : V_0 \to W_0$ be an isometry for $V_0 \subset V$ and $W_0 \subset W$, then $\sigma$ can be extended to an isometry of $V$. $O_n$: isometries. $O_n^+$: rotations, $O_n^-$: reflections. $\Omega_n = O_n'$.

If $n$ is odd, $1_V = Z(O_n^+)$. If $n$ is even, $\pm 1_V = Z(O_n^+)$. If $n = 2$ over $F_q$, the plane contains $q + 1$ lines: $< A + xB >, < B >$; if $V$ is isotropic, $\epsilon = 1$, otherwise $V$ contains no isotropic vectors and $\epsilon = -1$. There are $q - \epsilon$ non-isotropic lines. $O(V)$ has $q - \epsilon$ elements. Let $\varphi_n$ be the number of isotropic vectors in $V$ and $\lambda_n$ the number of hyperbolic pairs. If $< N, M >$ is a hyperbolic plane, $< N, M > \oplus < N, M >^* = V$. $< N^* >$ contains $q\varphi_{n-2}$ isotropic vectors. A type I form: TBD. Type I, II form: $\varphi_n = q^{n-1}$. Type III, IV form: $\varphi_n = q^{n-1} + cq^{\frac{n}{2}}$, $n \geq 1$. If $\Phi_n = |O_n^+(q)|$ or $|PSp_n(q)|$, $\Phi_n = \lambda_n\Phi_{n-2}$.

**Classical Groups Summary:** $SL_n(F) =< T_{ij}(b) >$, $T_{ij}(b) = 1 + be_{ij}$. $SL_n(F)' = SL_n(F)$, $n > 1$. $|PSL_n(q)| = (q^n - 1)(q^n - q)\ldots(q^n - q^{n-2})q^{n-1}/(d(q - 1))$, $d = (n - 1, q)$. $\tau_{u,c}(x) = x + cB(x, u)u$. Every orthogonal transform is the product of $\leq n$ reflections. If $U$ is defined by $(x, u) = 0$ and $\tau$ is a transvection, $\exists a \in U, x^\tau = x - (u, a)a$. $< transvections >= SL(V)$. If $G$ is one of $SL(V)$, $Sp(V)$, $SO(V)$ or $S\Omega(V)$, $G = BWB$, where $B$ is the **Borel subgroup** (upper triangular matrices) and $W$ is the **Weyl subgroup**

(the permutation matrices).

### 1.2.6 Fields

**Field extensions:** If $\alpha$ is the root of an irreducible polynomial $p(x) \in F[x]$ then $F(\alpha) = F[\alpha] = F[x]/(p(x))$. Isomorphisms between fields can be extended to isomorphisms of extensions over associated (under the isomorphism) polynomials.

Any two splitting fields of the same polynomial over $F$ are isomorphic. Proof: Let $\alpha$ and $\beta$ be two roots of and irreducible polynomial which divides a $f(x)$; let $E$ be the splitting field of $f(x)$. There is an isomorphism from $F(\alpha)$ into $F(\beta)$ which can be extended to an automorphism of $E$.

Definitions: $E$ is a **Galois** over $F$ if $E_G = F$. $E$ is **normal** over $F$ if an irreducible polynomial over $F$ with one root in $E$, **splits**.

**Artin:** Distinct automorphisms are linearly independent. Proof: Suppose not. Let $c_1\phi_1(x) + c_2\phi_2(x) + \ldots + c_r\phi_r(x) = 0$ be a minimal relation. Since the automorphisms are distinct, $\exists\beta : \phi_1(\beta) \neq \phi_r(\beta)$. Obtain two equations from the minimal relation, the first by substituting $\beta x$ into the equation for beta, the second by multiplying the equation by $\phi_r(\beta)$, then subtract them. This is a shorter relation.

If $G$ is a finite set of automorphisms fixing $F$, then $r = |E : F| \geq |G| = n$. Proof: Suppose not. Let $\{\omega_1, \ldots, \omega_r\}$ be a basis for $E$ over $F$. Consider the $r$ equations: $\phi_1(\omega_k)x_1 + \ldots + \phi_n(\omega_k)x_n = 0$ for $k = 1, 2, \ldots r$. Since $n > r$ there is a non trivial solution $c_1, c_2, \ldots, c_n$. Let $x = \sum_{i=1}^{r} a_i\omega_i$. Multiply the first equation by $a_1$, the second by $a_2$ and so on then add them to get $c_1\phi_1(x) + c_2\phi_2(x) + \ldots + c_n\phi_n(x) = 0$ for all $x$. This contradicts the Artin's result.

Let $G = \{\phi_1, \phi_2, \ldots, \phi_n\}$ be a finite group of $Aut(E)$, $F = E_G$, then $r = [E : F] = |G| = n$. Proof: Suppose $r > n$. Let $\{\omega_1, \ldots, \omega_r\}$ be a basis for $E$ over $F$. Consider the $n$ equations: $\phi_k(\omega_1)x_1 + \ldots + \phi_k(\omega_r)x_r = 0$ for $k = 1, 2, \ldots n$. This has a non trivial solution with $r - n$ more unknowns than equations. Set $a_i = \sum_{j=1}^{n} \phi_j(c_i)$; we can choose $c_1, \ldots, c_{r-n}$ so $a_1, \ldots, a_{r-n}$ are not 0. The $a_i$ are fixed by $G$ so they are in $F$. $\sum_{i=1}^{r} a_i\omega_i = \sum_{i=1}^{r}\sum_{j=1}^{n} \phi_j(c_i)\omega_i = \sum_{j=1}^{n}\sum_{i=1}^{r} \phi_j(c_i)\omega_i = \sum_{j=1}^{n} \phi_j(\sum_{i=1}^{r} c_i{\phi_j}^{-1}\omega_i) = 0$. But then $\sum_{i=1}^{r} c_i{\phi_j}^{-1}\omega_i = 0$ which contradicts the linear independence of the $\omega_i$'s. So $r \leq n$. Now $r \geq n$ by the previous result so $n = r$.

**Primitive Element Theorem:** If $E = F[\alpha_1, \ldots, \alpha_n]$ with $\alpha_2, \ldots, \alpha_n$ separable then $E = F[\alpha]$, some $\alpha$. Every separable finite extension is primitive. Proof: Assume $F$ is not finite, $E = F[\alpha, \beta]$ with $f, g$ the minimal polynomials for $\alpha = \alpha_1$ and $\beta = \beta_1$ respectively, $\alpha_i$ the roots of $f$ and $\beta_i$ the roots of $g$. Let $E$ be the splitting field of $f(x)g(x)$. $\alpha_i + x\beta_k = \alpha_1 + x\beta_1$ has one root for each $i, k$; pick $c$ such that $\alpha_i + c\beta_k \neq \alpha_1 + c\beta_1$ and set $\theta = \alpha + c\beta$. Claim: $E = F[\theta]$. $f(\theta - c\beta) = g(\beta) = 0)$ so $(f(\theta - cx), g(x)) = (x - \beta) \in F[\theta][x]$.

Let $E$ be a splitting field for $f(x)$ over $F[x]$. If $p(x)$ is irreducible and has one zero in $E$, then $p(x)$ splits in $E$. Proof: Let $L$ be the splitting field of $f(x)p(x)$. Set $E = F(a_1, a_2, \ldots, a_n)$ where $a_1, a_2, \ldots, a_n$ are the roots of $f(x)$. Suppose $p(\alpha) = 0, \alpha \in E$ and $p(\beta) = 0$. Let $\sigma : F(\alpha) \rightarrow F(\beta)$ be an isomorphism with $\sigma(\alpha) = \beta$. Extend $\sigma$ to $\tau : L \rightarrow L$. $\tau$ permutes the roots of $f(x)$ so $\tau(E) = E$. $\alpha = \frac{m(a_1,a_2,\ldots,a_n)}{n(a_1,a_2,\ldots,a_n)}$. So $\beta = \tau(\alpha) = \tau(\frac{m(a_1,a_2,\ldots,a_n)}{n(a_1,a_2,\ldots,a_n)}) \in E$.

Let $E$ be a finite extension of $F$, $char(F) = 0$. If $E$ is a splitting field of $f(x) \in F[x]$ then $|\mathcal{G}(E/F)| = [E : F]$. Proof: $E = F(w)$, $p(w) = 0$ and $p$ splits by foregoing. $deg(p) = [E : F] = |G|$.

Let $F \subseteq E$, $char(F) = 0$. If $G = \mathcal{G}(E/F)$ fixes $F$ then $E$ is a normal extension iff $F$ is the fixed field of $G$. Proof: $E = F(w)$, $|G| = [E : F]$. Let $K = \{a : \sigma(a) = a, \forall\sigma \in G\}$. $F \subseteq K \subseteq E$ and $E = K(w)$. STS if $g$ is irred over $F$ and $g(w) = 0$ then $g$ is irreducible over $K$. Let $p$ be an irreducible polynomial for $w$ over $K$. Applying elements of $G$, each root of $p$ is a root of $g$.

Let $E$ be a normal extension of $F$. $E \supset K \supset F$. If $\mathcal{G}(E/F) > S$ has $K$ as a fixed field then $\mathcal{G}(E/K) = S$.

$E$ is Galois over $F$ iff (i) every irreducible polynomial in $F[x]$ with one root in $E$ splits and (ii) $E = F(\theta)$. $GF(p^m) \subseteq GF(p^n)$ iff $m|n$. The following are equivalent: (1) $E$ is a splitting field over $F$ of a separable polynomial $f(x)$. (2) $F = E_G$. (3) $E$ is finite dimensional, normal and separable. Lemma: Let $K$ be the splitting field of $f(x)$ over $k$ and let $p(x)$ an irreducible factor of $f(x)$, if the roots of $p(x)$ are $\alpha_1, \ldots, \alpha_r$, there is a $\sigma_i \in \mathcal{G}(K/k)$ such that $\sigma_i(\alpha_1) = \alpha_i$.

**Galois:** Let $K$ be a normal, separable extension of $k$. Let $G = \mathcal{G}(K/k)$, $H < G$, $K \supset F \supset k$. There is a bijective pairing of $H, F$, such that (i) $H_1 \supset H_2 \leftrightarrow Inv(H_2) \supset Inv(H_1)$, (ii) $|H| = [K : Inv(H)]$, $[G : H] = [Inv(H) : k]$ and (iii) $H \lhd G \leftrightarrow Inv(H)$ is normal over F and $\mathcal{G}(Inv(H)/k) = G/H$.

If $f(x)$ is solvable by radicals, the Galois group of its splitting field is **solvable**. Galois group of an equation is a permutation group on its roots. Splitting field of $2x^5 - 10x + 5$ is $S_5$.

**Compute Galois group for arbitrary polynomial:** $f(t) = t^n - s_1 t^{n-1} + s_2 t^{n-2} - \ldots (-1)^n$. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the roots. For $\sigma$ in $S_n$, set $\beta = x_1\alpha_1 + \ldots + x_n\alpha_n$ and put $\sigma_x(\beta) = x_{\sigma(1)}\alpha_1 + \ldots + x_{\sigma(n)}\alpha_n$ and $\sigma_\alpha(\beta) = x_1\alpha_{\sigma(1)} + \ldots + x_n\alpha_{\sigma(n)}$. $\sigma_x(\beta) = \tau_x(\beta)$ iff $\sigma = \tau$ (since the roots are distinct). Set $Q = \prod_{\sigma \in S_n}(t - \sigma_\alpha(\beta))$ then $Q = \sum_{j=0}^{n!}(\sum_i g_i(s_1, s_2, \ldots, s_n) x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n}) t^j$. Factor $Q$ into irreducible factors: $Q = Q_1 Q_2 \ldots Q_k$ with $(t - \beta)|Q_1$. $Q_j = \prod_{\sigma \in T_j}(t - \sigma_x(\beta))$ and $\bigcup_j T_j = S_n$. Now $Q = \sigma_x(Q) = (\sigma_x Q_1) \ldots (\sigma_x Q_k)$, i.e. $\sigma_x$ permutes the irreducible factors of $Q$. Define $G = \{\sigma \in S_n : \sigma_x Q_1 = Q_1\}$. Theorem: $G = \mathcal{G}(E/K)$. Hint: If $g \in \mathcal{G}(K/k)$ then $g$ transforms $\beta$ into a conjugate; so does the $\sigma$. This lets us prove the following: Let $R$ be a UFD and $p$ a prime. Set $\overline{R} = R/(p)$ and let $Q_R$ and $Q_{\overline{R}}$ be their fields of quotients. Let $f(x)$ and $\overline{f(x)}$ be corresponding polynomials with no double roots with corresponding splitting fields $K$ and $\overline{K}$ respectively. Then $\mathcal{G}(\overline{K}/Q_{\overline{R}}) < \mathcal{G}(K/Q_R)$.

**Valuation:** $\varphi : K \rightarrow \mathbb{F}^{\geq 0}$ where $\mathbb{F}$ is an ordered field such that $\varphi(ab) = \varphi(a)\varphi(b)$, $\varphi(0) = 0$, $\varphi(x) > 0$ if $x \neq 0$ and $\varphi(a+b) \leq \varphi(a) + \varphi(b)$. If $a = \frac{s}{t}p^n$, $\varphi(a) = p^{-n}$ is a valuation. Ostowski: A non trivial valuation of $\mathbb{Q}$ is either (i) $\varphi(a) = |a|^\rho, 0 < \rho \leq 1$ (the Archemedean valuation) or (ii) $\varphi(a) = \varphi_p(a)$ (the $p-$adic valuation. $w(a) = log(\varphi(a))$ is the exponential valuation. Set $\wp = \{a : w(a) > 0\}$. Hensel: Let $K$ be complete in the exponential valuation $w$ and $f(x)$ a primitive polynomial in $K[x]$ with integral coefficients. Let $g_0, h_0$ be polynomials with integral coefficients such that $f(x) = g_0(x)h_0(x)$ $(\wp)$ then there are polynomials $f(x), h(x)$ with integral coefficients in $K$ such that (1) $f(x) = g(x)h(x)$, (2) $g(x) = g_0(x)$ $(\wp)$, (3) $h(x) = h_0(x)$ $(\wp)$ provided $(g_0(x), h_0(x)) = 1$ further $deg(g) = deg(g_0)$ $(\wp)$.

$F$ is **perfect** iff every irreducible polynomial is separable. $F$ is perfect if (1) $char(F) = 0$, (2) $char(F) = p$ and every element is a $p$th root, (3) $F = GF(q)$, (4) $F$ is algebraically closed.

Let $E = F[\theta]$ and $\rho = a_0 + a_1\theta + \ldots + a_{n-1}\rho^{n-1}$. $T(\rho) = \sum_{g \in \mathcal{G}(E/F)} \rho^g$ is the **trace** and $N(\rho) = \prod_{g \in \mathcal{G}(E/F)} \rho^g$ is the **norm**; both are in $F$.

For every $q = p^n$ there is, up to isomorphism, only one field $F = GF(q)$ and the multiplicative group is cyclic. Consider $f(x) = x^h - 1, h = q - 1$ whose roots are roots of 1. The automorphisms of $F$ are exactly $\sigma_i : x \mapsto x^{p^i}$. If $char(F) = p$, every irreducible polynomial $f(x)$ of degree $n$ either has distinct roots or is of the form $\phi(x^p)$ in which case all roots have the same multiplicity $p^l$ for some $l > 0$ with $n = n'p^l$ in which case there are $n'$ relative automorphisms. Thus in successive extensions there are $\prod_i n'_i$ relative automorphisms which have cardinality $[E : F]$ if $E$ is a separable extension and $< [E : F]$ if not.

If $G$ is solvable, $G^{(n)} = 1$ for some $n$. If $n > 4$, then $S_n^{(m)}$ contains every 3 cycle for every $m$.

Suppose $f \in k[x], deg(f) = n$ and let $\mathcal{G}_f(k)$ denote $\mathcal{G}(K/k)$ where $K$ is the slitting field for $f$ over $k$. Then $\mathcal{G}_f(k)$ is isomorphic to some subgroup of $S_n$ and if $f$ is irreducible, the group is transitive on $n$ symbols. Set $\Delta = \prod_{i<j}(u_i - u_j)$ and $Disc_k(f) = \Delta^2$, then if $f$ is irreducible, the Galois group is $A_3$ or $S_3$ according to whether $Disc_k(f) = \Delta^2$ is a square in $k$. If $f$ is a quartic with separated roots $u_1, u_2, u_3, u_4$ and $\alpha = u_1u_2 + u_3u_4$, $\beta = u_1u_3 + u_2u_4$, $\gamma = u_1u_4 + u_2u_3$; setting $K = k(\alpha, \beta, \gamma)$ and $[K : k] = m$, then $\mathcal{G}_f(k)$ is $S_4$ if $m = 6$, $\mathcal{G}_f(k)$ is $A_4$ if $m = 3$, $\mathcal{G}_f(k)$ is $\mathbb{Z} \times \mathbb{Z}$ if $m = 1$, and $\mathcal{G}_f(k)$ is $\mathbb{Z}_4$ or $D_4$ if $m = 2$.

Let $k \subset K \subset \overline{k}$ and $\sigma_1, \sigma_2, \ldots, \sigma_r$ be the distinct $k$-monomorphisms from $K \rightarrow \overline{k}$, for $u \in K$, define

$N_k^K(u) = (\prod_i \sigma_i(u))^{[K:k]_i}$ and $Tr_k^K(u) = (K : k]_i \sum_i \sigma_i(u)$. Note that distinct automorphisms are linearly independent. From now on, assume all extensions are separable (even Galois). $N_k^K(uv) = N_k^K(u)N_k^K(v)$ and $Tr_k^K(u+v) = Tr_k^K(u) + Tr_k^K(v)$; if $u \in k$, $N_k^K(u) = u^{[K:k]}$ and $Tr_k^K(u) = [K : k]u$; if $E$ is an intermediate field, $N_k^K(u) = N_k^E(N_E^K(u))$ and $Tr_k^K(u) = Tr_k^E(Tr_E^K(u))$. If $K$ is a cyclic extension of $k$ of degree $n$ with generator $\sigma$ then $Tr_k^K(u) = 0$ iff $\exists v \in K : u = v - \sigma(v)$ and $N_k^K(u) = 1$ iff $\exists v \in K : u = v(\sigma(v))^{-1}$. If $n = mp^t, (p, n) = 1$ where $char(k) = p \neq 0$, there are intermediate cyclic fields, all of which, except the last have degree $p$ and each of which is the splitting field of $f(x) = x^p - x + a$. If $char(k) = p \neq 0$, $K$ is a cyclic extension of degree $p$ iff $K$ is the splitting field of an irreducible polynomial $f(x) = x^p - x - a$ and $K = k(u), f(u) = 0$. Suppose $\zeta$ is a primitive $n$th root of unity over $k$ and $K = k(\zeta)$, if $d \mid n$, $\zeta^{n/d}$ is a primitive $d$-th root of unity and, $K$ is the splitting field over $k$ of an irreducible polynomial $f(x) = x^d - a, a \in k$. If $k$ contain a primitive $n$-th root of unity, $\zeta$, TFAE: (1) $K$ is cyclic of degree $d$ $d \mid n$, (2) $K$ is the splitting field over $k$ of $f(x) = x^n - a, a \in k$, (3) $K$ is the splitting field over $k$ of an irreducible polynomial $f(x) = x^d - a, a \in k$.

Let $(n, char(k)) = 1$, and $K$ a **cyclotomic extension** of $k$, then, (1) $K = k(\zeta)$ where $\zeta$ is a primitive $n$-th root of unity; (2) $K$ is an **abelian extension** of $k$ of dimension $d, d \mid \psi(n)$; (3) $|\mathcal{G}(K/k)| = d$ and is a subgroup of $\mathbb{Z}_n^*$.

**Radical extensions:** $K = k(u_1, u_2, \ldots, u_n)$ where $\exists n_1 : {u_1}^{n_1} \in k$ and $\exists n_m : {u_m}^{n_m} \in k(u_1, \ldots, u_{m-1})$. $f$ is said to be solvable by radicals if there is a radical extension containing the splitting field of $f$. If $K$ is a radical extension of $k$ and $E$ is an intermediate field then $\mathcal{G}(E/k)$ is solvable. If $E$ is a finite dimensional extension of degree $n$, $char(k) \nmid [E : k]$ and $\mathcal{G}(E/k)$ is solvable then there is a radical extension $K$ of $k$ containing $E$. If $char(k) \nmid n!$ and $f \in k[x], deg(f) = n$ then $f(x) = 0$ is solvable by radicals iff $\mathcal{G}_f$ is solvable.

### 1.2.7 Boolean Functions

For boolean functions, $f : GF(2)^n \rightarrow GF(2)$ and $g : GF(2)^n \rightarrow GF(2)$, define $C(f, g) = 2Prob(f(x) = g(x)) - 1$. Consider two real vectors, in $\mathbb{R}^{2^n}$,

$$\vec{a} = ((-1)^{f(0)}, (-1)^{f(1)}, \ldots, (-1)^{f(2^n-1)})$$

and

$$\vec{b} = ((-1)^{g(0)}, (-1)^{g(1)}, \ldots, (-1)^{g(2^n-1)})$$

We denote $< f, g >=< \vec{a}, \vec{b} >$ and $||f|| = \sqrt{< f, f >}$. With this notation, $C(f, g) = \frac{<f,g>}{||f|| \cdot ||g||}$. The vectors $\vec{w} = (-1)^{w \cdot x}$ as x varies over $GF(2)^n$ are called the linear **parities** and form an orthogonal basis for $\mathbb{R}^{2^n}$. The **correlation matrix**, $C$, for a boolean function $f$, is a row matrix (indexed by $w$) defined by $C(f(x), w^T \cdot x) =< (-1)^{f(x)}, (-1)^{w^T \cdot x} >$ and hence consists of the projections of the "reified" version of $f$ on each of the parities. The definition of a correlation matrix can be extended to a vector boolean function $h : GF(2)^n \rightarrow GF(2)^m$ (or $m$ boolean functions) and, in this case, the correlation matrix, $C$, is a $2^m \times 2^n$ matrix. This matrix has entries $C_{uw} = C(u^T \cdot h(a), w^T \cdot a)$ where $u$ indexes the rows and $w$ indexes the columns; thus the $u$ row is represented as $(-1)^{u^T \cdot h(a)} = \sum_w C^{(h)}_{u,w}(-1)^{w^T \cdot a}$. To emphasize the association with $h$, we sometimes write the correlation matrix as $C^{(h)}$.

**Hadamard-Walsh Transform and correlation:** For boolean function, $f : GF(2)^n \rightarrow GF(2)$, define $F(w) = 2^{-n} \sum_x (-1)^{f(x)+w \cdot x} = C(f(a), w^T a)$ and we say $\mathcal{W}(f) = F$ and call $\mathcal{W}$ the Walsh or Hadamard transform. Actually, owing to the factor $2^{-n}$ in front of the sum this is the normalized Walsh transform, the term "Walsh Transform" is also used for the operation without the $2^{-n}$ and to distinguish, we will describe this as the "un-normalized" Walsh transform. Basic results: $\sum_w F(w)^2 = 1$ (Parseval). If $f(x) = g(Mx+b)$, $M$, invertible, the absolute value of the spectrums of $F$ and $G$ are the same. $dist(f(v), u \cdot v) = \frac{1}{2}(2^n - 2^n \hat{F}(u))$. $dist(f(v), u \cdot v + 1) = \frac{1}{2}(2^n + 2^n \hat{F}(u))$. Define $A \otimes B = (a_{ij}B)$. The operation is associative but not commutative. $\mathcal{W}(f \oplus g) = \mathcal{W}(f) \otimes \mathcal{W}(g) = \sum_v F(v \oplus w)G(v)$. Also, $\mathcal{W}(fg) = \frac{1}{2}(\delta(w) + \mathcal{W}(f) + \mathcal{W}(g) - \mathcal{W}(f \oplus g))$. All correlation matrices are doubly stochastic. Involutions have symmetric correlation matrices. **Fast Hadamard Transform:** $H_{2^m} = H_2 \otimes H_{2^{m-1}}$. $H_{2^m} = M^{(1)}_{2^m} M^{(2)}_{2^m} \ldots M^{(m)}_{2^m}$, $M^{(i)}_{2^m} = I_{2^{m-1}} \otimes H_2 \otimes I_{2^{i-1}}$.

$W(\widehat{f})(t) = \widehat{F}(t) = \sum_x \widehat{f}(x)(-1)^{x \cdot t}$. If $f$ is boolean, $\widehat{f}(x) = (-1)^{f(x)}$. **Convolution:** $f * g(a) = \sum_x f(x)g(x+a)$. Theorem: $W^{-1}(F)(x) = f(x) = 2^{-n} \sum_t F(t)(-1)^{x \cdot t}$. $W(f * g) = W(f)W(g)$. For Boolean $f$,

$f(v_1, v_2, \ldots, v_m) = \prod_{a \in V^m} g(a) {v_1}^{a_1} {v_2}^{a_2} \ldots {v_m}^{a_m}$ where $g(a) = \sum_{b \subseteq a} f(b_1, b_2, \ldots, b_m)$ (subset means positions of 1's in a is a subset of b positions of 1's in b.) The "correlation coefficients" are $\hat{c}_{fg}(b) = C(f(a), g(a \oplus b)) = 2^{-n} \sum_a (-1)^{f(a) \oplus g(a \oplus b)} = \mathcal{W}^{-1}(FG)$.

A **balanced boolean function** is uncorrelated with either constant function. **Overall Question:** What is the best affine approximation of a balanced function? The question is important because if $E(k, x)$ is a block cipher on blocks of $n$ bits, each $E_i(k, x)$ is a balanced boolean function. How many inputs satisfy all approximations? Fail on all approximations? For the correct input, what are the expected number of equations that agree with it? Variance, etc.

**Theorem:** If $f$ is balanced, $\sum_w F(w) = \pm 2^n$. Proof: $\sum_w F(w) = \sum_w \sum_x (-1)^{f(x)+w \cdot x} = \sum_x (-1)^{f(x)} (\sum_w (-1)^{w \cdot x}) = \sum_x (-1)^{f(x)} 2^n \delta_{w,x}$, so $\sum_x (-1)^{w \cdot x + c} = (-1)^c 2^n, w = 0, 0, w \neq 0$. Let $F(w, c) = \sum_x (-1)^{f(x+w \cdot x+c)}$ then $\sum_{w,c} F(w, c) = 0$.

All Hadamard transform values of **bent functions** are equal to $\pm 2^{\frac{m}{2}}$ and hence the distance to any affine function is $2^m \pm 2^{\frac{m}{2}-1}$. If $f(x_1, x_2, \ldots, x_m)$ is bent and $m \geq 6$ then $f$ is indecomposable. $f(u_1, \ldots, u_m, v_1, \ldots, v_m) = g(v_1, \ldots, v_m) + \sum_i u_i v_i$ is bent. If $f(u_1, \ldots, u_m, v_1, \ldots, v_m) = \sum_i u_i v_i$, then $f + u_1 u_2, u_3$, $f + u_1 u_2, u_3 u_4$, $\ldots$, $f + u_1 u_2, u_3 \ldots u_m$ are all inequivalent bent functions.

In this paragraph, $F$ denotes the unnormalized Walsh transform of $f$. A function $z = f(x_1, \ldots, x_n)$ on $n$ variables $x_1, \ldots, x_n$ is $m$-th order **correlation immune** if for every subset of these variables or size $m$, $I(z; x_{i_1}, \ldots, x_{i_m}) = 0$. If $f$ has correlation immunity $m$ and non-linear order $k$, $m + k \leq n$. Let $N_{ab}(\omega) = |\{x : z = f(x) = a, \omega \cdot x = b\}|$ then $F(\omega) = N_{10}(\omega) - N_{11}(\omega)$. Denote $p_a = P(z = a)$ then $P(\omega \cdot x = b | z = a) = \frac{P(\omega \cdot x = b, z = a)}{P(z=a)} = p_a^{-1} 2^{-n} N_{ab}(\omega)$. We obtain the following: $P(\omega \cdot x = 0 | z = 1) = \frac{1}{2} + p_1^{-1} 2^{-n-1} F(\omega)$, $P(\omega \cdot x = 1 | z = 1) = \frac{1}{2} - p_1^{-1} 2^{-n-1} F(\omega)$, $P(\omega \cdot x = 0 | z = 0) = \frac{1}{2} + p_0^{-1} 2^{-n-1} F(\omega)$, $P(\omega \cdot x = 1 | z = 0) = \frac{1}{2} - p_0^{-1} 2^{-n-1} F(\omega)$. Let $h(t) = -t lg(t) - (1-t) lg(1-t)$. **Theorem 1:** Let $x_0, \ldots, x_{n-1}$ be independent and uniformly distributed arguments of the boolean function $f$ whose output is the random variable $z$; then $\forall \omega \neq 0, I(z; \omega \cdot x) = 1 - p_0 h(\frac{1}{2} - \frac{F(\omega)}{2^{n+1} p_0}) - p_1 h(\frac{1}{2} - \frac{F(\omega)}{2^{n+1} p_1})$. Moreover, when $z$ is uniformly distributed then $I(z; \omega \cdot x) = 1 - h(\frac{1}{2} - 2^{-n} F(\omega))$. $F$ thus describes the best affine approximation of $f$ (pick $\omega$ with largest coefficient, the coefficients of the best affine approximation has coefficients of 1 for the corresponding variables). This generalizes to **Theorem 2:** Let $x_0, \ldots, x_{n-1}$ be independent and uniformly distributed arguments of the boolean function $f_i \in \mathcal{F}$ where $\mathcal{F} = \{f_1, \ldots, f_m\}$, $p_f = \frac{1}{m}$ and the outputs of the randomly selected $f_i$ is the random variable $z$; then $\forall \omega \neq 0, I(z; \omega \cdot x) = 1 - p_0 h(\frac{1}{2} - \frac{\sum_{i=1}^m F_i(\omega)}{2^{n+1} m p_0}) - p_1 h(\frac{1}{2} - \frac{\sum_{i=1}^m F(\omega)}{2^{n+1} m p_1})$. Moreover, when $z$ is uniformly distributed then $I(z; \omega \cdot x) = 1 - h(\frac{1}{2} - 2^{-n+1} m^{-1} \sum_{i=1}^m F_i(\omega))$. Again, this provides the best affine approximation for the set of functions. Finally, this implies **Theorem 3:** A boolean function $f$ is correlation immune of order $m$ if $F(\omega) = 0, \forall \omega : 1 \leq wt(\omega) \leq m$.

**Counting Results:** Let $N = 2^n$ and $BF(n)$ denotes the set of boolean functions on $n$-bit values then $|BF(n)| = 2^N$. Let $BBF(n)$ be the balanced functions on $n$ bits then $|BBF(n)| = \binom{N}{\frac{N}{2}}$, $|GA(n)| \approx 2^{m^2+m}$.

**The natural isomorphism:** $\mathcal{L} : GF(2)^n \rightarrow \mathbb{R}^{2^n}$ by $a \mapsto (-1)^{a^T \cdot x}$. $\mathcal{L}(a + b) = \mathcal{L}(a)\mathcal{L}(b)$ by pointwise multiplication. Almost directly from the definitions, we get **Theorem:** $C^{(h)}(\mathcal{L}(a)) = \mathcal{L}(h(a))$.

If $h(x) = f(g(x))$ then $C^{(h)} = C^{(f)} C^{(g)}$ because $(-1)^{u^T \cdot h(a)} = \sum_v C^{(f)}_{u,v} (-1)^{v^T \cdot g(a)} = \sum_v C^{(f)}_{u,v} (\sum_w C^{(g)}_{v,w} (-1)^{w^T \cdot a})$. If $h$ is invertible, $(C^{(h)})^{-1} = (C^{(h)})^T$. (For a bijection, $C(u^T h^{-1}(a), w^T a) = C(u^T b, w^T h(b)) = C(w^T h(b), u^T b)^T$, so, $C^{(h^{-1})} = (C^{(h)})^{-1}$.)

**Theorem:** A boolean transformation is invertible iff its correlation matrix is invertible. The → direction follows from the inverse formula above. The proof of ←: $(-1)^{u^T h(a)} = \sum_w C^{(h)}_{u,w} (-1)^{w^T a}$. If $C^{(h)}$ is invertible, $(-1)^{w^T a} = \sum_u (C^{(h)})^{-1}_{w,u} (-1)^{u^T h(a)}$. If $\exists x \neq y : h(x) = h(y)$, substituting into the equation above, $(-1)^{w^T x} = (-1)^{w^T y}$ and that is just wrong.

**Correlation matrices for standard functions:** If $h(x) = x + k$, $C_{u,u} = (-1)^{u^T k}$. If $h(x) = Mx$, $C_{u,w} = \delta(M^T u \oplus w)$. If $h(x) = (b_{(1)}, b_{(2)}, \ldots, b_{(n)})$, $b_{(i)} = h_{(i)}(a_{(i)})$ and $C^{(i)} = C^{h_{(i)}}$ then $C_{u,w} = \prod_i C^{(i)}_{u(i),w(i)}$ (uses disjunct support). If $h(x) = g(x) + w^T x$, $H(u) = G(u \oplus w)$; if $V_f \cap V_g = \emptyset$, $w \in V_f$, $u \in V_g$,

$H(u+w) = F(w)G(u)$.

**Theorem:** $C^{(h)}_{u+v,x} = \sum_w C^{(h)}_{u,w+x} C^{(h)}_{v,w}$. Proof: $\mathcal{W}((u \oplus v)^T h(a)) = \mathcal{W}(u^T h(a)) \otimes \mathcal{W}(v^T h(a))$; note that first transform on right is $C^{(h)}_{u,w}$ and second is $C^{(h)}_{v,w}$. One consequence is: $C_{u \oplus v,0} = \sum_w C_{u,w} C_{v,w}$.

**Theorem:** A Boolean transformation is invertible iff every output parity is a balanced binary boolean function of the input bits. Proof of →: If $h$ is invertible, $CC^T = I$, $C_{00} = 1$ and the norm of every row and column is 1. $C(u^T h(a), 0) = \delta(u)$; all rows except row 0 are correlated to 0 hence the function is balanced for $u \neq 0$. For ←: The condition on output parities being balanced is $C_{u,0} = 0, u \neq 0$. i.e.- $C$ is orthogonal. $CC^T = I \leftrightarrow \sum_w C_{u,w} C_{v,w} = \delta(u \oplus v)$ ("*") also $\sum_w C_{u,w} C_{v,w} = C_{u \oplus v,0}$ but $C_{u,0} = 0, u \neq 0$ and $C_{00} = 1$ so "*" holds $\forall u, v$ hence $C$ is orthogonal. Let $\vec{f}$ and $\vec{g}$ be two surjective boolean transformations on $n$ variables and define $C(\vec{f}, \vec{g})$ in the obvious way. $C(\vec{f}, \vec{g})$ is invertible but not necessarily invertible. If $u$ and $w$ are parities then and $F^u$ denotes the normalized Walsh transform of $u^T \vec{f}(\vec{x})$ while $G^w$ denotes the normalized Walsh transform of $w^T \vec{g}(\vec{x})$ then $(C(\vec{f}, \vec{g}))_{u,w} = \sum_v F^u(v) G^w(v)$.

**Theorem:** The correlation coefficients and spectrum values for a boolean function over $GF(2)$ are integer multiples of $2^{1-n}$. Proof: The values are of the form $k + (2^n - k)(-1) = 2k - 2^n$ which is even.

**Theorem:** The elements of a correlation matrix corresponds to an invertible transform of $n$-bit vectors are integer multiples of $2^{2-n}$. The proof uses the restriction map and the fact that $\sum (F(w) + F(w+v))^2 = 2$.

For $F_q, q = 2^n$, $Tr_{F_q/F_2}(x) = Tr(x) = \sum_{i=0}^{n-1} x^{2^i}$. Theorem: $Tr(x) \neq 0$ for some $x$. $Tr(x+y) = Tr(x) + Tr(y)$. $Tr(x^2) = Tr(x)$. $Tr(x) \in F_2$. $Tr(\omega x)$ is linear in $x$. $Tr(\omega_1 x) = Tr(\omega_2 x) \rightarrow \omega_1 = \omega_2$. $Tr(\omega x)$ are exactly the linear functions.

$F : F_{2^n} \rightarrow F_{2^m}$ is **differentially** $\delta$ **uniform** if $\forall \alpha, \beta, \alpha \neq 0$: $|\{x : F(x+\alpha) + F(x) = \beta\}| \leq \delta$. Theorem: $F(x) = x^{2^k+1}$, $s = (k, n)$ then $F$ is differentially $2^s$-uniform. $N(F) = 2^{n-1} - 2^{\frac{n+s}{2}-1}$. Theorem: Let $G(x) = x^{-1}, x \neq 0; 0, x = 0$. $F$ is differentially 4 uniform. $N(G) \geq 2^{n-1} - 2^{\frac{n}{2}}$.

$a \vee b = a \oplus b \oplus ab$ as a boolean function. Let $\vec{x} = (x_4, x_3, x_2, x_1)$ with $x_1$ the least significant bit. $\vec{F}(\vec{x}) = (F_4(\vec{x}), F_3(\vec{x}), F_2(\vec{x}), F_1(\vec{x}))$. If $\rho = (0000, 0001)$ then $\vec{F}^{\rho}_i(\vec{x}) = x_i, i > 1$ and $\vec{F}^{\rho}_1(\vec{x}) = (\overline{x_2 \vee x_3 \vee x_4})(x_1 \oplus 1) \oplus (x_2 \vee x_3 \vee x_4)x_1 = 1 \oplus x_1 \oplus x_2 \oplus x_3 \oplus x_4 \oplus x_2x_3 \oplus x_2x_4 \oplus x_3x_4 \oplus x_2x_3x_4$. If $\sigma = (0000, 0001, \ldots, 1111)$, then $\vec{F}^{\sigma}_1(\vec{x}) = x_1 \oplus 1$, $\vec{F}^{\sigma}_2(\vec{x}) = x_1(x_2 \oplus 1) \oplus \overline{x_1}x_2 = x_1 \oplus x_2$, $\vec{F}^{\sigma}_3(\vec{x}) = (x_1x_2)(x_3 \oplus 1) \oplus (\overline{x_1x_2})x_3 = x_1x_2 \oplus x_3$, $\vec{F}^{\sigma}_4(\vec{x}) = (x_1x_2x_3)(x_4 \oplus 1) \oplus (\overline{x_1x_2x_3})x_4 = x_1x_2x_3 \oplus x_4$.

**Discrete Fourier Transform and FFT:** Let $c(x) = a(x)b(x)$ which corresponds to the convolution $\vec{c} = \vec{a} * \vec{b}$. Define the DFT as $F(\vec{a}) = A\vec{a}$, $A = \omega^{ij}$ with inverse $A^{-1} = \frac{1}{n}\omega^{-ij}$. Note that $F(\vec{b} * \vec{c}) = F(\vec{b}) \cdot F(\vec{c})$ (pointwise multiplication). Tukey-Cooley Idea: Suppose $n = pq$, set $j = j(j_1, j_2) = j_1 q + j_2, k = k(k_1, k_2) = k_2 p + k_1, 0 \leq j_1 < p, 0 \leq j_2 < q, 0 \leq k_1 < p, 0 \leq k_2 < q$. Then $\hat{f}(k_1, k_2) = \sum_{j_2=0}^{q-1} e^{\frac{2\pi i j_2 (k_2 p + k_1)}{n}} \sum_{j_1=0}^{p-1} e^{\frac{2\pi i j_1 k_1}{p}} f(j_1, j_2)$. This requires $p^2 q$ and $q^2 p$ operations respectively or $pq(p+q)$ rather than $(pq)^2$. Now do this recursively if $p, q$ factor further. **Strassen and FFT:** For matrix multiply, Strassen found 7 products that do the trick: $m_1 = (a_{12} - a_{22})(b_{21} - b_{22})$, $m_2 = (a_{11} + a_{22})(b_{11} + b_{22})$, $m_3 = (a_{11} - a_{21})(b_{11} + b_{12})$, $m_4 = (a_{11} + a_{12})b_{22}$, $m_5 = a_{11}(b_{21} - b_{22})$, $m_6 = a_{22}(b_{21} + b_{11})$, $m_7 = (a_{21} + a_{22})b_{11}$. $c_{11} = m_1 + m_2 - m_4 + m_6$, $c_{12} = m_4 + m_5$, $c_{21} = m_6 + m_7$, $c_{22} = m_2 - m_3 + m_5 - m_7$. $T(n) = 7T(\frac{n}{2}) + 18\frac{n}{2}^2$, which is $O(2^{lg(7)})$. $F_{i,j} = \omega^{ij}$. $F$ evaluates, $F^{-1}$, interpolates. $q_{l,m} = \prod_{j=l}^{l+2^m-1}(x - c_j)$ and $q_{l,m} = a_{l,m-1} q_{l+2^m,m-1}$. What is $Rem(\frac{p(x)}{q_{l,0}(x)}), \forall l$? If $q = q'q''$, $Rem(\frac{p(x)}{q'(x)}) = Rem(\frac{r_{l,m}(x)}{q'(x)})$, $q_{l,m} = x^{2^m} = \omega^{rev(l/2^m)}$. For algorithm, crucial step is $r_{l,m}(x) = \sum (a_j + \omega^s a_{j+2^m}) x^j$ and $r_{l+2^m,m}(x) = \sum (a_j + \omega^{s+\frac{n}{2}} a_{j+2^m}) x^j$.

**Theorem:** $RM(r, m)$ has minimum distance $2^{m-r}$. $R(1, 5)$ has 48 inequivalent affine classes.

Each possible Boolean transformation on $n$ bits is a permutaion on the $2^n, n$-bit values and so listing them in order, the columns are the possible $\vec{f}$ vectors representing the component functions. If we label these as points in $GF(2)^{2^n}$ and draw an edge between allowable co-components with the edges labeled by the correlation between these vectors, any allowable $n$ boolean functions form a complete graph with the label

0 on each edge. $C(f,g) = 1 - \frac{wt(f+g)}{2^{n-1}}$. **Generalized Balance Theorem:** For each $n \leq 128$ and each $1 \leq b_1 < b_2 < \ldots < b_n \leq 128$ and fixed $\vec{k}$, $(E_{b_1}(\vec{k},\vec{x}), E_{b_2}(\vec{k},\vec{x}), \ldots, E_{b_n}(\vec{k},\vec{x}))$ takes each value in $\mathbb{Z}_2^n$ as $\vec{x}$ varies over $\mathbb{Z}_2^n$. So does any non-trivial sum of any of these functions. **Theorem:** If $f : GF(2)^{n-1} \rightarrow GF(2)$ is any boolean function, $g(x_1, \ldots, x_n) = f(x_1, \ldots, x_{n-1}) + x_n$ is balanced.

Write $\epsilon = E_K$ and $\epsilon' = E_{K'}$. What does $[\epsilon^i, \epsilon'^j]$ reveal about $K$ for known $K'$. Let $\mathcal{P} = \{p_1, p_2, \ldots, p_m\}$ and let $l$ be given put $N = p_1^l \ldots p_m^l$ and denote the set of $n$-bit elements of the block by $S$; what is $\mathbb{C}_S(\epsilon^N)$? How do you characterize the $x : g(x) = x$ where, say, $g$ represents $N$ applications of $\epsilon$. In general, $\epsilon$ is complicated but $\epsilon^m = 1$ for some $m$ and $\epsilon^t$ many be much simpler for some $m < t$. Let $g_{(i)}^{(0)}(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = f(x_1, x_2, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)$. **Idea:** Suppose $\epsilon^i$ and $\epsilon^j$ are relatively easy to determine (low degree, good approximation whatever) and $(i,j) = 1$ then we can find $a, b : ai + bj = 1$ and calculate $\epsilon = (\epsilon^i)^a (\epsilon^j)^b = \epsilon$. Let $B_n(r, \vec{v}) = \{\vec{x} : wt(\vec{v} \oplus \vec{x}) = r\}$. $|B_n(\vec{v}, r)| = 2^{n-r}$. Motivation for idea is while there are lots of "far away" approximations of $\epsilon$ there aren't many near ones. However, there may be close approximations of $\epsilon^i$.

Let $f$ is a Boolean Function define $S_f^0 = \{x : f(x) = 0\}$ and $S_f^1 = \{x : f(x) = 1\}$. If $e_i(x) = E_i(k, x)$ then $|S_{e_1}^b \cap S_{e_2}^b \cap \ldots \cap S_{e_k}^b| = 2^{n-k}$. What are the permutations that fix such a set?

Let $f, g : GF(2)^n \rightarrow GF(2)$ and $N = 2^n$. Let $f, g : GF(2)^n \rightarrow GF(2)$ and $N = 2^n$. $C(f,g) = 2Pr[f(x) = g(x)] - 1$. Let $a$ be the number of positions where $f$ and $g$ agree and $d$ be the number of positions where $f$ and $g$ disagree, then $Pr[(f(x) = g(x)] = \frac{a}{2^n}$. Note that $wt(f \oplus g) = d = dist(f,g)$. Now suppose $g(x) = w \cdot x$, the linear function. $F(w) = \frac{1}{2^n} \sum_x (-1)^{f(x) = g(x)} = \frac{1}{2^n}(a - d)$ Since $a + d = 2^n$, $F(w) = 2\frac{a}{2^n} - 1$ and thus $C(f, w) = F(w)$. These yield $dist(f(x), w \cdot x) = 2^n(1 - F(w))$. Thus the best affine approximation is the one which maximizes $|F(w)|$ for some $w$.

Now let $f : GF(2)^n \rightarrow GF(2)$ be a bijective boolean transformation with component functions $f_1, f_2, \ldots, f_n$. All such transformations represent permutations in $S_{2^N}$ and the correlation matrices of these transformations is orthogonal $(CC^T = I)$. A block cipher gives rise to such transformations by setting $f(x) = E_K(x)$ for fixed $K$. Note that all balanced boolean functions can be obtained by applying a permutation in $S_{2^N}$ to a sequence of $\frac{N}{2}$, 1's and $\frac{N}{2}$, 0's.

With the foregoing notation: **Theorem 1:** $C(f_i, 1) = C(f_i, 0) = 0$, $C(f_i, f_j) = 0, i \neq j$, $wt(f_i) = 2^{n-1}, \forall i$, $wt(f_i f_j) = 2^{n-2}, i \neq j$ and in general, $wt(f_{i_1} f_{i_2} \ldots f_{i_k}) = 2^{n-k}$. Further, $C(f_i f_j, f_k) = \frac{1}{2}$ , $C(f_i, f_j, f_k f_l) = C(f_i f_j f_k, f_l)$ and in general $C(f_{i_1} f_{i_2} \ldots f_{i_k}, f_l) = 2^{n-k-1}$. Let $f$ be a boolean function. **Theorem 2:** Let $f$ be a boolean function. The $N$ functions $f_{i_1} f_{i_2} \ldots f_{i_k}$ form a basis for the space of boolean functions; that is, for any boolean function $g$, $\exists a_{i_1, i_2, \ldots, i_k}^{(g)}$ such that $g(x) = \sum_{1 \leq i_1 < i_2 < \ldots < i_k = n} a_{i_1, i_2, \ldots, i_k}^{(g)} f_{i_1} f_{i_2} \ldots f_{i_k}$. In particular, there are such coefficients such that $x_i = \sum_{1 \leq i_1 < i_2 < \ldots < i_k = n} a_{i_1, i_2, \ldots, i_k}^{(x_i)} f_{i_1} f_{i_2} \ldots f_{i_k}$. Define $Appx_i(f) = \{g : dist(f,g) \leq i\}$, then $|Appx_i(f)| = \sum_{j=0}^{i} \binom{N}{i}$.

$NL(f) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$, $NL(f) \leq 2^{n-1} + \sqrt{2^n + max_{e \neq 0}(F(D_e(f)))}$, where $D_e f = f(x) \oplus f(x \oplus e)$.

**Theorem (Rothaus):** Let $n \geq 4$ of even algebraic degree then any bent function on $GF(2)^n$ has degree $\leq \frac{n}{2}$. An $n$-Boolean function, $f$, is $m$-resilient iff $f$ is balanced and $F(u) = 0, \forall u : wt(u) \leq m$. Maiorana-MacFarland class $\mathcal{M} = \{f : f(x,y) = x\pi(y) \oplus g(y)\}$ where $\pi$ is a permutation on $GF(2)^{\frac{n}{2}}$ and $g$ is affine. $|\mathcal{M}| = (2^{\frac{n}{2}})! 2^{\frac{n}{2}}$. For **Bent Quadratics:** $\bigoplus_{1 \leq i,j \leq n} a_{ij} x_i x_j \oplus h(x)$, $h$, affine.

For this section, $f : GF(2)^m \rightarrow GF(2)$. The sensitivity of $v$ is defined by $S(v) = |\{v' : f(v) \neq f(v'), dist(v, v') = 1\}|$. The average sensitivity $aS(f) = \frac{1}{2^m} \sum_v S(v)$. The "influence" of $x_i$ is defined by
$I(x_i) = Prob(f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_m)$, the probability that the function is determined no matter what $y$ is.

**Theorem:** Let $f$ be a boolean function of $n$ variables with average sensitivity $aS(f) = k$. Let $\epsilon > 0$ and $M = \frac{k}{\epsilon}$ then (1) $\exists h$ depending on $exp((2 + \sqrt{\frac{2log(4M)}{M}})M)$ variables such that $Prob(f \neq h) \leq \epsilon$; and,

(2) $\exists g$ of degree at most $exp((2+\sqrt{\frac{2log(4M)}{M}})M)$ such that $Prob(f \neq g) \leq \frac{\epsilon}{2}$.

Basic Question: Let $F$ be a family of $m$ binary $n$-vectors. How densely packed is $F$?. Given $b \leq n$, $|F|$, what is the largest possible number of pairs of vectors in $F$ whose Hamming distance is less than $b$?

Trace and correlation in $GF(2^n)$: $C^f_{u,w} = 2^{-n} \sum_a (-1)^{Tr(wa)}(-1)^{Tr(uf(a))}$ so the terms are determined by the condition $Tr(wa+uf(a))=0$, if this is satisfied by $r$ values the entry is $r2^{1-n}$. If a function is linear over $GF(2^n)$, it is linear over $GF(2)$ but not vice versa.

**Theorem:** Let $r_n$ be the ratio of the number of invertible $n \times n$ matrices over $GF(2)$ to the number of $n \times n$ matrices over $GF(2)$, then $lim_{n \rightarrow \infty}(r_n) \approx 0.288$. Proof: The number of invertible $n \times n$ boolean matrices is $t_n = (2^n-1)(2^n-2)\ldots(2^n-2^{n-1})$. The number of $n \times n$ boolean matrices is $2^{n^2}$. $t_n = 2^{\frac{n(n-1)}{2}}(2^n-1)(2^{n-1}-1)\ldots(2-1)$. Define $s_n = (2^n-1)(2^{n-1}-1)\ldots(2-1)$. Now $t_{n+1} = 2^{\frac{n(n+1)}{2}}s_{n+1} = 2^{\frac{n(n+1)}{2}}2^{-\frac{n(n-1)}{2}}(2^{\frac{n(n-1)}{2}}s_n)(2^{n+1}-1) = 2^n(2^{n+1}-1)t_n$. Dividing both sides of this by $2^{(n+1)^2}$, we get $r_{n+1} = \frac{t_{n+1}}{2^{(n+1)^2}} = \frac{2^n}{2^{2n+1}}\frac{t_n}{2^{n^2}}(2^{n+1}-1) = r_n(1-2^{-(n+1)}))$. Using this recurrence, we get $r_n = \prod_{i=1}^n(1-2^{-n})$. The product approaches $\approx 0.288$ as $n \rightarrow \infty$.

### 1.2.8 Computational Algebra

**Hensel:** If $I \subseteq \mathbb{R}$, $f = gh \pmod{I}$ such that the pseudo $GCD(g,h) = 1$ then $\exists g^*, h^*$ such that (1) $f = g^*h^* \pmod{I^2}$, (2) $g = g^* \pmod{I}$, (3) $h = h^* \pmod{I}$, and pseudo $GCD(g^*,h^*) = 1 \pmod{I^2}$. If $g', h'$ satisfy the conditions also, $g' = g^*(1+u) \pmod{i^2}$ and $h' = h^*(1-u) \pmod{i^2}$.

**Bivariate Factoring:** If $|\mathbb{F}| > 4d^2$, $f \in \mathbb{F}$, $deg_x(f) \leq d$, $\exists \in \mathbb{F}$: $f_\beta(x,0) \in \mathbb{F}[x]$ has no repeated factors.

- 1a Obtain square free factorization
- 1b Find $\beta \in \mathbb{F}$ such that $f(x,\beta)$ is squarefree.
- 1c $f_\beta = f(x, y+\beta)$.
- 2a $f(x,y) = g(x,y)h(x,y) \pmod{y}$
- 2b Lift $f(x,y) = g_k(x,y)h_k(x,y) \pmod{y^k}$
- 3a Find $g''$ and $l_k$: $g'' = g_kl_k \pmod{y^{2^k}}$, $deg_x(g'') \leq deg_x(f)$, $deg_y(g'') \leq deg_y(f)$, $g'' \neq 0$.
- 3b Find $gcd(f,g)$ as polynomials in $F(y)[x]$.

$|Res(f,g,x)| \leq (m+1)^{\frac{n}{2}}(n+1)^{\frac{m}{2}}A^{\frac{m}{2}}B^{\frac{n}{2}}$.

**Extension Theorem:** Let $I =< f_1, ..., f_s > \in \mathbb{C}(x_1, x_2, ..., x_n)$ and $I_1$ is the first elimination ideal of $I$. For each $1 \leq i \leq s$ write $f_i = g(x_2, ..., x_n){x_1}^{N_i} + ....$ Suppose $c = (c_2, ..., c_n) \in V(I_1)$. If $c \notin V(g_1, g_2, ..., g_s)$, $\exists c_1$ such that $(c_1, c) \in V(I)$.

**Linear Programming:** $max(cx)$ subject to $Ax \leq b$, $x \geq 0$. **Quadratic Programming:** $max(\sum \rho_{ij}\sigma_i\sigma_jx_ix_j)$, subject to $\sum x_i = 1$, $x_i \geq 0$, $\sum x_iu_i \geq R$.

### 1.2.9 Algebraic Number Theory

**Gaussian Integers:** $\mathbb{Z}[i]$. Let $\alpha, \beta, \gamma, \delta$ represent gaussian integers. $N(x+yi) = x^2+y^2$. $\forall \alpha, \beta, \exists \gamma, \delta$ such that $\alpha = \beta\gamma + \delta$ with $0 \leq N(\delta) < N(\beta)$. $\alpha$ is a unit iff $N(\alpha) = 1$. Units are $1, -1, i, -i$. Let $S = \{\alpha\eta + \beta\gamma\}$, $\phi$ with minimal norm is the gcd. If $\pi$ is a Gaussian integer with $N(\pi) = p$ then $\pi$ is prime. If $\pi$ is a Gaussian prime and $\pi|\alpha\beta$ then $\pi|\alpha$ or $\pi|\beta$. Gaussian integers form a UFD. Let $\pi$ be a Gaussian prime, there is one and only one $p$ such that $\pi|p$. Note that $\pi = x+yi$, $N(\pi) = x^2+y^2$ divides $p$ or $p^2$ so $x = 0, 1, 2 \pmod{4}$. Characterization of Gaussian primes: $p = 2$: $p = -i\pi^2$. $p = 3 \pmod{4}$, $p = \pi$. $p = 1 \pmod{4}$, $p = \pi\overline{\pi}$ and $\pi$ and $\overline{\pi}$ are non-associated primes. If $p = 1 \pmod{4}$ then $p \mid (z^2+1)$. If $\pi \mid p$, $\pi|(z+i)(z-i)$ so $\pi|(z-i)$.

$x$ is **integral** over $A$ if $x$ is a root of a monic polynomial $f$ with coefficients in $A$. If $A$ is a subring of $R$, the **integral closure** of $A$ in $R$ is the set $A_c$ of elements of $R$ that are integral over $A$. Note that $A \subseteq A_c$. We say $A$ is integrally closed in $R$ if $A_c = A$. If $A$ is an integral domain with quotient field $K$, and $A$ is integrally closed in $K$ we simply say that $A$ is integrally closed without reference to $R$.

**Integral Ring Extensions:** Let $M$ be an $A$-module. $M$ is faithful if $aM = 0 \rightarrow a = 0$. Let $A \subseteq B$, $\alpha \in B$. The following are equivalent: (1) $\alpha$ is a root of $f(x) = x^n + a_{n-1}x^{n-1} + ... + a_0$; (2) $A[\alpha]$ is a finitely generated $A$ module; (3) $\exists$ a faithful module over $A[\alpha]$ which is a finitely generated $A$-module.

If $A$ is an entire ring and a UFD then it is integrally closed. If $B$ is integral over $A$ and $\phi$ is an embedding of $A$ into its algebraic closure, $L$, $\phi$ extends to $B$.

$N_{E/F}(x) = det(m(x))$, $Tr_{E/F}(x) = trace(m(x))$. If $\alpha = x + yi$, $Tr(\alpha) = 2x$, $N(\alpha) = \alpha\overline{\alpha}$. $S(\alpha) = \sum_\sigma \alpha^\sigma$ is an integer, so is $N(\alpha) = \prod_\sigma \alpha^\sigma$. $\alpha$ is a unit iff $|N(\alpha)| = 1$. $\alpha$ is an integer of $Q(\sqrt{d})$ iff $T(\alpha)$ and $N(\alpha)$ are integers.

**Quadratic integers:** $I_d = \{x + y\omega_d, x, y \in \mathbb{Z}\}$, $\omega_d = \sqrt{d}$ if $d = 2, 3 \pmod 4$, $\frac{1+\sqrt{d}}{2}$, if $d = 1 \pmod 4$. Ideal Theory: $P = (2, 1 + \sqrt{-5})$, $Q = (3, 1 + \sqrt{-5})$. $P^2 = (2)$ and $Q\overline{Q} = (3)$. Fermat analogue: $\alpha^{N(\pi)-1} = 1 \pmod \pi$.

Rational algebraic integers are integers. If $\theta$ is an algebraic number, there is an integer $n$ such that $n\theta$ is an algebraic integer. Every basis for $R(\theta)$ has $n$ elements. $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_n) = det({\alpha_i}^{\sigma_j})^2$. Alternatively, $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_n) = det(S(\alpha_i\alpha_j))$. $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is an integer. If $\{\alpha_i\}$ and $\{\beta_i\}$ are basis with $\alpha_j = \sum_k a_{jk}\beta_k$ then $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_n) = det(a_{ij})^2\Delta(\beta_1, \beta_2, \ldots, \beta_n)$. $\{\alpha_i\}$ is a basis iff $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_n) \neq 0$. If $\{\alpha_i\}$ is an integral basis for $R(\theta)$ then $\Delta(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is minimal, in which case it is called the **discriminant** of $R(\theta)$ and written $Disc(R(\theta))$. All integral basis have the same discriminant.

**Every ideal contains a basis.** If $A$ is an ideal of $Q(\theta)$ then $\mathbb{Z} \cap A \neq \emptyset$. If $A, B$ are ideals in $R(\theta)$, $A|B$ iff $A = BC$ iff $B \subseteq A$. If $A$ is an ideal in $R(\theta)$, $\exists B$ such that $AB = (a)$ for some $a$ in $R(\theta)$. If $A, B$ are ideals in $R(\theta)$,with $AC = BC$ then $A = B$. If $P|AB$ and $P$ does not divide $A$ then $P|B$. Every ideal has finitely many distinct divisors. Every prime ideal must divide the principal ideal of a rational prime. Every ideal can be written as a product of prime ideals. The factorization is unique apart from order. Every rational integer belongs to finitely many ideals. Rational prime is ramified if its principal ideal factors into prime ideals in which one prime ideal is repeated. If this happens, $p|\Delta(\alpha_1, \ldots, \alpha_n)$.

**An ideal has finitely many divisors.** If $A$ is an ideal with basis $\alpha_i = \sum_j a_{ij}\omega_j$ then $N(A) = det(a_{ij})$. $A \sim B$ iff $\exists \alpha, \beta$ such that $(\alpha)A = (\beta)B$. each equivalence class is called an ideal class. There are finitely many ideal classes $h$ of $R(\theta)$ and $A^h \sim (1)$. Proof: For $K = R(\theta)$, $\exists C(K) : \forall A, \exists 0 \neq \alpha \in a : |N(\alpha)| \leq C(N(A)$. Use this to show $\exists B : N(B) \leq C$ so there are a finite number of ideals containing $B$. $\exists \alpha : (\alpha) = AD$. $AN \sim AD$.

The ring of integers $\mathcal{D}_K$ in the number field, $K$, has the following properties: $\mathcal{D}_K$ is a domain with field of fractions $K$. $\mathcal{D}_K$ is noetherian (Use the fact that $\mathcal{D}_K$ is a free abelian group of degree $n = K : \mathbb{Q}$.) A $\mathcal{D}$, $\mathfrak{a}$ is a fractional ideal if $\exists c \in \mathcal{D}$: $c\mathfrak{a} \subseteq \mathcal{D}$. Every non zero prime ideal $\mathfrak{p}$ of $\mathcal{D}$ is maximal. ( $\mathcal{D}/\mathfrak{a}$ is a finite integral domain.) Fractional ideals form an abelian group. Every non-zero ideal of $\mathcal{D}$ can be factored into prime ideals ($\mathcal{D}$ is noetherian).

**Norm of an ideal:** $N(\mathfrak{a}) = |\mathfrak{D}/\mathfrak{a}|$; if $\mathfrak{a} =< a >$ is principal $N(a) = N(\mathfrak{a})$. $N(\mathfrak{ab}) = N(\mathfrak{a})N(\mathfrak{b})$. $\Delta_{K/\mathbb{Q}}(\alpha_1, \alpha_2, \ldots, \alpha_n) = [det(\sigma_i(\alpha_j))]^2$. Every non-zero ideal of $\mathcal{D}$ has a finite number of divisors. A non-zero rational integer belongs to a finite number of ideals of $\mathcal{D}$. Only a finite number of ideals of $\mathcal{D}$ have a given norm. If $\mathfrak{a} \neq \mathfrak{b}$ are ideals of $\mathcal{D}$ then $\exists \alpha \in \mathfrak{a} : \alpha\mathfrak{a}^{-1} + \mathfrak{b} = \mathcal{D}$. Let $\mathfrak{a} \neq 0$ be an ideal of $\mathcal{D}$ and $0 \neq \beta \in \mathfrak{a}, \exists \alpha \in \mathfrak{a} : \mathfrak{a} =< \alpha, \beta >$.

**Minkowski:** $X$ is convex if $x, y \in X \rightarrow \lambda x + (1 - \lambda)y \in X, \forall \lambda \in [0, 1]$. $X$ is symmetric if $x \in X \rightarrow -x \in X$. Let $L$ be an $n$-dimensional lattice in $\mathbb{R}^n$ with fundamental region $T$ and let $X$ be a bounded, convex, symmetric subset of $\mathbb{R}^n$; if $v(X) > 2^n v(T), \exists \alpha \in X \cap L, x \neq 0$. Let $L$ be a lattice then $\mathbb{R}^n/L \cong T^n$ (a torus).

Let $T$ be a fundamental region of $L$, $\phi : T \rightarrow T^n$ then $v(X) = v(\phi^{-1}(X))$. If $\nu : \mathbb{R}^n \rightarrow T^n$ is the natural homomorphism with $ker(\nu) = L$. If $X$ is a bounded subset of $\mathbb{R}^n$, $\nu$ exists and $v(\nu(X)) \neq v(X)$ then $\nu_{|X}$ is not injective. Four squares: If $p = 4k + 1$ then $p = a^2 + b^2$. ($< g >= \mathbb{Z}_p$ is cyclic $g^k = u$ and $u^2 = -1$. Let $L = \{(a, b) : b = ua \pmod{p}\}$, $\mathbb{Z}^2 : L = p^2$, $vol(T_L) = p$. $C_r : \{x : ||x|| < r\}$ and $\pi r^2 > 4p$, $r^2 = \frac{3p}{2}, 0 \neq a^2 + b^2 \leq r^2 < 2p$.

Examples in algebraic fields: In $R = \mathbb{Z}[\sqrt{-3}]$: $\frac{-1+\sqrt{-3}}{2}$ is a unit note that $2 \times 2 = -1 + \sqrt{-3} \times -1 - \sqrt{-3}$. In $R = \mathbb{Z}[\sqrt{-5}]$ ideals are not all principal; note that $2 \times 3 = -1 + \sqrt{-5} \times -1 - \sqrt{-5}$. Pell related: There are two equivalence classes of forms of determinant 5: $x^2 + 5y^2$ and $2x^2 + 2xy + 3y^2$ and the class number of $\mathbb{Z}[\sqrt{-5}]$ is 2. If $p$ is a rational prime and $K/\mathbb{Q}$ is a Galois extension then $G = \mathcal{G}(K/\mathbb{Q})$ acts transitively on the ideal divisors of $(p)$, the exponent of the ideal divisors are called the ramification index. The ideal generated by a rational ideal $(p)$ factors into indecomposable factors in an algebraic number field, $O_F$, in one of three ways: (a) $(p)$, (b) $(p) = P\sigma(P)$ ("$p$ splits"), or (c) $(p) = P^2$ ("$p$ ramifies").

**Analytic formulas:** $f * g(n) = \sum_{d|n} f(d)g(\frac{n}{d})$. This is commutative, associative and has an inverse. $\Lambda(n) = ln(n)$, if $n = p^m$, $\Lambda(n) = 0$, otherwise. Note: $ln(n) = \sum_{d|n} \Lambda(d)$. $\sigma_\alpha(n) = \sum_{d|n} d^\alpha$. $\psi(x) = \sum_{n \leq x} \Lambda(n)$, $\vartheta(x) = \sum_{p \leq x} ln(p)$. $\frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \leq \frac{ln(x)^2}{2\sqrt{x} ln(2)}$. $L(1, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n}$, $\chi$, a non-principal character. **Dirichlet:** If $k > 0$ and $(h, k) = 1$, $\forall x > 1$ $\sum_{p \leq x, p=h \pmod{k}} \frac{ln(p)}{p} = \frac{1}{\phi(k)} ln(x) + O(1)$. $\pi_a(x) = \sum_{p \leq x, p=a \pmod{k}} 1$. $\pi_a(x) \approx \frac{\pi(x)}{\phi(k)}$, $x \rightarrow 0$, $\forall a$, $(a, k) = 1$ and $\pi_a(x) \approx \pi_b(x)$ when $(a, k) = (b, k) = 1$.

### 1.2.10 Group Theory

**Isomorphism Theorems:** (1) If $\varphi : G \to H$ is a homomorphism, $G/ker(\varphi) \cong Im(\varphi)$, (2) If $G \rhd H$ and $G \rhd N$ and $N \subseteq H \subseteq G$ then $G/H \cong (G/N)/(H/N)$, (3) If $G = HN$, $G \rhd N$ then $HN/N \cong H/(H \cap N)$.

**Derived series:** $G^{[0]} = G$, $G^{[i+1]} = [G, G^{[i]}]$. $G$ is solvable iff derived series terminates at 1. Subnormal Series: $G = G_0 \rhd G_1 \rhd \ldots \rhd G_k = H$, if this happens, we say $G \rhd \rhd H$. **Normal series:** Subnormal series where $G \rhd G_i, \forall i$. **Chief series:** Normal series with no repeated terms and no normal subgroup properly lying between two series elements. Zassenhaus. If $A \lhd A^*$ and $B \lhd B^*$ then $A(A^* \cap B) \lhd A(A^* \cap B^*)$ and $B(B^* \cap A) \lhd B(B^* \cap A^*)$; further, $\frac{A(A^* \cap B^*)}{A(A^* \cap B)} \cong \frac{B(B^* \cap A^*)}{B(B^* \cap A)}$. The following are equivalent: (1) $G$ is solvable, (2) $G$ has a normal series terminating at the identity whose factor groups are cyclic of prime order, (3) $G$ has a subnormal series with abelian quotients.

**Schreier:** Two normal series for $G$ have equivalent refinements. Two compositions series for $G$ are equivalent. Proof: By induction on length ($l$) of shortest such series. If $l = 1$, $G$ is simple. Suppose $G = G_0 \geq G_1 \geq \ldots \geq G_r = 1$ and $H = H_0 \geq H_1 \geq \ldots \geq H_t = 1$ and assume $l = r > t$ and that the theorem is true for all series of length less than $l$. If $H_1 = G_1$ then we are done by induction on the shortened series. Assume $G_1 \neq H_1$, $H_1 \lhd G, G_1 \lhd G$ then $G_1 H_1 = G$ and $G/G_1 \cong H_1/K, K = G_1 \cap H_1$. Consider the two series $G_1 \geq G_2 \ldots \geq G_r = 1$ and $G_1 \geq K \geq K_1 \ldots \geq K_t = 1$. By induction, $r - 1 = t + 1$ and they are equivalent. Thus, $H_1 \geq H_2 \ldots \geq H_s = 1$ and $H_1 \geq K \geq K_1 \ldots \geq K_{r-2} = 1$ so $r = s$ and they are equivalent.

$\phi$ is a **normal endomorphism** iff $\phi(a^{-1}xa) = a^{-1}\phi(x)a$, $\forall x, a \in G$. Lemma 1: If $G$ satisfies ACC or DCC then $G$ is the direct product of indecomposable groups. Lemma 2: If $G$ satisfies ACC (resp. DCC) on normal subgroups and $f$ is a normal endomorphism of $G$, then $f$ is an automorphism iff $f$ is an epimorphism (resp automorphism). Lemma 3 (**Fitting**) Let $G$ satisfy both chain conditions. If $\phi$ is a normal endomorphism of $G$ then $G = Ker(f^n) \times Im(f^n)$, some $n \geq 1$. If $G$ is an indecomposable group satisfying ACC and DCC on normal subgroups and if $f$ is a normal endomorphism then $f$ is nilpotent or an automorphism. **Krull-Schmidt:** If $G$ has both chain conditions on normal subgroups and $G = H_1 \times \ldots \times H_s = K_1 \times \ldots \times K_t$ are two decompositions into indecomposable factors then $s = t$ and, after reindexing, $H_i \cong K_i$ and for each $r < t$, $G = G_1 \times G_2 \times \ldots \times G_r \times H_{r+1} \times H_t$. Proof: Let $P(0)$ be the statement $G = G_1 \times G_2 \times \ldots \times G_s$ and for $1 \leq r \leq min(s,t)$ let $P(i)$ be the statement $G = G_1 \times G_2 \times \ldots \times G_r \times H_{r+1} \times \ldots H_t$. $P(0)$ is true by assumption, assume $P(r-1)$. Let $\pi_i$ (resp $pi'_i$ be the canonical epimorphisms from $G_1 \times G_2 \times \ldots \times G_s$ (resp. $G_1 \times G_2 \times \ldots \times G_r \times H_{r+1} \times H_t$ and $\lambda_i$ (resp $\lambda'_i$) be the inclusion maps, $\varphi_i = \lambda_i \pi_i$ and $\phi_i = \lambda'_i \pi'_i$. $\varphi_r \phi_i = 0_{|G}$ for $i < r$ and $\varphi_1(1_{|G}) = \varphi_r \phi_1 + \ldots + \varphi_r \phi_t = \varphi_r \phi_r + \ldots + \varphi_r \phi_t$ so $(\varphi_r \phi_j)_{|G}$ is an automorphism of $G_r$. $\varphi_j \phi_r$ must be an automorphism of $H_j$ and $\phi_j : G_r \to H_j$ is and isomorphism and so is $\varphi_r : H_j \to G_r$ reindexing we have the first half of $P(r)$. Let $g = g_1 g_2 \ldots g_{r-1} h_r h_{r+1} \ldots h_t$ define $\theta(g) = g_1 g_2 \ldots g_{r-1} \varphi(h_r) h_{r+1} \ldots h_t$. $G = Im(\theta) = G^* = G_1 \times G_2 \times \ldots \times G_r \times H_{r+1} \times H_t$ which completes the argument.

**Lower Central Series:** $L_1(G) = G$, $L_{n+1}(G) = [L_n(G), G]$. $G$ is **nilpotent** if $L_n(G) = 1$ for some $n$. Note $L_n(G)/L_{n+1}(G) \subseteq Z(G/L_{n+1}(G))$. **Upper Central series:** $Z_0(G) = 1$; Let $H^* = H/Z_n(G)$, define $Z_{n+1}(G)^* = Z(G/Z_n(G))$. Upper and Lower central series have same length. Finite nilpotent groups are direct products of their Sylow subgroups.

$G$ is an **extension** of $K$ by $Q$ if $G \rhd K$ and $G/K \cong Q$. If $1 \to N \to_i G \to_\varphi Q \to 1$, the following are equivalent (1) $\exists Q^* \subseteq G : Q^* \to Q$ and (2) $\exists s : Q \to G$ such that $\varphi \cdot s = id$. (3) $G$ is a semi-direct product of $N$ by $Q$ written $N \ltimes Q$; in this case, we say $G$ is a split extension of $N$ by $Q$.

$G$ is **complete** if it is centerless and every automorphism is inner. in which case $G \cong Aut(G)$. $S_n$ is complete if $n \neq 2, 3$. Proof: Let $T_k$ be the set of $k$ disjoint transpostions so $x \in T_k \to x^2 = 1$; note that if $\theta \in Aut(S_n), \theta(T_1) = T_k$ for some $k$. Also observe that $\theta$ preserves transpositions iff $\theta \in Inn(S_n)$. Now we can show $|T_1| = \frac{n(n-1)}{2}$ and $|T_k| = \frac{(n-2k+1)!}{(n-2k)!k!2^k}$. Comparing the two $|T_1| = |T_k|$ is possible only if $k = 2, 3$ and in fact, only if $k = 3$. If $\theta \in Out(S_6)$ and $\tau$ is a transposition, $\theta(\tau)$ must be a product of three transpositions and such an automorphism exists. If $G$ is a non-abelian simple group, then $Aut(G)$ is complete. If $K \lhd G$ and $K$ is complete, $G = K \times Q$. $Hol(K) \subset S_K$ is $< K^l, Aut(K) >$, $K^l \lhd Hol(K)$, $Hol(K)/K^l \cong Aut(K)$ and $C_{Hol(K)}(K^l) = K^r$. If $K$ is a direct factor whenever $K$ is a normal subgroup then $K$ is complete.

Suppose $G$ is an extension of $N$ by $H$ and let $\phi : H \to G/N$. Pick $s : G \to H$ such that $s(1) = 1$

and $\phi(h) = Ns(h)$, then $\exists f : H \times H \to N : s(h_1h_2) = f(h_1,h_2)s(h_1h_2)$ and $f(h_1,h_2)f(h_1h_2,h_3) = f(h_2,h_3)^{s(h_1)}f(h_1,h_2h_3)$. Note that $\theta_h : n \mapsto s(h)ns(h)^{-1}$ is in $Aut(N)$ and $\theta_{h_1}(\theta_{h_2}(n)) = \theta_{h_1h_2}(n)^{f(h_1,h_2)}$. Given $N, H$ with $\theta_h \in Aut(N)$ and $\theta_1 = 1$ and a map $f : H \times H \to N$ with $f(1,h) = f(h,1) = 1$ and $f(h_1,h_2)f(h_1h_2,h_3) = \theta_{h_1}(f(h_2,h_3))f(h_1,h_2h_3)$, suppose $f$ is compatible in the sense that $\theta_{h_1}(\theta_{h_2}(n) = \theta_{h_1h_2}(n)^{f(h_1,h_2)}$ then the operation $(n_1,h_1)\cdot(n_2,h_2) = (n_1\theta_{h_1}(n_2)f(h_1,h_2), h_1h_2)$ defines a group $G$ which is an extension of $N$ by $H$.

Suppose $T$ is a subset consisting of a representative of each coset of an $G/K$ which is called a **transversal**. If $\pi : G \to Q$ is a surjective homomorphism with kernel $K$, $l : Q \to G$ is a **lifting** if $\pi(l(x)) = x$. $G$ realizes $(Q, K, \theta)$ with $K' = 1$, $\theta : Q \to Aut(K)$ and $l : Q \to G$ if $G$ is an extension of $K$ by $Q$ and every transversal $l : Q \to G$ satisfies $xa = \theta_x(a) = l(x) + a - l(x)$. Note additive notation for non-abelian operation. If $\pi : Q \to G$ is a surjective homomorphism with kernel $K$ and $l : Q \to G$ is a transversal with $l(1) = 0$ then $f : Q \times Q \to K$ defined by $l(x) + l(y) = f(x,y) + l(xy)$ is called a **factor set**. **Cocycle identity:** $xf(y,z) - f(xy,z) + f(x,yz) - f(x,y) = 0$. Note $xf(y,z) = l(x)f(y,z)l(x)^{-1}$. Given "data," $(Q,K,\theta)$, $f : Q\times Q \to K$ is a factor set iff it satisfies the cocycle identity and $f(1,y) = 0 = f(x,1)$. Proof: Let $G = \{(a,x) : a \in K, x \in Q\}$. With $(a,x)+(b,y) = (a+xb+f(x,y), xy)$. This is a group if the conditions hold.

Let $G$ realize $(Q,K,\theta)$ and $l$ and $l'$ be transversals with $l(1) = l'(1) = 0$ giving rise to factor sets $f$ and $f'$ then there is an $h : Q \to K$ with $h(1) = 0$ such that $f'(x,y) - f(x,y) = xh(y) - h(xy) + h(x), \forall x, h \in Q$ and $g$ is called a **coboundary**. The set of all coboundaries is $\mathbf{B^2(Q,K,}\theta)$. $\mathbf{Z^2(Q,K,}\theta)$ is the set of all **factor sets**. $\mathbf{H^2(Q,K,}\theta) \cong \mathbf{Z^2(Q,K,}\theta)/\mathbf{B^2(Q,K,}\theta)$. Two extensions are equivalent if the difference of their two factor sets is in $B^2(Q,K,\theta)$. There is a bijection from $H^2(Q,K,\theta)$ and the set of equivalence classes of extensions realizing $(Q,K,\theta)$ taking 0 to the class of the semidirect product. See proof of Schur-Zassenhaus.

$G$, an extension of $K$ by $Q$, is a **central extension** if $K < Z(G)$. Functorially, a central extension $G$ is a pair $(H,\pi)$ satisfying $\pi : H \to G, ker(\pi) \subseteq Z(H)$. A cyclic extension $G$ of $N$ is one where $G/N$ is cyclic. Solvable groups are built from cyclic extensions. $\alpha : (H_1,\pi_1) \to (H_2,\pi_2)$ is a morphism in this category. If $(\tilde{G},\tilde{\pi})$ is universal if $\forall(H,\sigma), \exists!\alpha : (\tilde{G},\tilde{\pi}) \to (H,\sigma)$. $G$ possesses a universal central extension iff $G$ is perfect. If $(\tilde{G},\pi)$ is a universal central extension then $ker(\pi)$ is the Schur multiplier. **Homological version:** If $G > N$ and $H > K$ are normal subgroups isomorphic under $\phi$, the pullback is $(g,h)$ where $gN = \phi(hK)$. $(Q,K,\theta)$ is trivial iff every extension realizing $(Q,K,\theta)$ is a central extension. There's a bijection between $H^2(Q,K,\theta)$ and central extensions. **Schur multiplier:** $M(Q) = H^2(Q,\mathbb{C}^\times)$ ($\theta$ is trivial). Here $f(1,y) = f(x,1) = 1$, $f(x,y)f(xy,z)^{-1}f(x,yz)f(x,y)^{-1} = 1$, $g : Q \times Q \to \mathbb{C}^\times$ is a coboundary iff $\exists h : Q \to \mathbb{C}^\times$ with $h(1) = 1$ such that $g(x,y) = h(y)(h(xy))^{-1}h(x)$. Assume $G$ is perfect then a central extension $(E,\phi)$ of $G$ is universal iff (a) $E$ is perfect and (b) all central extensions of $E$ are trivial. In that case, $1 \to R \to F \to G \to 1$, $F$, free and $E = [F,F][F,R] \to [F,F]/R = G$.

**Central Product:** $G =< G_i >$, $[G_i,G_j] = 1$ for $i \neq j$. Equivalently, $\rho : (x_1,x_2,\ldots,x_n) \mapsto x_1x_2\ldots x_n$ is a surjective homomorphism from $(G_1 \times G_2 \times \ldots \times G_n)$ to $G$ with $\rho(D_i) = G_i$ where $\pi_i(G_1,\ldots,G_n) = D_i$ and $ker(\rho) \cap D_i = 1$, $ker(\rho) \subseteq Z(G)$. Let $Z < Z(A)) \cap Z(B)$, $A \times B/Z$ is a central product. Both $D_8$ and $Q_8$ are central products of $Z_2$ by $Z_2 \times Z_2$. Let $G_i, 1 \leq i \leq n$ be a family of groups with $Z(g_1) = Z(G_i)$ and $Aut_{G_i}(Z(G_i)) = Aut(Z(G_i))$. Then up to isomorphism there is a unique central product with $Z(G_1) = Z(G_i)$.

**Wreath Product:** $G^* = G^X$ - maps from $X$ to $G$. $fg(x) = f(x)g(x)$. Let $H$ act on $X$: $f^h(x) = f(xh^{-1})$. Let $\phi$ be the natural action of $H$ induced on $G^{|H|}$, then $G \wr H = H \rtimes_\phi G^*$. If $G_x = \{f : f(y) = 1\ if\ x \neq y\}$. $G^* = \prod_X G_x$. Put $g_x(y) = g(y)$ if $x = y$, 1 otherwise. Note that ${g_x}^h = g_{xh}$. If $H$ is finite and $G/K = H$, $G$ can be embedded in the regular wreath product $K \wr H$: **Universal Embedding Theorem:** Let $G \rhd N$ and $K \equiv G/N$, $\exists\phi : G \to N\ \wr\ K$ such that $\phi$ maps $N$ onto $im(\phi) \cap \bigotimes_i N$. $exp(G) = min\{e : x^e = 1, \forall x \in G\}$. If $Q$ is finite then $M(Q)$ is a finite abelian group and $exp(M(Q)) \mid |Q|$.

**Representations (1):** $M$ is a simple $R$-module if it has no non-trivial submodules. Let $M$ be a non-zero $R$-module. The following are equivalent ("**semisimple**"): (1) $M$ is a sum of simple modules, (2) $M$ is a direct sum of simple modules, (3) if $N \subseteq M$ is a submodule, there is another submodule $N'$: $M = N \oplus N'$. **Schur:** If $f \in Hom_R(M,N)$ and $M, N$ are simple then $f = 0$ or is an isomorphism. If $M$ is simple, $A = End_R(M)$ is a division ring. Let $M$ be a semi-simple $R$-module, $A = End_R(M)$ and $f \in End_A(M)$, if $m \in M, \exists r \in R : f(m) = rm$. $End_R(V^n) \cong M_n(End_R(V))$. **Jacobson:** Let $M$ be a semi-simple $R$-module,

$A = End_R(M)$ and $f \in End_A(M)$, $m_1, m_2, \ldots, m_n \in M, \exists r \in R : f(m_i) = rm_i$. Corollary: Let $M$ be a faithful simple $R$-module, $D = End_R(M)$, if $M$ is finite dimensional over $D$ then $End_D(M) \cong R$. If $R$ is a ring and $I$ is an ideal, say $I$ is simple if it is simple as a left module of $R$; say $I$ is semi simple if it is semi-simple if it is semi-simple as a module and all simple left ideals are isomorphic. If $R$ is a semi-simple ring then all non-zero $R$ modules are semi-simple. Let $I$ be a simple left ideal in a semi-simple ring $R$ and let $M$ be a simple $R$-module; either $IM = M$ and $I \cong M$ or $IM = 0$.
From now on let $R$ be a semisimple ring. $B_i = \sum_{I \subset R, I \cong I_i} I$. If $I_j$ is not isomorphic to $I_k$ then $B_j B_k = 0$; $R = \sum B_i$ and each $B_i$ is a two-sided ideal. There are only finitely many isomorphism classes of left ideals. If $R = \bigoplus_{i=1}^t B_i$ and $1 = \sum_{i=1}^t e_i$. If $b_i \in B_i$ then $e_i b_i = b_i = b_i e_i$ and $B_i = Re_i$. Each $B_i$ is a simple ring. If $M$ is a simple module it is isomorphic to some $I_k$ so there are only finitely many isomorphism classes of simple $R$-modules. Let $M$ be a non-zero $R$-module, define $M_i$ as the sum of all simple $R$ modules isomorphic to $M_i$, then $\bigoplus_{i=1}^t B_i M, M_i = e_i M$. A semi-simple ring, $R$ is ring isomorphic to the direct product of simple rings. Let $R$ be a simple ring and $V$ a simple $R$-module with $D = End_D(V)$, then $V$ is a finite dimensional vector space over $D$, $R \cong End_D(V) \cong M_n(D^o)$. Let $B = b_1 \oplus \ldots \oplus B_n$ be a direct sum of simple algebras then two sided ideals of $B$ are of the form $J_1 \oplus \ldots \oplus J_n$ where the $J_i$'s are 2 sided simple ideals of the $B_i$'s. Let $S_1, S_2, \ldots, S_r$ be distinct simple $A$-modules; for each $i$, let $U_i$ be a direct sum of copies of $S_i$ and $U = U_1 \oplus U_2 \oplus \ldots \oplus U_r$ then $End_A(U) = End_A(U_1) \oplus End_A(U_2) \oplus \ldots \oplus End_A(U_r)$. If $S$ is a simple $A$-module then $End_A(nS) \cong M_n(End_A(S))$ and if $F$ is algebraically closed then $End_A(S) \cong F$. **Wedderburn:** Let $R$ be a semi-simple ring then (1) $R$ is isomorphic to the direct sum of simple rings $B_1, B_2, \ldots, B_t$, (2) there are $t$ isomorphism classes of simple$R$-modules; if $V_1, V_2, \ldots, V_t$ are representatives, let $D_i = End_R(V_i)$ then $B_i \cong End_{D_i}(V_i) \cong M_n(D_i{}^o)$, and (3) $B_i V_j = 0, i \neq j, B_i V_i = V_i$. **Maschke:** If $G$ is a finite group and $k$ is a field with $char(k) \nmid |G|$ then $kG$ is semi-simple. Let $K = End_R(E)$ with $E$ semi-simple over $R$ and $f \in End_K(E)$; further, let $x \in E$ then $\exists \alpha \in R$: $f(x) = \alpha x$. Proof: $E = Rx \oplus F$ let $\pi \in End_K(E)$ be the projection on the first factor. $f(x) = \pi f(x) = f(\pi x) \in Rx$. Jacobson: Let $K = End_R(E)$ with $E$ semi-simple over $R$ and $f \in End_K(E)$ let $x_i \in E, i = 1, 2, \ldots, n$ then $\exists \alpha \in R$: $f(x_i) = \alpha x_i$. Let $R$ be a ring, $\psi \in End_R(R)$, $\exists \alpha \in R$: $\psi(x) = z\alpha$. $\psi(x) = \psi(x1) = x\psi(1)$. Rieffel: Let $R$ be a ring without non-trivial two sided ideals. Let $L$ be a non zero left ideal and $R' = End_R(L)$, $R'' = End_{R'}(L)$, then there is a natural map $\lambda : R \rightarrow R''$. Definition: $R$ is simple iff it has no non-trivial two sided ideals. If $R$ is semi-simple, $R = R_1 \oplus R_2 \oplus \ldots R_k$ with each $R_i$ simple. The decomposition is unique apart from order. Proof: Let $R_1$ be a minimal 2 sided ideal, $R = R_1 \oplus \overline{R}_1$, $R_1 = Re$, $\overline{R}_2 = R(1-e)$. Both are idempotent so sums and products act on each summand separately. Regularity is inherited by the summands. Now you can decompose $\overline{R}_2$ into a further sum. We get $R = R_1 \oplus \ldots \oplus R_s$ and $e = e_1 + e_2 + \ldots + e_s$, ${e_i}^2 = e_1$, $e_i e_j = 0$ and each $e_i$ is in $r_i$ and is in the center of $R_i$.

Two $FG$ modules afford equivalent representations iff they are isomorphic. Every irreducible ordinary representation of $G$ occurs as a component of the regular representation $R(G)$. The number of inequivalent irreducible representations is the number of conjugacy classes of $G$. If $\rho_1, \rho_2, \ldots, \rho_r$ are inequivalent representations and $deg(\rho_i) = n_i$ then $dim(\rho_i) = {n_i}^2$ and $\rho_i$ occurs $n_i$ times in $R(G)$. $|G| = \sum_{i=1}^r {n_i}^2$. Proof: Extend $F$ to a suitable algebraic extension so that the center of $R_G$ is the direct sum of $r$ matrix rings: $R_1, R_2, \ldots R_r$. if $dim(r_i) = {n_i}^2$, $deg(\rho_i) = n_i$. Since $dim(R_G) = |G|$, $dim(R_G) = \sum_{i=1}^r {n_i}^2$. Each $R_i$ is the direct sum of the $n_i$ right ideals: $e_{11}R, \ldots, e_{n_i n_i}R$. So $\rho_i$ occurs $n_i$ times in $R_G$. $\mathbb{Z}(G)$ has an irreducible faithful representation iff it is cyclic.

**Representations (2):** Let $V$ be an $\mathbb{C}G$ module, $V = U_1 \oplus U_2 \oplus \ldots \oplus U_r$ with $U_i$ irreducible. If $V, W$ are $\mathbb{C}G$-modules and $\theta : V \rightarrow W$ is a $\mathbb{C}G$ module homomorphism, $\exists U$, a submodule of $V$ such that $V = ker(\theta) \oplus U$. $Hom_{\mathbb{C}G}(V, W)$ is a vector space over $\mathbb{C}$. If $V, W$ are irreducible $\mathbb{C}G$ modules, $dim_{\mathbb{C}}(Hom_{\mathbb{C}G}(V, W))$ is 1 if $V \cong W$ and 0 otherwise. $dim_{\mathbb{C}}(Hom_{\mathbb{C}G}(V, W)) \neq 0$ if $V$ and $W$ have a common composition factor. Let $V$ be an $\mathbb{C}G$ module, $V = U_1 \oplus U_2 \oplus \ldots \oplus U_r$ with $U_i$ irreducible; (a) if $W$ is an irreducible $\mathbb{C}G$ module then $dim_{\mathbb{C}}(Hom_{\mathbb{C}G}(V, W)) = dim_{\mathbb{C}}(Hom_{\mathbb{C}G}(W, V))$ is the number of $U_i \cong W$; (b) each $U_i$ is a composition factor in the Jordan Holder series. $\mathbb{C}G = U_1 \oplus U_2 \oplus \ldots \oplus U_r$ with $U_i$ irreducible; if $G$ is finite, there are finitely many irreducible $\mathbb{C}G$ modules. $dim(Hom_{\mathbb{C}G}(V_1 \oplus \ldots \oplus U_r, W_1 \oplus \ldots \oplus W_s) = \sum_{i=1,j=1}^{r,s} dim(Hom_{\mathbb{C}G}(V_i, W_j))$. $dim(Hom_{\mathbb{C}G}(\mathbb{C}G, U)) = dim(U)$. [Proof: Let $d = dim(U)$ and $u_1, u_2, \ldots, u_d$ be a basis for $U$. Define $r\phi_i = u_i r$. The $\phi_i$ are a basis for $Hom_{\mathbb{C}G}(\mathbb{C}G, U)$]. If $V_1, V_2, \ldots, V_r$ are a complete set of irreducible $\mathbb{C}G$-modules then $|G| = \sum_{i=1}^r dim(V_i)^2$. [Proof: $V = U_1 \oplus U_2 \oplus \ldots \oplus U_k$, of these, $dim(V_i)$ are isomorphic to $V_i$ and each of these had dimension $dim(V_i)$].

If $G \subseteq S_n$, $\alpha : G \rightarrow \mathbb{C}$ by $\alpha(g) = |fix(g)| - 1$, then $\alpha$ is a character of $G$. Define $ker(\rho) = \{g : \chi_\rho(g) = \chi_\rho(1)$;

$\rho$ is faithful iff $ker(\rho) = 1$. $N = \{n : |\chi(n)| = \chi(1)\} \lhd G$. If $N \lhd G, \exists \chi_i : \bigcap_{i=1}^{r} ker(\chi_i) = N$. $g \sim h$ iff $\chi(g) = \chi(h), \forall \chi$. Let $x \in A_n$; if there is an odd permutation that commutes with $x$, $ccl_{A_n}(x) = ccl_{S_n}(x)$ otherwise $ccl_{S_n}(x)$ splits into two conjugacy classes in $A_n$. Let $C_i = \sum_{x \in ccl(y)} x$ then the $C_i$ form a basis for $\mathbb{Z}(FG)$.

Suppose $\chi$ is a character of a $\mathbb{C}G$-module, $V$, and $g \in G$ has order $m$ then (1) $\chi(1)$ $dim(V)$, (2) $\chi(g)$ is a sum of $m$-th roots of unity, (3) $\chi(g^{-1}) = \overline{\chi(g)}$ and (4) $\chi(g)$ is real iff $g \sim g^{-1}$. If $\chi$ is an irreducible character, $\chi(1) \mid |G|$ [If $g_i$ is in the $i$th conjugacy class, $\frac{|G|}{|C_G(g_i)|} \frac{\chi(g_i)}{\chi(1)}$ and $\overline{\chi(g)}$ are algebraic integers so $\sum_{i=1}^{k} \frac{|G|}{|C_G(g_i)|} \frac{\chi(g_i)}{\chi(1)} \overline{\chi(g)} = \frac{|G|}{\chi(1)}$ is.]

**Burnside's Theorem:** $|\frac{\chi(g)}{\chi(1)}| \leq 1$ if $|\frac{\chi(g)}{\chi(1)}| \neq 1$ it is not an algebraic integer. Let $p$ be a prime and $G$ a finite group with conjugacy class of size $p^r, r \geq 1$, then $G$ is not simple. Every group of order $p^a q^b$ is solvable. Let $\chi$ be an irreducible character and $C$ a conjugacy class. If$(\chi(1), |C|) = 1$ then either $C \subseteq Z(\chi)$ or $\chi(C) = 0$. If $G$ is a non-abelian simple group $\{1\}$ is the only class with prime power order.

$\chi_{reg} = \chi_1(1)\chi_1(g) + \chi_2(1)\chi_2(g) + \ldots + \chi_r(1)\chi_r(g)$. Let $U, V$ be non-isomorphic irreducible $\mathbb{C}G$ modules with characters $\chi, \psi$, then $< \chi, \chi >= 1$ and $< \chi, \psi >= 0$. $\chi(g)$ is real iff $\chi(g) = \chi(g^{-1}), \forall \chi$. $N \lhd G$ iff $\exists \chi_i, i = 1, \ldots, k$ such that $\bigcap_{i=1}^{k} ker(\chi_i) = N$. $G$ is not simple iff $\exists \chi, g \neq 1 : \chi(g) = \chi(1)$. $G$ has $|G/G'|$ linear characters. If all irreducible representations of $G$ have dimension 1, $G$ is abelian.

**Feit's moduleless treatment.** Maschke: If $char(F)$ does not divide $|G|$, then $F$-representations of $G$ are completely reducible. For $\phi$ irreducible, if $\exists S : \forall g, S\phi(g) = \phi(g)S$ then S is non-singular. If $A(g), B(g)$ are $k$-irreducible then (i) if $A$ is not similar to $B$, and, $\sum_g a_{is}(g) b_{tj}(g^{-1}) = 0$; or, (ii) $A$, $B$ are absolutely irreducible and $\sum_g a_{is}(g^{-1}) a_{tj}(g) = \frac{|G|}{n} \delta_{ij} \delta_{st}$, where $n \times n$ is the dimension of $(a_{is}(g))$. If $A^s$ is absolutely irreducible then $a_{ij}^s(g)$ are linearly independent and $\sum_{s=1}^{k} n_s^2 \leq |G|$.

Define $(\theta, \eta) = \frac{1}{|G|} \sum_g \theta(g) \overline{\eta(g)}$. If $U = U_1 \otimes ... \otimes U_s$, the number of these similar to $U_1$ is $\frac{(\theta,\eta)}{(\eta,\eta)}$. $(\theta, \rho_G) = \theta(1)$, $(\chi_i, \chi_j) = \delta_{ij}$, $\sum_g \chi(g) = |G| \delta_{i1}$, $\sum_i \chi_i^2(1) = |G|$. $\omega_i(R_j) = |r_j| \chi_i(g)/\chi_i(1)$, $\omega_t(R_i)\omega_t(R_j) = \sum_s a_{ijs} \omega_t(R_s)$. $\sum_t \chi_t(g_i) \overline{\chi}_t(g_j) = \frac{|G|}{|R_j|} \delta_{ij}$. The number of conjugacy classes = number of irreducible representations. $\omega_i(R_j)$ is an algebraic integer. $\chi_i || G|$, $(|R|, \chi(1)) = 1, \chi^{irred} \rightarrow |\chi(g)| = 1$ or $\chi(g) = 0$. Let $H$ be the kernel of $\theta$ then (i) $|\theta(g)| \leq \theta(1)$, (ii) $\theta(g) = \theta(1)$, iff $g \in H$, (iii) $|\theta(g)| = \theta(1)$, iff $gH$ is in the center if $G/H$.

**Induced representations:** If $H \leq G$ and $\varphi$ a class function on $H$, define $\varphi^G(g) = \frac{1}{|H|} \sum_{x \in G} \varphi^*(x^{-1} g x)$. **Frobenius Reciprocity:** $(\varphi^G, \theta) = (\varphi, \theta_{|H})$.

**Brauer's Characterization of Characters:** $p$-elementary groups are the products of a cyclic $p'$ group and $p$ group. Every irreducible character is an induced character of a linear character of a $p$ elementary subgroup for some $p$.

**RSK correspondence** for representations of the symmetric group: $\exists$ bijection between $S_n$ and the set of ordered tableau of the same shape $g \leftrightarrow (S, T)$, further $g^{-1} \leftrightarrow (T, S)$. **Young's diagram:** $D(\lambda)$, $n = n_1 + n_2 + \ldots + n_k$, $n_1 \geq n_2 \geq \ldots \geq n_k$. Number of tableaus with shape $\lambda$: $f_\lambda = \frac{n!}{\prod_{i,j \in D(\lambda)} h(i,j)}$, where $h(i,j)$ = number of cells in hook $H_{i,j}$.

Let $G$ be a transitive permutation group on $X$ and $1 \neq g \in G$ fixes no more than one element then $N = \{g : X_g = \emptyset\}$ is a normal subgroup of $G$. Thompson showed any finite group having a fixed point free automorphism is nilpotent.

**Characters and group structure:** The character table determines the normal subgroups and the nilpotent groups. General procedure for calculating characters: (1) Derive a faithful representation, (2) generate group elements, (3) determine conjugacy classes, (4) determine structure constants $(|C_i||C_j| = \sum_k \alpha_{ijk} |C_k|)$, (5) get characters from structure constants.

**Schrier and coset enumeration:** Let $G =< g_1, g_2, \ldots, g_m >$. Let $k_1, k_2, \ldots, k_s$ be a group of coset representatives for a subgroup $H < G$. $\overline{g}$ is the coset representative for $g$ in $G/H$ and $k_1 = 1$ then $H =< (k_i g_j)\overline{(k_i g_j)}^{-1} >$ for $i = 1, 2, 3, \ldots, s$ and $j = 1, 2, 3, \ldots, m$. Maintain following tables: Coset, relation table for each relation, subset table. Column headers are generators, rows are right coset labels. To calculate $|G|$, calculate orbit of point. Calculate point stabilizer by completing paths in Schrier tree and using the resulting relations.

$\mathcal{B} =< \beta_1, ..., \beta_n >$ is a base for $G \leq Sym(\Omega)$ if $G_{\mathcal{B}} = 1$. If $G^{[i]} = G_{\beta_1,...,\beta_i}$ and $G = G^{[1]} \geq ... \geq G^{[m+1]} = 1$ then $S$ is a **strong generating set** relative to $\mathcal{B}$ if it is a generating set and $S \cap G^{[i]} = G^{[i]}$. Can use this to get orbit sizes. Schrier-Sims calculates base and strong generating set.

**Coxeter groups:** $M = (m_{ij}), 1 \leq i, j \leq n, m_{ii} = 1$, $m_{ij} \in \mathbb{Z}, m_{ij} \geq 2$. Associate to each such matrix a graph with nodes $i, 1 \leq i \leq n$, $(i, j)$ is an edge if $m_{ij} > 0$ if $m_{ij} > 2$, label it with $m_{ij} - 2$. The Coxeter group is $G$ generated by $S = \{s_i\}, 1 \leq i \leq n$ with $(s_i s_j)^{m_{ij}} = 1$. Note the $s_i$'s must be involutions, $\theta = \frac{\pi}{m_{ij}}$. Geometrically: If $\Delta = \{r_1, \ldots, r_n\}, ||r_i|| = 1$ is a **root system** with each $r_i$ defining a reflection along its associated hyperplane by $S_r(x) = x - 2(r, x)r$ and $\alpha_{ij} = -cos(\frac{\pi}{p_{ij}}) = (r_i, r_j)$. Associate a marked graph with edges labeled by $p_{ij}$ (unmarked edged have $p_{ij} = 3$) and associated quadratic form $Q(\vec{x}) = \sum \alpha_{ij} x_i x_j$. The Coxeter group is generated by the involutions $S_r$ and $S_{r_i} S_{r_j}$ has order $p_{ij}$. The quadratic forms are positive definite and the associated forms are irreducible iff the graphs are connected. The root system is effective iff the roots generate the underlying vector space. Union of the fundamental region under each element of $G$ is the vector space.

**Classical Groups:** Every **transvection** in $SL_n(F)$ is conjugate if $n > 2$. Group orders: $PSL_n(q) = \frac{1}{(q-1)(n,q-1)}(q^m - 1)(q^m - q)(q^m - q^2)...(q^m - q^{m-1})$, simple if $n > 2$ or $q > 3$. $PSp_{2l}(q) = \frac{1}{(2,q-1)} q^{l^2}(q^2 - 1)(q^4 - 1)...(q^{2l} - 1)$, simple unless $(2l, q) = (2, 2), (2, 3), (4, 2)$. $PSU_n(q^2) = \frac{1}{(n,q+1)}(q^{\frac{n(n-1)}{2}} - 1)(q^2 - 1)(q^3 + 1)(q^4 - 1)...(q^n - (-1)^n)$, simple unless $(2l, q) = (2, 4), (2, 9), (3, 4)$. For next two, set $\Omega_n(q) = (O_n(q))' \subseteq SO_n(q)$. $P\Omega_{2l+1}(q) = \frac{1}{(n,q-1)} q^{l^2}(q^2 - 1)(q^4 - 1)...(q^{2l} - 1)$, simple if $l > 1$. Note $P\Omega_{2l+1}(q)$ is not isomorphic to $PSp_{2l}(q)$ despite having the same order. For $|\epsilon| = 1$, $P\Omega^{\epsilon}{}_{2l}(q) = \frac{1}{(4,q^l-\epsilon)} q^{l(l-1)}(q^2 - 1)(q^4 - 1)...(q^{2l-2} - 1)(q^l - \epsilon)$, if $q = 2^k$, simple if $l > 2$.

**Finite Simple Group Families:** $\mathbb{Z}_p$, Schur Multiplier: 1. $\Sigma'_n$ simple if $n > 4$, Schur Multiplier: 6 if $n = 6, 7$, 2 if $n = 5, n > 7$. $A_n(q) = PSL_{n+1}(q)$ simple if $n \geq 1$, Schur Multiplier: $(n + 1, q - 1)$ except $A_1(4)[2]$, $A_1(9)[6]$, $A_2(4)[48]$, $A_3(2)[2]$. $B_n(q) = P\Omega_{2n+1}(q)$ simple if $n \geq 1$ , Schur Multiplier: $(2, q - 1)$ except $B_2(2)$, $B_3(2)[2]$, $B_2(2)[6]$; $C_n(q) = PSp_{2n}(q)$ simple if $n > 2$ , Schur Multiplier: $(2, q - 1)$ except $C_3(2)[2]$. $D_n(q) = P\Omega^+_{2n}(q)$ simple if $n \geq 4$, Schur Multiplier: $(2, q - 1)$ except $D_4(2)[4]$. $E_6(q)$ of order $\frac{1}{(3,q-1)} q^{36}(q^{12} - 1)(q^9 - 1)(q^8 - 1)(q^6 - 1)(q^5 - 1)(q^2 - 1)$, Schur Multiplier: $(3, q - 1)$. $E_7(q)$ of order $\frac{1}{(3,q-1)} q^{63}(q^{18} - 1)(q^{14} - 1)(q^{12} - 1)(q^{10} - 1)(q^8 - 1)(q^6 - 1)(q^2 - 1)$, Schur Multiplier: $(2, q - 1)$. $E_8(q)$ of order $q^{120}(q^{30} - 1)(q^{24} - 1)(q^{20} - 1)(q^{18} - 1)(q^{14} - 1)(q^{12} - 1)(q^8 - 1)(q^2 - 1)$, Schur Multiplier: 1. $F_4(q)$ of order $q^{24}(q^{12} - 1)(q^8 - 1)(q^6 - 1)(q^2 - 1)$, Schur Multiplier: 1 except $F_4(2)[4]$. $G_2(q)$ simple except $G_2(2)$ of order $q^6(q^6 - 1)(q^2 - 1)$, Schur Multiplier: 1 except $G_2(3)[3]$, $G_2(4)[2]$. ${}^2A_n(q^2) = PSU_{n+1}(q)$ simple if $n \geq 2$, Schur Multiplier: $(n+1, q+1)$ except ${}^2A_3(2^2)[2]$, ${}^2A_3(3^2)[36]$, ${}^2A_5(2^2)[12]$ . ${}^2D_n(q) = P\Omega^-_{2n}(q)$ simple if $n \geq 4$, Schur Multiplier: $(4, q^n + 1)$. ${}^3D_4(q^3)$ of order $q^{12}(q^8 + q^4 + 1)(q^6 - 1)(q^2 - 1)$, Schur Multiplier: 1 . ${}^2E_6(q)$ of order $q^{36}(q^{12} - 1)(q^9 + 1)(q^8 - 1)(q^6 - 1)(q^2 - 1)$, Schur Multiplier: $(3, q + 1)$ except ${}^2E_6(2^2)[12]$. ${}^2B_2(2^{2m+1}) = Sz(2^{2m+1})$ simple if $m > 1$ of order $q^2(q^2 + 1)(q - 1)$, Schur Multiplier: $1, n > 2$. ${}^2F_4(2^{2m+1})$ (Ree) simple if $m > 1$ of order $q^{12}(q^6 + 1)(q^4 - 1)(q^3 + 1)(q - 1)$, Schur Multiplier: $1, m > 1$. ${}^2G_2(3^{2m+1})$ (Ree) simple if $m > 1$ of order $q^3(q^3 + 1)(q - 1)$, Schur Multiplier: $1, m > 1$.

**Sporadic Groups:** $M_{11}$ $(2^4 \cdot 3^2 \cdot 5 \cdot 11)$, Schur: 1. $M_{12}$ $(2^6 \cdot 3^3 \cdot 7 \cdot 11)$, Schur: 2. $M_{22}$ $(2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11)$, Schur: 12. $M_{23}$ $(2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 23)$, Schur: 1. $M_{24}$ $(2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23)$, Schur: 1. $J_1$ $(2^3 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19)$, Schur: 1. $J_2 = HJ$ $(2^7 \cdot 3^3 \cdot 5^2 \cdot 7)$, Schur: 2. $J_3 = HJM$ $(2^7 \cdot 3^5 \cdot 5 \cdot 17 \cdot 19)$, Schur: 3. $J_4$ $(2^{21} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11^3 \cdot 23 \cdot 29 \cdot 31 \cdot 37 \cdot 43)$, Schur: 1. $Co_1$ $(2^{21} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23)$, Schur: 2. $Co_2$ $(2^{18} \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23)$, Schur: 1. $Co_3$ $(2^{10} \cdot 3^7 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23)$, Schur: 1. $HS$ $(2^9 \cdot 3^2 \cdot 5^3 \cdot 7 \cdot 11)$, Schur: 2. $Mc$ $(2^7 \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11)$, Schur: 3. $Sz$ $(2^{13} \cdot 3^7 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13)$, Schur: 1. $Ly$ $(2^8 \cdot 3^7 \cdot 5^6 \cdot 7 \cdot 11 \cdot 31 \cdot 37 \cdot 57)$, Schur: 1. $He$ $(2^{10} \cdot 3^3 \cdot 5^2 \cdot 7^3 \cdot 17)$, Schur: 1. $Ru$ $(2^{14} \cdot 3^3 \cdot 5^3 \cdot 7 \cdot 13 \cdot 29)$, Schur: 1. $O'N{-}S$ $(2^9 \cdot 3^4 \cdot 5 \cdot 7^3 \cdot 11 \cdot 19 \cdot 31)$, Schur: 3. $F_{22}$ $(2^{17} \cdot 3^9 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13)$, Schur: 6. $F_{23}$ $(2^{18} \cdot 3^{13} \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 23)$, Schur: 1. $F_{24}$ $(2^{21} \cdot 3^{16} \cdot 5^2 \cdot 7^3 \cdot 11 \cdot 13 \cdot 17 \cdot 23 \cdot 29)$, Schur: 3. $F_3$ (Thompson) $(2^{15} \cdot 3^{10} \cdot 5^3 \cdot 7^2 \cdot 13 \cdot 19 \cdot 31)$, Schur: 2. $F_5$ (Harada) $(2^{14} \cdot 3^6 \cdot 5^6 \cdot 7 \cdot 11 \cdot 19)$, Schur: 1. $F_2$ (Baby Monster) $(2^{41} \cdot 3^{13} \cdot 5^6 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17 \cdot 19 \cdot 23 \cdot 31 \cdot 47)$, Schur: 2. $F_1$ (Monster) ($2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$), Schur: 1.

Let $\Delta$ be an orbit of $G$ and let $\delta \in \Delta$. For each $\gamma \in \Delta$ let $v(\gamma) \in G$ be such that $\delta \mapsto \gamma$. Finally, suppose $S$ generates $G$. Then $G_\delta =< v(\gamma)sv(\gamma^s)^{-1} | \gamma \in \Delta, s \in S >$.

**System of imprimitivity for permutation group $G$:** $\mathcal{B} = \{\Delta_i\}$ $|\Delta_i| > 1$ with the property that for $\Delta \in \mathcal{B}, g \in G$ either $\Delta \cap \Delta^g = \phi$ or $\Delta = \Delta^g$. **Primitive:** No set of imprimitivity. $\Gamma$ is $G$ invariant if $\Gamma^G = \Gamma$ so $\Gamma$ is a union of $G$ orbits. $G/G_\Gamma \equiv G^\Gamma$. If $\Delta \subseteq \Gamma$ and $\alpha \in \Omega$ then $\psi = \bigcap_{\alpha \in \Delta^g} \Delta^g$ is a block of a transitive group $G \subseteq Sym(\Omega)$. A transitive group is imprimitive iff $\exists Z$: $G_\alpha < Z < G$. $G$ is primitive iff $G_\alpha$ is maximal. Let $G$ act transitively on $\Omega$, $H \lhd G$ then (1) The orbits of $H$ are blocks of $G$, (2) If $\Delta$ and $\Delta'$ are two $H$ orbits then they are permutation isomorphic, (3) If any point lies is fixed by all elements of $H$ then $H$ lies in the kernel of the action on $\Omega$, (4) The group $H$ has at most $|G : H|$ orbits, if finite, it divides $|G : H|$, (5) If $G$ acts primitively on $\Omega$ then either $H$ is transitive or it lies in the kernel of the action.

Define $\mathcal{G}(G, \Omega)$ as the graph of $G$ acting on $\Omega$ as follows: $G$ acts on $\Omega \times \Omega$. Diagonal orbital is $\Delta_1 = \{(\alpha, \alpha)\}$. If $\Delta = \{(\alpha, \beta)\}$, $\Delta^* = \{(\beta, \alpha)\}$. Self paired if $\Delta^* = \Delta$. $\Delta(\alpha) = \{\beta : (\alpha, \beta) \in \Delta\}$ — corresponds to orbits of $G_\alpha$. The **rank of the permutation group** is number of orbitals. On a self-paired orbit $\Delta$, the graph $\mathcal{G} = (G, X, \Delta)$ is symmetric and $G$ is transitive on edges. Let $G$ be a transitive permutation group of even order and rank 3 with two necessarily self-paired non-diagonal orbits $\Delta$ and $\Gamma$. $G$ is primitive iff $\mathcal{G}$ is connected.

A transitive permutation group is **regular** if $|X| = |G^X|$ or, equivalently $|G_x| = 1, \forall x \in X$ and $G^X$, transitive. Let $X$ be a faithful primitive $G - set$ with $G_x$ simple. The either $G$ is simple or every non-trivial normal subgroup $H$ of $G$ is a regular normal subgroup. **Iwasawa:** Let $G = G'$ and $X$ be a faithful primitive $G - set$. If there is an $x \in X$ and an Abelian normal subgroup $K \lhd G_x$ whose conjugates generate $G$ then G is simple. Permutation representation: Let $H \leq G$ and $Hg_1, ..., Hg_n$ be the cosets; the map $\pi(g) :< Hg_1, ..., Hg_n > \mapsto < Hg_1g, ..., Hg_ng >$ is a map from $G$ to $\Sigma_n$ whose kernel is the largest normal subgroup of $G$ in $H$. Corollary: If $H < G$ and $G$ is simple then $|G| \mid |G : H|!$. If $G^X$ is primitive and $1 \neq N \lhd G^X$ then $N^X$ is transitive. If $G^X$ is primitive and $G_x$ is simple then either (1) $G$ is simple, or (2) $\exists N \lhd G : N^X$ is regular. If $N$ is a regular normal subgroup of $G^X$ then $G_x$ acts on $N^\#$. If $A$ is transitive on $H^\#$ then $H \cong (Z_p)^n$, if 2-transitive, $H \cong (Z_2)^n$ or $Z_3$, if 3-transitive, $H \cong (Z_2)^2$.

**Frobenius group:** Transitive permutation group with non-trivial stabilizers but only the identity fixes more than one letter. If $G$ is a Frobenius group then the set $S$ of elements which fix no points together with $e$ form a normal subgroup of order $|G : G_a|$; Thompson showed this normal subgroup is nilpotent.

**Metacyclic:** $\exists H \lhd G : G/H, H$ are cyclic. $Core_G(H) = \bigcap_{g \in G} H^g$ (Can use this to show $|G : Core_G(H)| \leq |G : H|!$). $O^{\mathcal{A}}(G) = \bigcap_{A \lhd G, G/A \in \mathcal{A}} A$. $O_{\mathcal{A}}(G) = \prod_{A \lhd G, A \in \mathcal{A}} A$. **Socle:** $soc(G) =< M >$ where $M$ is a non-trivial minimal normal subgroup of $G$. $O_\pi(G) =$ maximal normal $\pi-$subgroup of $G$. $O^\pi(G) =$ smallest normal subgroup of $G$ such that $G/O^\pi(G)$ is a $\pi$-group. $G$ is $p-$closed if $O_p(G) \in S_p(G)$. $SCN(P) =$ set of self centralizing normal subgroups of $P$. $SCN(p) = SCN(P)$ where $P \in SCN(P)$. $\mathcal{N}_G(A, \pi) =$ set of all $A-$ invariant $\pi$ subgroups of $G$. $\mathcal{N}^*_G(A, \pi) =$ maximal subgroups in $\mathcal{N}_G(A, \pi)$. For a $p-$group, $P$, $\Omega_n(P) =< x \in P : x^{p^n} = 1 >$ and $\mho_n(P) =< x^{p^n} : x \in P >$. $H \subseteq G$ and $S$ an $H$-invariant subset of $G$, $H$ is said to control fusion in $S$ if for $s \in S$, $s^G \cap S = s^H$. Let $X \leq H \leq G$. $X$ is **weakly closed** in $H$ with respect to $G$ if $X^g \cap H = \{X\}$. $G$ is $p-$**solvable** if it has a normal series whose factors are either $p-$groups or $p'$-groups. $G$ is $p$-**constrained** if $P \in S_p(O_{p',p}(G))$ implies $C(P) \subseteq O_{p',p}(G)$. $G$ is $p$-**stable** if $p \neq 2$ and if $A \in p(N(P))$ with $[P, A, A] = 1$ implies $AC(P)/C(P) \subseteq O_p(N(P)/C(P))$. $m_p(P)$ is the rank of the largest elementary abelian $p$-group in $P$. $O_\infty(G) =$ largest solvable normal subgroup of $G$. $F(G)$ is the unique maximal normal, nilpotent subgroup of $G$ and $F(G) = \prod_p O_p(G)$ . $E_{p^n}$ denotes the elementary abelian $p-$group of rank $n$. $m_{2,p}(G) = max\{m_p(H)\}$, where $H$ is 2-local. $e(G) = max\{m_{2,p}(G), p \neq 2\}$ ($e(G)$ is a good approximation of the Lie rank.).

**Modular Property:** If $A, B, C \leq G$ and $A \leq C$ then $AB \cap C = A(B \cap C)$. $[ab, c] = [a, c]^b[b, c]$ and $[a, bc] = [a, c][a, b]^c$. **Jacobi:** $[x, y^{-1}, z][y, z^{-1}, x][z, x^{-1}, y] = 1$. If $x, y \in C(z), z = [x, y]$ then $[x^n, y^m] = z^{mn}$ and $(yx)^n = y^n x^n z^{\frac{n(n-1)}{2}}$. **Three Subgroups:** $A, B, C \subseteq G$ and $N \lhd G$ with $[A, B, C] \subseteq N$ and $[B, C, A] \subseteq N$ then $[C, A, B] \subseteq N$.

Let $G$ be a group with $G/Z(G)$ finite, then $G^{(1)}$ is finite. Proof: Let $n = |G/Z(G)|$. For $z \in Z(G)$ and $g, h \in G$: $[g, hz] = [g, h] = [gz, h]$ so the set of commutators, $\Delta$, is of order at most $n^2$. Claim: $g \in G^{(1)}$ then $g = x_1 x_2 \dots x_m$, $x_i \in \Delta$ and $m \leq n^3$.

**Critical subgroup of a $p$-group:** $H$ *char* $G$ with $\Phi(H) \leq Z(H) \geq [G, H]$. $C_G(H) = Z(H)$. Every $p-$group has a critical subgroup. A $p-$group $P$ is **special** if $\Phi(G) = Z(G) = G'$ and **extra-special** if $Z(G)$ is cyclic. Let $G$ be a non-abelian group of order $p^n$ with cyclic subgroup $H$ of index $p$ then $G \cong < p^n >, D_{2^n}, SD_{2^n}, Q_{2^n}$.

**O-Nan-Scott:** Let $G$ be a finite primitive permutation group of degree $n$ and $H = soc(G)$. Then either (1) $H$ is a regular elementary abelian $p$ group for some $p$ and $G$ is isomorphic to a subgroup of $AGL_m(P)$ ; or, (2) $H$ is isomorphic to $T^m$ where $T$ is a non-abelian simple group with a bunch of conditions.

**Mathieu Groups:** $M_{11}$: $\pi_1 = (123)(456)(789), \pi_2 = (147)(258)(369)$, $< \pi_1, \pi_2 >= \mathbb{Z}_3 \times \mathbb{Z}_3$, $\rho_1 = (2437)(5698), \rho_2 = (2539)(4876)$, $< \rho_1, \rho_2 >= Q \cong Q_8$. Set $M_9 =< \pi_1, \pi_2, \rho_1, \rho_2 >$, $|M_9| = 72$. Now set $\sigma = (1, 10)(4, 5)(6, 8)(7, 9)$, $\mu = (4, 7)(5, 8)(6, 9)(10, 11)$, $\theta = (4, 9)(5, 7)(6, 8)(11, 12)$. $M_{10} = M_9 \cup M_9 \sigma M_9$, $(M_{10})_x = M_9$, $M_{11} = M_{10} \cup M_{10} \mu M_{10}$, $(M_{11})_x = M_{10}$, $M_{12} = M_{11} \cup M_{11} \theta M_{11}$, $(M_{12})_x = M_{11}$. $|M_{11}| = 7920$. $|M_{24}| = 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 48$. $M_{11}$ is simple: Let $N$ be a non-trivial normal subgroup, it is regular and all Sylow 11 subgroups are contained in it (there are 144 by sylow) and $G : N = 5$. All Sylow 3 subgroups of $M_{11}$ are in N and $\psi = \pi_1 \sigma \pi_2^2 \sigma^{-1}$ has order 5 which is a contradiction. Note symmetries of $S(4, 5, 11)$ also generate it. Note that $(M_{11})_a = PSL_2(9)$ and $(M_{22})_a = PSL_3(4)$.

**Schur-Zassenhaus:** Let $G$ be a finite group, $H \lhd G$ and $(|H|, |G : H|) = 1$ and either are solvable then $G$ splits over $H$ and $G$ is transitive on $H$ complements.
Proof of existence by induction: Suppose it holds for all groups of order $< G$ and that $|G| = nm; (m, n) = 1; N \lhd G; |N| = n$. If $\exists K \leq G : |K| = m$ then the theorem is true. Let $P \in S_p(N)$. (1) We may assume $P \lhd N$: If not $G = N_G(P)N, N_N(P) = N_G(P) \cap N \lhd N_G(P)$ and $m = |G/N| = |N(P)N/N| = |N_G(P)/(N_G(P) \cap N| = |N_G(P)/N_N(P)|$ and $N_G(P)$ has a normal Hall group $N_N(P)$ so by induction $\exists K \subseteq N_G(P)$ with $|K| = m$ and $N_N(P)K = N_G(P)$, so $NK = G$. (2) We may assume $P = N$: If not, $|(G/P)/(N/P)| = m$ so $\exists L/P : (N/P)(L/P) = G/P$ and $|L| = m|P|$, $|L \cap N| \mid (|L|, |N|)$; but $(m, |N|) = 1$ so $L \cap N \subset P$ and $L < G$ and $\exists K \subset L : |K| = m$. (3) May assume $N = P$ is abelian: If not $1 \neq Z = Z(N)$ *char*; $N \lhd G$ and $|(G/Z)/(N/Z)| = m$ so $\exists L/Z : (L/Z)(N/Z) = (G/Z)$ and $L \cap N = Z, L < G$ and $(|Z|, |L/Z|) = 1$ and $L$ and hence $G$ has a desired subgroup $K$. (4) So it suffices to show the theorem if $N$ is a normal abelian Hall $p-$group. Let $\overline{H} = G/N$. If $h \in \overline{H}$ and $t, u$ are two elements of $h$ then $t^{-1}u \in N$ so $tnt^{-1} = unu^{-1}$. Define ${}^h x = txt^{-1}, t \in h$. $H$ acts on $N$ - i.e. $H \subset Aut(N)$. Select a transversal $\{t_h | h \in H\}$. $t_{h_1 h_2}^{-1} N = (t_{h_1 h_2} N)^{-1} = (h_1 h_2)^{-1} = {h_1}^{-1} {h_2}^{-1}, \forall h_1, h_2 \in H$, so $t_{h_1} t_{h_2} t_{h_1 h_2}^{-1} \in N$. Define $f : H \times H \rightarrow N$ by $f(h_1, h_2) t_{h_1 h_2} = t_{h_1} t_{h_2}$. Since $t_{h_1}(t_{h_2} t_{h_3}) = (t_{h_1} t_{h_2}) t_{h_3}$, we get ${}^{h_1} f(h_2, h_3) + f(h_1, h_2 h_3) = f(h_1, h_2) + f(h_1 h_2, h_3)$. If $\exists c : H \rightarrow N : f(h_1, h_2) = c(h_1 h_2) - c(h_1) - {}^{h_1} c(h_2)$, then $c(h_1 h_2) t_{h_1 h_2} = c(t_1) t_{h_1} c(t_2) t_{h_2}$, this would be an isomorphism whose image would satisfy the requirements of $K$. Define: $e : H \rightarrow N$ by $e(h) = \sum_{k \in H} f(h, k)$. $mf(h_1, h_2) = -e(h_1 h_2) + e(h_1) + {}^{h_1} e(h_2)$. Since $(m, |N|) = 1$, $\frac{x}{m}$ is well defined for $x \in N$ and $c(x) = \frac{-1}{m} e(x)$ satisfies the desired properties.
Proof of conjugacy: Suppose $G/N$ is solvable and $\pi$ is the set of primes dividing $m = |G : N|$ and $H, K \leq G$ and $|H| = |K| = m$, put $R = O_\pi(G)$ so $O_\pi(G/R = 1$. Let $L/N$ be a minimal normal subgroup of $G/N$ then $L/N$ is an elementary abelian $p-$group for some $p$. $H \cap L \in S_p(L)$ and $S = (H \cap L) = (K \cap L)^g = K^g \cap L$. $S \lhd < H, K^g >= J$. If $J = G$, $S \lhd J$ and $S \subseteq R = 1$; thus $L$ is a $p'-$group which is a contradiction. So $J \neq G$ and by induction $K$, $K^g$ are $J$-conjugate. This concludes this case. Suppose $N$ is solvable and again $|H| = |K| = m = |G : N|$. $HN'/N' \cong KN'/N'$ so $h^g \subseteq KN'$ and again by induction, $H^{gk} = K$.

**Philip Hall's Theorem:** Let $G$ be a solvable group and $\pi$ a set of primes then (i) $G$ has a $\pi$-Hall subgroup, (ii) $G$ acts transitively on its Hall $\pi$-subgroups via conjugation, (3) any $\pi$ subgroup is contained in a Hall $\pi$ subgroup. Proof: By induction on $|G|$. Let $N$ be a minimal normal subgroup of $G$ then $1 \neq N \lhd G$. $N$ is elementary abelian for some $p$ and $p \mid mn$. If $p \mid m, |G/N| = \frac{m}{p}$ and $\exists L : |L/N| = \frac{m}{p}, |L| = m$ and we're done. If $p \mid n, \exists H : |H/N| = m, |H| = |N|m$. If $|H| < |G|$, we're done by induction. Otherwise $H = G, N \lhd G, |N| = n, |G : N| = m$ and $(m, n) = 1$ so by Schur Zassenhaus, $\exists K : |K| = m$.

**Theorem:** Let $G$ be a finite group possessing a Hall $\pi'$ subgroup for each $p$, then $G$ is solvable. (Proof requires Burnside $p^aq^b$ theorem.)

$|A|, |H| < \infty, (|A|, |H|) = 1$. Suppose $A \rightarrow Aut(H)$ and either are solvable then (1) $\exists A-$invariant Sylow $p-$group of $H$, (2) $C_H(A)$ is transitive on the $A-$invariant sylow $p-$subgroups of $G$, (3) If $K$ is an $A-$invariant normal subgroup of $H$ and $H^* = H/K$ then $C_{H^*}(A) = N_{H^*}(A) = (C_H(A))^*$. (5) Every $A-$invariant $p-$subgroup of $H$ is contained in an $A-$invariant Sylow $p-$group of $H$.

**Frattini subgroup:** $\Phi(G)$ is the intersection of all maximal subgroups of $G$. $\Phi(G)$ *char* $G$. If $H =< X, \Phi(H) >$ then $H =< X >$. If $P$ is a $p-$group $P/\Phi(P)$ is elementary abelian. Frattini Argument: $H \lhd G$, $P \in S_p(H)$ then $G = HN_G(P)$.

If $A$ is a maximal abelian normal subgroup of $P$ and $Z = \Omega_1(A)$. Then (1) $(C_P(A/Z) \cap C(Z))^{(1)} \leq A$, (3) if $p$ is odd $\Omega_1(C_P(Z)) \leq C_P(A/Z)$. If $p$ is odd and $Z$ is a maximal elementary abelian subgroup of $P$ then $Z \setminus \Omega_1(C_P(Z))$.

**Co-prime action 1:** In this paragraph $A$ acts on $G$ and $(|A|, |G|) = 1$ with either $A$ or $G$ solvable. If $U \leq G$ is $A-$invariant and $g$ satisfies $(Ug)^A = Ug$ then $\exists c \in C_G(A)$: $Ug = Uc$. If $N$ is an $A-$invariant normal subgroup of $G$ then (1) $C_{G/N}(A) = C_G(A)N/N$ (This shows $G = [G, A]C_G(A)$.) and (2) if $A$ acts trivially on $N$ and $G/N$ then $G$ acts trivially on $G$. If $p \mid |G|$ (the analogous results hold for $\pi$) then (1) $\exists S \in S_p(G) : S^A = S$, (2) all such $A-$invariant Sylow $p-$groups are conjugate under $C_G(A)$, (3) every $A-$invariant $p$-group of $G$ is contained in an $A-$invariant Sylow $p-$group. If $T = \bigcap_{S \in S_p(G), S^A = S} S$, the $T$ is the largest $A-$invariant $p-$subgroup of $G$ normalized by $C_G(A)$. If $P$ is an $A-$invariant Sylow $p-$group and $H \leq G$ with $H^A = H, H^{C_G}(G) = H$ then $P \cap H \in S_p(H)$. If $A = P \times Q$ acts on $M$ and $P, M$ are $p-$groups and $Q$ is a $p'-$group with $C_M(P) \leq C_M(Q)$ then $[M, Q] = 1$. If $A$ acts trivially on $G/\Phi(G)$ then $\underline{A}$ acts trivially on $G$ and if $\Phi(G)$ is a $\underline{p}-$group then so is $A/C_A(G)$. Applying $P \times Q$: If $p \in \pi(G)$ and $\overline{G} = G/O_{p'}(G)$ with $C_{\overline{G}}(O_p(\overline{G})) \leq O_p(\overline{G})$ then $\forall P \in p(G), O_{p'}(N_G(P)) = O_{p'}(G) \cap N_G(P)$.

**Co-prime action 2:** If $P$ is a $p-$group and $Q$ a $p'-$ group with $Q \mapsto Aut(P)$ then $Q$ is faithful on $P/\Phi(P)$. A group of automorphisms $A$ of a group $P$ stabilizes a chain $1 = P_n \subseteq P_{n-1} \subseteq \ldots \subseteq P_0 = P$ if $[A, P_i] \subseteq P_{i+1}$. If $P$ is a $\pi$ group stabilized by $A$ then $A$ is a $\pi$ group. Proof: $a \in A$ is a $\pi'$ automorphism. $x^a = xy, y \in P_1$. Similarly, $x^{a^{|a|}} = xy^{|a|} = x$, so $y = 1$ and $[a, P] = 1$. If $A$ is a $\pi'$ group of automorphisms on a $\pi$ group $P$ with $[P, A, A] = 1$ then $[P, A] = 1$. Proof: $A$ stabilizes $[P, A, A] \subseteq [P, A] \subseteq P$. Let $A$ be a $\pi'$ group of automorphisms of a $\pi$ group $P$. Let $Q$ be an $A-$invariant normal subgroup of $P$. Then $C_{P/Q}(A) = (C_P(A)Q)/Q$. Proof uses Schur-Zassenhaus. $P$ is a $\pi$ group, $A$ is a $\pi'$ group. $P = [P, A]C_P(A)$. Proof: $[P, A] \subseteq P$ and $A$ centralizes $P/[P, A]$. $P$ is an abelian $\pi$ group, $A$ is a $\pi'$ group. $P = [P, A] \oplus C_P(A)$. Proof: $\theta = \frac{1}{|A|} \sum_a a$.

If $G$ is solvable, (1) $C(F(G)) \subseteq F(G)$, (2) if $P$ is a $p-$group of $G$ then $O_{p'}(C(P)) \subseteq O_{p'}(G)$ and $O_{p'}(N_G(P)) \subseteq O_{p'}(G)$. If $P \in p(G)$ with $N_G(P)$ $p-$constrained then $C_G(P)$ is also $p-$constrained.

**Transfer:** $|G| < \infty, H \leq G$ . $|G : H| = n$ and $\{l_1, l_2, \ldots, l_n\}$ be a left traversal and suppose $gl_i = l_j x_i$ then $V(g) = \prod_{i=1}^n x_i H'$. $\exists h_1, h_2, \ldots, h_m \in H$ and $n_1, n_2, \ldots, n_m$. (1) $h_i \in \{l_1, l_2, \ldots, l_n\}$, (2) $h_i^{-1} g^{n_i} h_i \in H$, (3) $\sum_{i=1}^m n_i = |G : H|$, (4) $V(g) = \prod (h_i^{-1} g^{n_i} h_i H'$. If $Q$ is an abelian subgroup of finite order $n$ in $G$ and if $Q \subseteq Z(G)$ then $V(g) = g^n, \forall g \in G$. Let $Q \in S_p(G)$; if $g, h \in C(Q)$ and $g$ and $H$ are $G$ conjugate then they are $N(Q)$ conjugate. Let $Hx_ig^j, 1 \leq i \leq r, 0 \leq j \leq n_i$, cycles of $g$ on $G/H$. $X = \{x_ig^j\}$ then (a) $(g^{n_i})^{x_i^{-1}} \in H$ for $1 \leq i \leq r$, (b) $\sum_{i=1}^r n_i = |G : H|$ and (c) $V(g) = \prod_{i=1}^r ((g^{n_i})^{x_i^{-1}})^\alpha$.

Let $G$ be a finite group $H \leq G$, $(p, |G : H|) = 1, K \lhd H$, $H/K$ abelian, $g$ a $p-$element in $H \setminus K$: $g^{ma} \in g^m K, \forall m$, all $a \in G$ such that $g^{ma} \in H$ then $g \notin G^{(1)}$.

**Fusion:** Let $p$ be a prime, $T \in S_p(G), W \leq T$ with $W$ weakly closed in $T$ with respect to $G$ and $D = C_G(W)$. Then $N_G(W)$ controls fusion in $D$. $P \in S_p(G)$. $X \in p(G)$ is a tame intersection of $Q, R \in S_p(G)$ if $X = Q \cap R$ and $N_Q(X), N_R(X) \in S_p(N(X))$. **Alperin's Fusion Theorem:** If $P \in S_p(G), g \in G$ and $< A, A^g > \subseteq P$. Then for $1 \leq i \leq n$, $\exists Q_i \in S_p(G)$ and $x_i \in N(P \cap Q_i)$ such that (1) $g = x_1x_2...x_n$, (2) $P \cap Q_i$ is a tame intersection of $P$ and $Q_i$ for each $i$, (3) $A \subseteq P \cap Q_1$ and $A^{x_1x_2...x_i} \subseteq P \cap Q_{i+1}$. Supporting lemmas: $R, Q \in S_p(G)$. Say $R \rightarrow Q$ if $\exists Q_i \in S_p(G), X_i \in N_G(P \cap Q_i)$ such that (1) $P \cap Q_i$ is tame, (2)

$P \cap R \le P \cap Q_1$ and $P \cap R)^{x_1 x_2 \ldots x_i} \le P \cap Q_i$ and (3) $R^x = Q, x = x_1 x_2 \ldots x_n$. Sometimes say $R \rightarrow_x Q$. (1) $Q \rightarrow P, \forall Q \in S_p(G)$. (2) $P \rightarrow P$. (3) $\rightarrow$ is transitive. (4) $S \rightarrow_x P$, $Q^x \rightarrow P$ and $P \cap Q = P \cap S$ then $Q \rightarrow P$. (5) Assume $P \cap Q$ is tame and $S \rightarrow P, \forall S \in S_p(G)$ with $|S \cap P| > |Q \cap P|$ and $S \rightarrow P$ then $Q \rightarrow P$.

**Gaschutz:** Let $K$ be a normal abelian p-subgroup of a finite group $G$ and let $P \in S_p(G)$. Then $K$ has a complement in $G$ iff $K$ has a complement in $P$. If $K$ is an abelian normal subgroup of $G$ with $(|K|, |g : K|) = 1$ then $K$ has a complement. Proof: Set $\sigma(x) = \sum_{y \in Q} f(x, y)$.

**Focal Subgroup Theorem:** $S \in S_p(G)$ then $S \cap G' =< x^{-1}y | x, y \in S, x \sim_G y >$. Suppose $P \in S_p(G)$ and $A_1, A_2 \lhd G$, if $A_1^g = A_2$, then $\exists y \in N_G(P) : A_1^g = A_2$. **Burnside Normal $p$-complement**: (proved using transfer): If $P \in S_p(G)$ and $P \subseteq Z(N(P))$ then $P$ has a normal $p$-complement. If $P \in S_p(G), P' = 1$ then $P \cap G' = P \cap N_G(P)'$. **Frobenius Normal $p-$complement:** The following are equivalent: (1) $G$ has a normal $p-$complement, (2) Each $p-$local subgroup of $G$ has a normal $p-$complement, (3) $Aut_G(P)$ is a $p-$group $\forall P \in p(G)$. If $H \le G$ and $H \cap H^g = 1, \forall g \in (G \setminus H)$ then $G = NH, N \lhd G$.

**Thompson:** Let $a$ be a $\pi'$ automorphism of a $\pi$ group $P$ and suppose $X \lhd \lhd P$ such that $[a, X] = 1 = [a, C_P(X)]$ then $a = 1$. **$\mathbf{P \times Q}$ Lemma:** Let $A = P \times Q$, $P$ a $p-$group, $Q$ a $p'$-group. Suppose $M$ is a $p$-group and $C_M(P) \le C_M(Q)$. Then $Q$ acts trivially on $M$.

**Thompson subgroup:** $A(P)$: abelian subgroups of $P$ of maximal order. $J(P) =< \{A | A \in A(P)\} >$. If $O_p(G) \ne 1$, $G$ is $p$-stable and $p$-constrained, $p \ne 2$. If $P \in S_p(G)$ then $G = O_{p'}(G)N(Z(J(P)))$. **Thompson Factorization:** Let $G$ be solvable with $F(G) = O_p(G)$, $P \in S_p(G)$, $Z = \Omega_1(Z(P))$, $V =< Z^G >$, $G^* \cong G/Z$. The either (i) $G = N_G(J(P))C(Z)$; or (ii) $p \le 3$ and $J(G)^*$ is a direct product of copies of $SL_2(p)$ permuted by $G$ and $J(P)^* \in S_p(J(G)^*)$. Note if $p = 3$ and $G$ has an abelian Sylow $2-$subgroup, so (i) holds. **Thompson Normal $p-$Complement:** Let $p \ne 2$ and $P \in S_p(G)$. Assume $N_G(J(P))$ and $C_G(\Omega_1(Z(P)))$ have a normal $p-$complement then so does $G$. By Burnside transfer, $A \in SCN(p) \rightarrow C_G(A) = A \times Q, Q \in p'(G)$. Property PC: If $G$ is a group in which the normalizer of every $p$ group is $p$-constrained we say $PC(G)$.. **Thompson Transitivity Theorem:** If $PC(G)$ and if $A \in SCN_3(p)$ then $C_G(A)$ permutes all maximal $A$-invariant $q$ groups of $G$, $q \ne p$. Consequence: Under the TTT conditions, if $P \in S_p(G), A \in SCN_3(P)$ and $\forall q \ne p$, $P$ normalizes some $A-$invariant $q-$subgroup of $G$; so if $P$ normalizes no $p'$ subgroup of $G$, neither does $A$. Used to show the **Maximal Subgroup Theorem:** If $P \in S_p(G), SCN_3(P) \ne \emptyset, p \ne 2$ and every element of $N^*(P)$ is $p-$constrained and $p-$stable and $\exists 1 \ne H \lhd P$: $[Q, P] = 1$ if $H \in p'(G)$ and $H^P = H$ then $N^*(P)$ has a unique maximal element.

**Baer-Suzuki:** $X \in p(G)$ then either $X \le O_p(G)$ or $\exists g \in G$ with $< X, X^g >$ not a $p-$group. Thompson (from N-group paper): $G$ is not solvable iff $\exists x, y, z \in G \setminus \{1\}$ with $(|x|, |y|) = (|y|, |z|) = (|x|, |z|) = 1$ such that $xy = z$. If $G$ is a non-abelian simple group all of whose $p-$locals are solvable then $G$ is isomorphic to one of the following: (1) $PSL_2(q), q > 3$, (2) $Sz(q), q = 2^{2m+1}, m \ge 1$ or (3) $A_7$, $PSL(2(3)$, $U_3(3)$, or $M_{11}$.

**Quadratic action:** If $V$ is an abelian $p-$group then $a$ acts quadratically on $V$ if $[V, a, a] = 1$ or $v^{(a-1)^2} = 0$. If $G$ acts quadratically on $V$ then (a) $[v^n, a] = [v, a^n] = [v, a]^n$, (b) $|V| \le |C_V(a)|^2$, (c) $G/C_G(V)$ is an elementary abelian $p-$group. If $G$ acts on an $F_q$ vector space $W \ne 0$, $q = p^m$. Suppose $G =< a, b >$ and $a, b$ act quadratically on $W$, $G/C_G(W)$ is not a $p-$group, $|ab| = p^e k, k \mid (p-1)$ then $\exists \varphi : G \rightarrow SL_2(q)$. $G$ is $p-$stable if $\forall a \in G, [V, a, a] = 1$ implies $aC_G(V) \in O_p(G/C_G(V))$. Let $p \ne 2$ and $G$ be faithful on $V$. Suppose (1) $G =< a, b >$ where $a$ and $b$ act quadratically on $V$ and (2) $G$ is not a $p-$group then (1) the Sylow 2 subgroups of $G$ are not abelian and (2) If $Q$ is a normal $p'$-subgroup of $G$ and $[Q, a] \ne 1$ then $p = 3$ and there is a section of $G$ isomorphic to $SL_2(3)$. If $p \ne 2$. Suppose the action of $G$ on $V$ is faithful and not $p-$stable then (1) the Sylow 2-subgroups of $G$ are non-Abelian and (2) if $G$ is $p-$separable ($G$ is said to be **$p$-separable** if two non conjugate elements of $G$ remain non-conjugate in some finite $p$-group endomorphic image of $G$.) then $p = 3$ and there is a section of $G$ isomorphic to $SL_2(3)$. Suppose $G$ acts faithfully on $V$ and $E_1, E_2$ are two subnormal subgroups of $G$ such that $[V, E_1, E_2] = 1$ then $[E_1, E_2] \le O_p(G)$. Let $G$ be a group and $C_G(O_p(G)) \le O_p(G)$ then $V =< \Omega(Z(S)) | S \in S_p(G) >$ is an elementary abelian normal subgroup of $G$ and $O_p(G/C_G(V)) = 1$.

$Q_8 =< \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} >$. Let $m = max\{|A|, A \in \mathcal{E}(G)\}$, $\mathcal{A}(G) = \{A \in \mathcal{E}(G) | |A| = m\}$

and $J(G) =< A|A \in \mathcal{A}(G)\}$. Let $A \in \mathcal{A}(G)$ acts quadratically on $V$ and $A_0 = [V, A]C_A([V, A])$ then $A_0$ is in $\mathcal{A}(G)$ and acts quadratically on $V$ and if $[V, A] \neq 1$ then $[V, A_0] \neq 1$. **Thompson factorizable** with respect to $p$ if $G = O_{p'}(G)C_G(\Omega(Z(S)))N_G(J(S))$. Let $O_{p'}(G) = 1$ and $V =< \Omega(Z(S))|S \in S_p(G) >$ then $G$ is Thompson factorizable iff $J(G) \leq C_G(V)$.

**Weilandt:** If $A \lhd \lhd G$ and $B \lhd \lhd G$ then $< A, B > \lhd \lhd G$; if $A \lhd \lhd < A, A^g >, \forall g \in G$ then $A \lhd \lhd G$. **Quasi-simple:** $L' = L$ and $L/Z(L)$ is simple. $L$ is a **component** of $H$ if $L \lhd \lhd H$ and $L$ is quasi-simple. Let $Comp(G) = \{H : H$ is a component of $G\}$. $E(G) =< Comp(G) >$ where $H$ is a component of $G$. If $K \in Comp(G), U \lhd \lhd G$ then $K \subseteq U$ or $[K, U] = 1$. $\mathbf{F^*(G) = F(G)E(G)}$. $C_G(F^*) \subseteq F^*(G)$. $G$ is of **characteristic** $p-$**type** if $F^*(H) = O_p(H)$ for every $p-$local, $H$ (Groups of Lie type over characteristic $p$ are, for example.). $G$ is of characteristic $p-$type if $P \in p(G), N = N_G(P) \rightarrow F^*(N) = O_p(N)$. $PSL_n(p^m)$ is of characteristic $p-$type. Let $G$ be a non-abelian simple group, $G$ is of characteristic $p-$type iff $F^*(N(P)) = O_p(N(P))$ for every maximal $p-$local. If $F^*(G)$ is a $p-$group then so is $F^*(N(P)), \forall P \in p(G)$ (use $P \times Q$).

**Amalgams:** $P_1, P_2 \leq G$, $|P_i| < \infty$. Construct a graph $\Gamma(G, P_1, P_2) = \Gamma$ as follows: $\Gamma$ has verticies consisting of right cosets of $P_1$ and $P_2$; the verticies $P_ig_j$ and $P_ng_m$ are joined by an edge if $P_ig_j \neq P_ng_m$ and $P_ig_j \cap P_ng_m \neq \emptyset$. $\Delta(\alpha)$ denotes the verticies adjacent to $\alpha$. $G$ act on graph by right multiplication on cosets. $G \rightarrow Aut(\Gamma)$. $\Gamma$ is connected iff $G =< P_1, P_2 >$. **Theorem:** (a) $G$ has 2 orbits. Every vertex stabilizer $G_\alpha$ is a $G-$conjugate of $P_1$ or $P_2$. (b) $G$ acts transitively on edges of $\Gamma$; every edge stabilizer in $G$-conjugate of $P_1 \cap P_2$. (c) $G$ acts transitively on $\Delta(\alpha)$. $|\Delta(\alpha) : \Delta(\alpha, \beta) = |G_\alpha : G_{\alpha,\beta}|, \beta \in \Delta(\alpha)$. (d) $(P_1 \cap P_2)_G$ (the largest normal subgroup of $G$ in $P_1 \cap P_2$) is the kernal of the action of $G$ on $\Gamma$. **Condition** $\mathcal{A}$: Let $G$ be a finite group generated by $P_1, P_2$, $T = P_1 \cap P_2$ satisfying: $C_{P_i}(O_2(P_i)) \leq O_2(P_i)$, $T \in S_2(P_i)$, $T_G = 1$, $P_i/O_2(P_i) \approx S_3$ and $[\Omega(Z(T)), P_i] \neq 1$. **Goldschmidt:** If $\mathcal{A}$ holds either (i) $P_1 \approx P_2 \approx S_4$ or (ii) $P_1 \approx P_2 \approx C_2 \times S_4$.

Some examples motivating components and classification by **centralizers of involutions:** Brauer proved if $G = PSL_3(q), q = 3 \pmod 4$ and $x \in Inv(G)$ then $C_G(t) \cong GL_2(q)$ and that the converse is true for $q > 3$; if $q = 3$ other possibilities are $PSL_3(3)$ and $M_{11}$. Classifications fall into two steps: (I) Given $H = C_G(t), t \in Inv(G)$, find $|G|$ and its structure and (II) find $C(t)$ for simple groups. Note that all simple groups are determined by their character table. Step (I) consists of two steps: (A) $\forall v \in Inv(H)$, determine $C_G(v)$ and the fusion patterns of $Inv(C_G(v))$, (B) if $G$ has more than one conjugacy class, this determines the order, if not we must examine all if $H$ using characters. Let $L = SL_n(q), G = PSL_n(q) = L/Z(L)$, $t \in Inv(G)$ corresponds to $T \in L$ with $T^2 = \lambda I_n$ putting $Z = \{\lambda I_n, \lambda^n = 1\}, d = |Z| = (n, q - 1)$ and $C = \{X \in L : XT = \mu TX\}$, $C_G(t) = C/Z$. Let $p \neq 2$ and the eigenvalues of $T$ be $\rho, -\rho$ then $T$ is conjugate to $\begin{pmatrix} \rho I_r & 0 \\ 0 & -\rho I_s \end{pmatrix}$, or $\begin{pmatrix} 0 & I_m \\ -\lambda I_m & 0 \end{pmatrix}$, depending on whether the minimum polynomial is $(x + \rho)(x - \rho)$ or $(x^2 - \lambda)$ which depends on whether the eigenvalue is in $GF(q)$ or $GF(q^2) \setminus GF(q)$. Let $X \in C$ with $X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}$, so either $X_2 = X_3 = 0$ and $det(X_1)det(X_4) = 1$ or $r = s$ and $X_1 = X_4 = 0$ and $det(X_2)det(-X_3) = 1$; let $\delta : X \mapsto det(X_1), K = ker(\delta)$ then $K = SL_r(q) \times SL_s(q)$. Put $E = KZ/Z$, $E \lhd C/Z$ and $E = K/(K \cap Z)$ and $E$ is a central product.

Here are a bunch of results on the centralizers of the classical groups: Let $G = PSL_n(q)$, $q$ odd, $t \in Inv(G)$, (1) if $n$ is odd $\exists N \lhd C(t)$ with $N$ the minimal central product of $SL_r(q)$ and $SL_s(q)$, $r + s = n$ (type *) and both $C(t)/N$ and $Z(N)$ are cyclic groups with orders dividing $q - 1$; (2) if $n$ is even there is a centralizer as above and centralizers of two additional types: (A) $\exists C_0 : |C(t) : C_0| = 2$ and $E \lhd C(t)$ of type * with $r = s$ and $C(t)/E$ is **dihedral** and $C_0/Z$ and $Z(E)$ are cyclic — there is an element of order 2 outside $C_0$ that interchanges the factors of $E$, (B) $\exists C_0 : |C(t) : C_0| = 2$ and $E \lhd C(t)$ of type * with $r = s$ and $E/Z(E) \cong PSL_r(q^2)$ and $Z(E)$ is cyclic with order dividing $q + 1$ and $C(t)/E$ is dihedral of order $q + 1$ or $2(q + 1)$; further, there is an element of order two in $C(t) \setminus C_0$ which transforms elements in $E/Z(E)$ like the element of order 2 in the Galois group of $GF(q^2)/GF(q)$. If $G = PSp_{2m}(q)$ with $q$ odd and $t \in Inv(G)$ then either (1) $C(t)$ is a minimal central product of $Sp_{2r}(q)$ and $Sp_{2s}(q)$ with $r + s = m, r \neq s$, or (2) $\exists C_1 \lhd C(t)$ with $C_1$ a minimal central product of two copies of $Sp_{2l}(q), 2l = m$ and there is an element of order two in $C(t) \setminus C_1$ that interchanges the two, or (3) $\exists C_1 \lhd C(t)$ with $C_1 \cong GL_m(q)/\{\pm I\}$ and there is an element of order two in $C(t) \setminus C_1$ that corresponds to $A \mapsto^t A^{-1}$ and $q = 1 \pmod 4$, or (4) $\exists C_1 \lhd C(t)$ with $C_1 \cong U_m(q)/\{\pm I\}$ and there is an element of order two in $C(t) \setminus C_1$ that corresponds to $A \mapsto A^\tau$ and $q = 3$

(mod 4), $\tau$ the generator of the Galois group. If $G = PSU_n(q)$ with $q$ odd and $t \in Inv(G)$ then either (1) $\exists N \lhd C(t)$ with $N$ a minimal central product of $SU_r(q)$ and $SU_s(q)$ with $r+s=m, r \neq s$, both $C(t)/N$ and $Z(N)$ are cyclic with orders dividing $q+1$, (2) if $n$ is even there is a centralizer as above and centralizers of two additional types: (A) $\exists C_0 : |C(t) : C_0| = 2$ and $E \lhd C(t)$ of type * with $r = s$ and $C(t)/E$ is dihedral and $C_0/Z$ and $Z(E)$ are cyclic — there is an element of order 2 outside $C_0$ that interchanges the factors of $E$, (B) $\exists C_0 : |C(t) : C_0| = 2$ and $E \lhd C(t)$ with $r = s$ , $Z(E)$ cyclic of order dividing $q-1$ and $E/Z(E) \cong PSL_r(q^2)$ and there is an element of order two in $C(t) \setminus C_1$ that corresponds to $A \mapsto^t (A^\tau)^{-1}$ $\tau$ the generator of the Galois group. If $G = P\Omega_n(q)$ with $q$ odd and $t \in Inv(G)$ then either (1) $\exists E \lhd C(t)$ with $C(t)/E$ solvable, $E' = E$ and $E$ is either $SL_m(q)/\{\pm I\}$ and $SU_m(q)/\{\pm I\}$ ($2m = n$ in both cases) or a central product of $\Omega_r(q)$ and $\Omega_s(q)$. For $G = A_n$, let $H_1 = \Sigma_k, H_2 = Z_2 \wr \Sigma_l$ and $C(t) = H_1 \times H_2$ with $(\sigma, \rho) \in C(t), sign(\sigma) = sign(\rho)$.

Since $C(F^*(G)) \subseteq F^*(G)$, $G \rightarrow Aut(G)$ has kernel $Z(F^*(G))$; further, $F^*(G)$ is uncomplicated and its embedding in $G$ is well behaved. Want to study relationship of $F^*(G)$ and its $p-$locals. Hard when $F^*(G)$ is a $p-group$ but then we can use Thompson factorization. Thompson $p-$complement $\rightarrow$ nilpotence of Frobenius kernel.

Let $X/Z(X)$ be a non-abelian simple group then $X = X'Z(X)$ and $X'$ is quasi-simple. Let $X$ be quasi-simple and $H \lhd \lhd X$, then $X = H$ or $H \leq Z(X)$. $H \lhd \lhd X \rightarrow Comp(H) = Comp(x) \cap H$. $L \in Comp(G), H \lhd \lhd G$, then $L \in Comp(H)$ or $[L, H] = 1$. Distinct components commute. Let $L \in Comp(G)$, $H$ and $L-$ invariant subgroup, then (a) $L \in Comp(H)$ or $[L, H] = 1$, (b) If $H$ is solvable, $[L, H] = 1$. $E^* = E(G)/Z(E(G))$ then (a)$Z = Z(L) : L \in Comp(G) >$, (b) $E^*$ is a direct product of $< L : L \in Comp(G) >$, $E$ is a central product of its components.

**Signalizers:** $r$, prime, $G$ finite and $A$ an abelian $r-$subgroup of $G$. An $A-$signalizer is a map $\theta : A^\# \rightarrow \mathcal{S}$ where $\mathcal{S}$ is a set of $r'$ $A-$invariant subgroups such that $a, b \in A^\#$ and $\theta(a) \leq C_G(a)$ and $\theta(a) \cap C(b) \leq \theta(b)$. $\theta$ is complete if $\exists \theta(G)$ an $r'$, $A-$invariant subgroup such that $\theta(a) = C_{\theta(G)}(a)$ for each $a \in A^\#$. $\theta(a) = C_X(a)$ is one such function; if $m(A) \geq 3$ then every $A-$signalizer functor is complete. Under these conditions, for a solvable $A-$signalizer, $\mathcal{N}_\theta(A)$ has a unique maximal element. Goldschmidt proved this for solvable signalizer functors.

$O_{p'}(G)$ is called the $p$**-core** of $G$. $O_{2'}(G)$ is often called the **core** of $G$. Walter: Let $G$ be a group with 2 rank $\geq 5$ and $O_{2'}(G) = 1$ with the property that the centralizer of every involution is $2-$constrained then $O_{2'}(C(x)) = 1$ for every involution $x$.

Semi-regular action: $C_G(a) = 1, \forall a \in A^\#$. Suppose $A$ acts semi-regularly on $G$. Then (1) $|G| = 1 \pmod{|A|}$, (2) $A$ is semi-regular on each $A-$invariant subgroups factor group of $G$, (3) $\forall p \in \pi(G), \exists! A-$invariant Sylow $p-$subgroup of $G$, (4) $\forall a \in A, g \mapsto [g, a]$ is a permutation of $G$, (5) if $2||A|, \exists t : |t| = 2, t \in A : g^t = g^{-1}, g \in G$ and $G^{(1)} = 1$.

Let $p, q \in \pi(A)$ then for $S \subseteq A$. (1) $p \neq 2$, $S_p(A) \rightarrow S$ is cyclic. (2) $S \in S_2(A)$ is cyclic or quaternion. (3) $|S| = pq \rightarrow S$ is cyclic. (4) $|S| = 1 \pmod 2 \rightarrow S$ is metacyclic.

If $x, y$ are two involutions in $G$ then $< x, y >$ is dihedral of order $2|xy|$. Let $G$ be even order with $Z(G) = 1$, let $m$ be the number of involutions in $G$ and $n = |G|/m$. Then $G$ possesses a proper group of order at most $2n^2$.

Let $G$ be a simple group of even order, $t$ and involution and $n = |C_G(t)|$. Then $|G| \leq (2n^2)!$. From this we get: **Brauer-Fowler:** Let $H$ be a finite group. There are at most a finite number of finite simple groups with $H \cong C_G(t)$.

**Feit-Thompson:** The only finite simple groups or odd order are $\mathbb{Z}_p, p \neq 2$. The proof follows the CN classification.

**Thompson Order Formula:** Assume $G$ has more than two congugacy classes of involutions $\{{x_i}^G\}$ and let $n_i$ be the number of ordered pairs $(u, v)$ with $u \in {x_1}^G, v \in {x_2}^G$ and $x_i \in < uv >$ then $|G| = |C(x_1)||C(x_2)| \sum_{i=1}^k \frac{n_i}{|C(x_i)|}$.

Let $\Omega$ be a collection of subgroups. Define $\mathcal{D}(\Omega)$ as the graph formed by joining $A, B \in \Omega$ if $[A, B] = 1$. If $k > 0$ let, $\mathcal{E}^p_k(G)$ be the elementary abelian subgroups of $p$-rank at least $k$. $G$ is said to be $k-connected$ for prime $p$ if $\mathcal{D}(\mathcal{E}^p_k(G))$ is connected.

If $G$ is a non-abelian finite simple group with $m_2(G) \leq 2$ then either (1) a Sylow 2-group is either dihedral, semi-dihedral or $Z_{2^n}\ wr\ Z_2$ and $G \cong L_2(q)$, $G \cong L_3(q)$, $G \cong U_3(q)$ $q$,odd, or $M_{11}$; or, (2) $G \cong U_3(4)$. Note that $Q_8 \in S_2(SL_2(3))$ and $\left(\begin{array}{cc} 2 & 0 \\ 0 & 2 \end{array}\right)$ is the unique involution.

If $G$ is a non-abelian finite simple group with $m_2(G) > 2$ and assume $G$ has a proper 2-generated 2-core, then either $G$ is a group of Lie type of characteristic 2 and Lie rank 1 or $G \cong J_1$.

**Glauberman $ZJ$:** If $C_G(O_p(G)) \leq O_p(G)$ and the action of $G$ on its chief factors of $G$ is $p-$stable then $G = N_G(Z(J(S)))$. Every group admitting a fixed-point-free automorphism of prime order is nilpotent. **Glauberman's $Z^*$ Theorem:** Let $G$ be a finite group and $t$ and involution in $G$ which is weakly closed in $C(t)$. Then $t^* \in Z(G^*)$ where $G^* = G/O_{2'}(G)$.

**$B_p$ property:** Suppose $O_{p'}(G) = 1$ and $x \in G, |x| = p$ then $O_{p',E}(C(x)) = O_{p'}(C(x))E(C(x))$. A **standard subgroup** for the prime $p$ is a group $H = C_G(x), |X| = p$ such that $H$ has a unique component, $L$, and $C_G(L)$ has a cyclic Sylow $p-$group. **Component Theorem:** Let $G$ be a finite group with $F^*(G)$ satisfying the $B_2$ property and with in involution, $t$ such that $O_{2',E}(C(t)) \neq O_{2'}(C(t))$ then $G$ possesses a standard subgroup for the prime 2. **Standard Form** problem for $(L, r)$: Determine all finite groups, $G$, possessing a standard subgroup $H$ for the prime $r$ with $E(H) \cong L$.

Let $G$ be a minimal counter-example to the classification theorem and assume $G$ is generic of even characteristic. Then one of the following holds: (1) $G$ possesses a standard subgroup for some $p \in \sigma(G)$; (2) there is an involution $t \in G$ such that $F^*(C(t))$ is a 2-group of symplectic type; or, (3) $G$ is in the uniqueness case.

In real simple groups $O_{2'}(C(t))$ is cyclic and almost central. **Bender's Theorem:** For any group $X$, we have $C_X(F^*(X)) \leq F^*(X)$ and if $W \lhd X$ and $C_X(W) \leq W$ then $E(X) \leq W$. If $O_{p'}(X) = 1$ then $F(X) = O_p(X)$ and every component of $X$ has order divisible by $p$ so $X$ is $p$-constrained iff $E(X) = 1$ or, equivalently, $C_X(O_p(X)) \leq O_p(X)$. Let $\overline{X} = E(X/O_{p'}(X))$, $L$ is a minimal normal subgroup subject to $\overline{L} = E(\overline{X})$, $\overline{L_i}$ is a component of $E(\overline{X})$, $L_i = O^{p'}(L_i)$, $[L_i, L_i] = L_i$ and $[L_i, L_j] \leq O_{p'}(X)$, $L$ is called the **$p$-layer**. $F^*(X)$ controls embedding of $X$ of $p'$-cores and the $p$-layer of every $p$-local. $O_\pi((X/O_\pi(X))) = 1$. If $O_\pi(X) = 1$ then $F(X)$ is divisible by $p \in \pi$ and every component is divisible by some $p \in \pi'$.

Recall signalizers. The idea is that $A-$invariant $p'$ subgroups of $G$ can be glued into a single $p'$ subgroup $\theta(G, A)$ which is either normal or strongly $p-$embedded in $G$. $M \subseteq G$ is **strongly $p-$embedded** if $p||M|$ but $p$ does not divide $|M \cap M^g|$ for $g \in G - M$. **Tightly embedded:** $p = 2$. If $M$ is strongly embedded, $G$ fixes one point when acting on the cosets of $M$. Bender identified all simple groups with strongly 2-embedded subgroups, namely, $SL_2(2^n), SZ(2^n), PSU_3(2^n)$. No simple group of $p - rank \geq 3$ has a strongly $2-$embedded $2'$ local subgroup.

Let $G$ be a finite simple group and $S \in S_2(G)$ then one of the following holds: (a) $S$ is **dihedral**, (b) $S$ is semidihedral, (c) $G$ has a strongly embedded subgroup, (d) $S$ has a non-cyclic characteristic elementary abelian subgroup, $A$, and $E = N_G(A)$ has conjugacy classes, $< z_i^G >$, that do not fuse in $G$ such that $G =< E, C_G(z_i) >$. If $G$ is a finite simple group and $H < G$ with $\mathbb{Z}(H)$ of even order and $h \approx C_H(z)$ then $G$ is said to be of $H$-type. Note we can construct a faithful transitive permutation representation of $G$ given a presentation of $H$. A group has an $H$-satellite if there are non-isomorphic groups of $h$-type. A finite simpe group, $G$, is uniquely determined by $C_H(z)$ for a $2-$central involution, $z$, if $G$ does not have any non-isomorphic $H$-satellites.

**Netto:** Let $x, y \in S_n$ be selected randomly. $Pr[< x, y >= S_n] = \frac{3}{4}$. **Irreducible characters of the symmetric group $S_n$:** $n = n - m, \mu_1, \ldots, \mu_j$, $d_n(\mu)$ is the dimension of the irreducible character determined by: $l_{j+1} = \mu_j$, $l_j = \mu_{j-1} + 1$, $l_{j-1} = \mu_{j-2} + 2$, $\ldots$, $l_1 = n - m + j$. $d_n(\mu) = \prod_{s>r}(l_r - l_s)$.

## 1.3 Algebraic Geometry

### 1.3.1 Basics

Every conic in the affine space over $R$ is equivalent under an affine transformation to one of the following: (1) $X^2+Y^2+P=0$ (ellipse, point, empty set), (2) $X^2-Y^2+P=0$ (hyperbola, intersecting lines), (3) $X^2+Y+P=0$ (parabola), (4) $X^2+P=0$ (parallel lines, point empty). In projective space (1), (2), (3) are equivalent. In the projective space over $\mathbb{C}$, they are all projectively equivalent. $f(x,y)$ is rational if $\exists \phi, \psi$: $f(\phi(t), \psi(t))=0$. Any conic (2nd order equation) in 2 variables has either infinitely many rational solutions or none.

$rad(I) = \sqrt{I} = \{a : a^n \in I\}$. Radical ideals $\leftrightarrow$ varieties, prime ideals $\leftrightarrow$ subvarieties, maximal ideals $\leftrightarrow$ points. $\Gamma[V] = k[x_1, x_2, \ldots, x_n]/I(V)$. $\Gamma(V)$ is the quotient field of $\Gamma[V]$. $\mathfrak{O}_P(V)$ denotes the rational functions on $V$ defined on $P$. $k \subseteq \Gamma(V) \subseteq \mathfrak{O}_P(V) \subseteq \Gamma(V)$. $\mathfrak{M}_P(F)$ denotes the maximal ideal of $\mathfrak{O}_P(F)$. $0 \rightarrow \mathfrak{M}^n/\mathfrak{M}^{n+1} \rightarrow \mathfrak{O}/\mathfrak{M}^{n+1} \rightarrow \mathfrak{O}/\mathfrak{M}^n \rightarrow 0$. $\chi(n) = dim(\mathfrak{O}/\mathfrak{M}^n)$ = Hilbert polynomial. **Pullback:** $\phi : A^n(k) \rightarrow A^m(k)$, $f \in k[y_1, \ldots, y_m]$; the pullback $\phi^* : \phi^* \circ f = f \circ \phi$. **DVR:** Noetherian, local, maximal ideal is principal. If a form, F, does not vanish on an irreducible projective variety X then $dim(X_F) = dim(X) - 1$. $M_P$ is the maximal ideal associated with $(T-P)$.

**Intersection multiplicity:** Multiplicity of root of $f(t) = gcd(F_1(ta), \ldots, F_m(ta))$. L touches X at O if its intersection multiplicity is greater than 1. Locus of points touching X at x is the tangent space, $\Theta_{x,X}$.

Let $k$ be algebraically closed. An affine irreducible algebraic set is an *algebraic variety*. There is a one to one correspondance between polynomial maps $\varphi : V \rightarrow W$. and the homomorphisms $\tilde{\varphi}\Gamma[W] \rightarrow \Gamma[V]$. Let $\mathcal{T}(V,k) = \{f : f : V \rightarrow W\}$. If $\varphi : V \rightarrow W$, $\tilde{\varphi} : \mathcal{T}(W,k) \rightarrow \mathcal{T}(V,k)$. Two affine varieties $V, W$ are isomorphic if $\exists \phi, \psi : \phi \circ \psi = id_W$. The following are equivalent: (1) The set of non-units in $R$ form an ideal; (2) $R$ has a unique maximal ideal. The following are equivalent and define a discrete valuation ring (DVR): (1) $R$ is Noetherian and its maximal ideal is principal; (2) $\exists t \in R : \forall 0 \neq z \in R : z = ut^n$, where $u$ is a unit. Reimann Roch: Let $X$ be a non-singular projective plane curve. $\exists g \geq 0 : \forall D, dim_k(L(D)) \geq deg(D) + 1 - g$. The minimum such $g$ is called the genus. A variety is *rational* if it is birationally equivalent to $A^n$ for some $n$.

A closed set is union of solutions of polynomial equations. Every closed set is the union of finitely many irreducible ones. Every irreducible closed set is birationally isomorphic to a hypersurface in $A^n$. Two curves are **birationally equivalent** iff their fields of functions are isomorphic. $k[X] = k[T]/U_X$. Every irreducible curve of degree 2 is rational. $x^n + y^n = 1$ is not rational for $n > 2$. Let $C$ be a plane curve with only ordinary multiple points, $r_P = m_P(C)$ and $n = deg(C)$ then $g = \frac{(n-1)(n-2)}{2} - \sum_{P \in C} \frac{r_P(r_P-1)}{2}$.

**Weak Bezout:** If two curves of dimension $m$ and $n$ meet at more than $mn$ points (counting multiplicity) then they have a common component. Strategy of Proof: (S-1) $\#(C_1 \cap C_2 \cap \mathcal{A}^2) \leq dim(\frac{R}{(f_1,f_2)}) \leq n_1 n_2$, (S-2) first inequality is an equality, (S-3) first inequality can be strengthened to $I(C_1 \cap C_2, P) \leq dim(\frac{R}{(f_1,f_2)})$, (S-4) inequality in 4 is an equality, (S-5) $I$ is invariant under projective transformations — transform so the line at infinity does not intersect $C_1 \cap C_2$. Notation: Let $f_1(x,y)$, and $f_2(x,y)$, defining curves $C_1$ and $C_2$, have dimension $m, n$ respectively. $R = k[x,y]$, $(f_1(x,y), f_2(x,y)) = Rf_1 + Rf_2$.

S-1: $C_1 \cap C_2 \leq dim_k(\frac{R}{(f_1,f_2)}) \leq mn$. [Argument: If $P_1, P_2, \ldots P_r$ are distinct, $\exists h_i(x,y)$ with $h_i(P_j) = \delta_{ij}$, so if there are $r$ common root of $f_1$ and $f_2$, $\sum_{i=1}^r c_i h_i(x,y) = r_1 f_1(x,y) + r_2 f_2(x,y)$ implies $c_i = 0$.]
Let $R_d$ be polynomials of degree $\leq d$ then $dim_k(R_d) = \phi(d) = \frac{(d+1)(d+2)}{2}$. Let $W_d = R_{d-m}f_1 + R_{d-n}f_2$, for $d \geq (m+n)$. $R_{d-m}f_1 \cap R_{d-n}f_2 = R_{d-m-n}f_1 f_2$. $dim_k(R_d) - dim_k(W_d) = mn$. $g = \sum_i^l c_j g_j$ has a non-trivial dependency for $l > mn$ with $g \in W_d$.
S-2: Second inequality is equality if $C_1 \cap C_2$ don't meet at infinity. Let $f^*$ be the homogeneous polynomial consisting of the highest degree terms in $f$. If $\infty \notin C_1 \cap C_2$ then $f_1^*$, $f_2^*$ have no common factor. If $f_1^*$ and $f_2^*$ have no common factor then $(f_1, f_2) \cap R_d = W_d$. Under the conclusion of the previous sentence, if $d \geq n_1 + n_2$ then $dim(\frac{R}{(f_1,f_2)}) \geq n_1 n_2$ which proves the result.
Define $O_P = \{F \in K(x,y) : F(P) \text{ exists }\}$, $M_P = \{f \in O_P : f(P) = 0\}$. $M_P$ is a unique maximal ideal of $O_P$. $(f_1, f_2)_P = f_1 O_P + f_2 O_P$. Now define $I(C_1 \cap C_2; P) = dim(\frac{O_P}{(f_1,f_2)_P})$.
S-3: $\frac{O_P}{(f_1,f_2)_P} \leq \frac{R}{(f_1,f_2)} < \infty$. $O_P = (f_1,f_2)_P + R$. If $P \notin C_1 \cap C_2$ then $I(C_1 \cap C_2, P) = 0$; If $P \in C_1 \cap C_2$ then

$(f_1, f_2)_P \subset M_P$; $I(C_1 \cap C_2; P) = 1 + dim(\frac{R}{(f_1,f_2)_P})$ iff $(f_1, f_2) = M_P$. If $P \in C_1 \cap C_2$ and $r \geq dim(\frac{O_P}{(f_1,f_2)_P})$ then $M_P^r \subset (f_1, f_2)_P$. If $P, Q \in C_1 \cap C_2 \cap \mathcal{A}^2, \psi \in O_P$ then $\exists g \in R$: $g = \psi \pmod{(f_1, f_2)_P}$ and $g = 0 \pmod{(f_1, f_2)_Q}$ if $P \neq Q$.
S-4: Kernel of natural map $R \rightarrow \prod_{P \in (C_1 \cap C_2 \cap \mathcal{A}^2)} \frac{O_P}{(f_1,f_2)_P}$ is just $(f_1, f_2)$ where the natural map is: $f \mapsto (\ldots, f \pmod{(f_1, f_2)}, \ldots)$. $dim(\frac{R}{J}) = \sum_P dim(\frac{O_P}{(f_1,f_2)}) = \sum_P I(C_1 \cap C_2, P)$. The last equality holds iff $J \subset (f_1, f_2)$. Define $L = \{g \in R : gf \in (f_1, f_2)\}$ and $1 \in L$. $L$ is an ideal $(f_1, f_2) \subset L \subset R$. $P \in \mathcal{A}^2$, $\exists g \in L : g(P) = 0, P \in L$. $\exists a \in k : 1 \notin L + R(x - a)$ and $\exists b \in k : 1 \notin L + R(x - a) + R(y - b)$.
S-5: Properties of intersection multiplicity. $I((y - x^m), y; 0) = m$. Show the definitions make sense and that there is a line $L$ which does not contain any of the intersection points. The proof requires knowing there are only a finite number of points in the intersection.

**Genus** for non-singular curve: $g_f = \frac{(n-1)(n-2)}{2} - d$. $L(D) = \{f : K(C)^* : div(f) \geq -D\}$. $l(D) = dim(L(D))$.
**Riemann-Roch:** $l(d) = l(K - D) + deg(D) - g + 1$.

The $r$ forms $f_1, f_2, ..., f_r$ with indeterminate coefficients possess a resultant system of integral polynomials $b_k$ such that for special values of the coefficients in $K$ (algebraically closed). The vanishing of all resultants is a necessary and sufficient condition for $f_1 = f_2 = ... = f_r = 0$ to have a solution $\neq 0$. The $b_k$ are homogeneous in the coefficients of every form $f_i$ and satisfy $x_k^{s_r} b_k = 0 \pmod{(f_1, f_2, ..., f_r)}$.

**Bezout's Theorem.** If $f, g$ are two curves of degree $n, m$ respectively that have no common component then they intersect in $mn$ points counting multiplicity. Notes: A homogeneous system $f_1 = f_2 = ... = f_r = 0$ has solutions $(\xi_1^{(a)}, \xi_2^{(a)}, ...\xi_n^{(a)}), a = 1, 2, 3, ..., q$. Set $l_x = u_1 x_1 + u_2 x_2 + ... + u_n x_n$. Form resultant system $b_1(u), ..., b_t(u)$. The common zeros of $b_1, ...$ are $\prod l_a$. By Nullstellensatz, $(\prod_a l_a)^\tau = 0(b_1(u), b_2(u), ..., b_t(u))$ $\rightarrow D(u) = \prod {l_a}^{\rho_a}$ and $(b_i(u))^{r_i} = 0(\prod l_a) \rightarrow D(u) = (f_1, ..., f_r, l)$. $R(u)$ is the same as the u-resultant so $\sum \rho_a$ is the degree of $R(u) = \prod deg(f_i)$.

Example: $F_1(x, y, z) = x^2 + y^2 - 10z^2 = 0$, $F_2(x, y, z) = x^2 + xy + 2y^2 - 16z^2 = 0$, add $F_3(x, y, z) = u_0 z + u_1 x + u_2 y$. $Res_{1,2,2}(F_0, F_1, F_2) = (u_0 + u_1 - 3u_2)(u_0 + 2\sqrt{2}u_1 + \sqrt{2}u_2)(u_0 - 2\sqrt{2}u_1 - \sqrt{2}u_2)$. Solutions are $(1, -3, 1)$, $(-1, 3, 1)$, $(2\sqrt{2}, 2\sqrt{2}, 1)$, $(-2\sqrt{2}, -2\sqrt{2}, 1)$.

Proofs with Generics. Example: $F_1(X, Y, Z) = X - Y^2$, $F_2(X, Y, Z) = XY - Z$, $(X, Y, Z) \rightarrow (t^2, t, t^3)$ is generic because it is a solution for any specialization of $t$ and any solution is obtainable this way.

Let $D$ be a domain and $\Omega = \Omega_D = \overline{D(t_1, t_2, ...)}$ is called a universal field. Note that $\Omega \leftrightarrow$ prime ideals over $D[X_1, ...]$.

**Theorem:** $x_1, ...x_n \in \Omega$. $I = \{f : f(x_1, ..., x_n) = 0\}$ is a prime ideal. If $I$ is a prime ideal and $1 \notin I$ then $I$ has a generic 0. Any extension $K(\alpha_1, ..., \alpha_m)$ can be embedded in $\Omega$.
Hints: look at $E = D[X]/I$. Under this homomorphism the image of $(X_1, ...X_n)$ is generic.

**Theorem:** If $\xi_1, ..., \xi_n$ are elements of an arbitrary extension of $K$ then If $\Re = K[X_1, ..., X_n]$ and $\wp = \{f : f(\xi_1, ..., \xi_n) = 0\}$. $1 \notin \Re$ and $\wp$ is a prime ideal. Every prime ideal has a generic element.

**Theorem:** Any ideal $g = (f_1, ..., f_n)$ which has no zeros in $\Omega$ is the unit ideal. Proof: Otherwise a maximal ideal would correspond to a non-zero generic point.

Extension of **Nullstellensatz**: If $p_1, ..., p_s$ all vanish at the common zeros of $(f_1, ..., f_n)$, then $\exists q$ such that powers of the $p_i$'s of degree $q$ are in $(f_1, ..., f_n)$. Proof: For $s = 1$, this is the simple Nullstellensatz. Let the exponent for each $i$ be $q_i$. Set $q = q_1 + q_2 + ... + q_n - n + 1$. Nullstellensatz bound: $\rho \leq 13d^n$ where $d$ is the degree and $n$ is the number of variables.

Let $N_q$ be the number of products $X_j$ of degree $q$. **Theorem:** Suppose $F_1, F_2, ..., F_r$ are forms. $(0, ..., 0)$ is the only common zero iff all products $X_j$ can be expressed as linear combinations of the $X_{ki} F_i$ with coefficients in $K$. Note: This means they are linearly independent. So there are other common zeros is there are fewer than $N_q$. Note that $X_1, ..., X_n$ satisfy the Extension conditions. If the $X_{ki} F_i = \sum a_{kij} X_j$ are not linearly independent, the determinant families, $R_i(a)$, form a resultant set.

**Multivariate resultants:** If we fix degrees $d_0, d_1, \ldots d_n$ then there is a unique polynomial $Res \in \mathbb{Z}[u_i, \alpha]$ such that (a) if $F_0, F_1, \ldots, F_n$ are homogeneous polynomials of degrees $d_0, d_1, \ldots d_n$ then $F_0 = \ldots = F_n = 0$ has a nontrivial solution over $\mathbb{C}$ iff $Res(F_0, \ldots, F_n) = 0$, (b) $Res(x_0^{d_0}, \ldots, x_n^{d_n}) = 1$, (b) , (c) $Res$ is irreducible in $\mathbb{C}[u_i, \alpha]$. If $PP = PP(x_1, x_2, \ldots, x_n)$ is a set of power products in the $x_i$, there are $N_m = \binom{m+n-1}{n-1}$ $PP$'s of degree $m$. Example: $A_3 = a_3x^2b_3y^2 + c_3z^2$, $A_2 = a_2x + b_2y + c_2z$, $A_1 = a_1x + b_1y + c_1z$. $S_i = \frac{PP_i^d}{x_i^d}$, $S_1 =< x^2, xy, xz >, S_2 =< y^2, yz >, S_3 =< z^2 >$.

$$\left(\begin{array}{c|cccccc} & x^2 & xy & xz & y^2 & yz & z^2 \\ \hline xA_1 & a_1 & b_1 & c_1 & 0 & 0 & 0 \\ yA_1 & 0 & a_1 & 0 & b_1 & c_1 & 0 \\ zA_1 & 0 & 0 & a_1 & 0 & b_1 & c_1 \\ yA_2 & 0 & a_2 & 0 & b_2 & 0 & 0 \\ zA_2 & 0 & 0 & a_2 & 0 & b_2 & c_2 \\ A_3 & a_3 & 0 & 0 & b_3 & 0 & c_3 \end{array}\right).$$

### 1.3.2 Elliptic Curves

Elliptic Curves: $Y^2Z = X^3 + aXZ + bZ^3$, $P_i = (x_i, y_i)$, $O = (0 : 1 : 0)$. We want to calculate $R = P_1 + P_2$. If $P_1$ or $P_2$ is $O$, result is obvious. If $x_1 = x_2$ and $y_1 = -y_2$, $R = O$. If $x_1 \neq x_2$, set $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$. If $x_1 = x_2$ and $y_1 \neq -y_2$, set $\lambda = (3{x_1}^2 + a)(y_1 + y_2)^{-1}$. In either case, $x_3 = \lambda^2 - x_1 - x_2$, $y_3 = \lambda(x_1 - x_3) - y_1$ and $R = (x_3 : y_3 : 1)$. $|\epsilon_p| \leq 2\sqrt{p}$. Multiple roots iff $-(4a^3 + 27b^2) = 0$. Usually pick $Z$ axis tangent to $O$, or $(0, 1, 0)$ as the point at $\infty$, If this intersects $C$ at $P$, pick $X$ axis tangent to $C$ at $P$.

**Mordell:** If a non-singular cubic curve has a rational point then the rational points are finitely generated as a $k$-module. Use $H(\frac{m}{n}) = max(|m|, |n|)$. Let $P = (x, y)$. Define $H(P) = H(x)$ and $h(P) = log(H(P))$. From now on assume $C$ is given by $y^2 = x^3 + ax^2 + bx + c$.

To prove it, need four lemmas: Lemma 1: There are a finite number of points $P$: $h(P) < M$. Lemma 2: Fix $P_0$ on $C$, $\exists K_0(P_0, a, b, c) : h(P + P_0) \leq 2h(P) + k_0$. Show that if $P$ is on $C(Q)$, $P = (\frac{m}{e^2}, \frac{n}{e^3})$. Then show $n \leq KH(P)^{\frac{3}{2}}$. Use this to get $k_0$. Lemma 3: Fix $\exists K(a, b, c) : h(2P) \geq 4h(P) - K$. Lemma 4: $|\{C(Q) : 2C(Q)\}| < \infty$.

For lemma 4, assume $y^2 = x^3 + ax^2 + bx$ (so the curve always has a rational point), and use $\Gamma = C(Q)$ and $\Delta = 2\Gamma$. Define the map $\phi(x, y) = (x + a + \frac{b}{x}, y\frac{x^2 - b}{x^2})$. Define $\psi$ similarly. Note that $\psi(\phi(P)) = 2P$ and $ker(\phi) = \{0, (0, 0)\}$. ${Q^*}^2$ $\alpha(x, y) = x \pmod{{Q^*}^2}$. $im(\phi) \subseteq ker(\alpha)$. Let $p_i|b$, $i = 1, 2, \ldots t$ then $|\Gamma : \phi(\Gamma)| \leq 2^{t+1}$. $|\Gamma : \phi(\Gamma)| \leq 2^{t+1}$. Use the following lemma: If $A$ and $B$ are abelian $A \rightarrow B \rightarrow A$ and $|B : \phi(A)| < \infty$, $|A : \phi(B)| < \infty$, then $|A : 2A| \leq |B : \phi(A)||A : \phi(B)|$.

Proof given lemmas: Let $Q_0, \ldots, Q_{m-1}$ be the coset representatives. $P - Q_{i_1} = 2P_1$ is in the subgroup, $P_1 - Q_{i_2} = 2P_2$, repeatedly doing this yields: $P = Q_{i_1} + 2Qi_2 + \ldots + 2^{m-1}Q_{i_m} + 2^mP_m$, $h(P_j) \leq \frac{3}{4}h(P_{j-1})$. Since there are a finite number of $Q_i$ there's a $k'$ so that $h(P - Q_i) \leq 2h(P) + k'$ for all $P$. Using the inequalities $h(P_j) \leq \frac{h(P_{j-1})}{2} + \frac{k+k'}{4}$. So the group is generated by the $Q_i$ and the (finite number of) points of $ht \leq \frac{k+k'}{4}$.

Let $C$ be a non-singular cubic curve $C : x^3 + ax^2 + bx + c$. Set $D = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2$. Let $\Phi$ be the set of points of finite order. Let $\phi$ be the reduction map mod $p$. If $(p, 2D) = 1$ then $\phi$ is an injection into $C(F_p)$. **Nagel-Lutz:** Same as above with $P(x, y)$ as a rational point of finite order $y = 0$ or $y|d^2$.

**General Weierstrauss Form:** $E(F) : y^2 + a_1xy + a_3 = x^3 + a_2x^2 + a_4x + a_6$. If $E_q$ is an eliptic curve over a finite field of characteristic $p$, $E_q$ is said to be **supersingular** if $E_q[p] = \{\infty\}$. (1) $char(F) \neq 2, 3$, $(x, y) \mapsto (\frac{x - 3a_1^2 - 12a_2}{36}, \frac{y - 3a_1x}{216} - \frac{a_1^3 + 4a_aa_2 - 12a_3}{24})$, sends the general equation to $E_q(a, b) : y^2 = x^3 + ax + b, \Delta = -16(4a^3 + 27b^2)$. (2) $char(F) = 2, a_1 \neq 0$, $(x, y) \mapsto (a_1^2x + \frac{a_3}{a_1}, y + \frac{a_1^2a_4 - a_3^2}{a_1^2})$, sends the general equation to $E_q(a, b) : y^2 + xy = x^3 + ax + b, \Delta = b$. This is non-supersingular. (3) $char(F) = 2, a_1 = 0$, $(x, y) \mapsto (x + a_2, y)$, sends the general equation to $E_q(a, b) : y^2 + cy = x^3 + ax + b, \Delta = c^4$. This is supersingular.

(4) $char(F) = 3, a_1^2 \neq -a_2$, $(x, y) \mapsto (x + \frac{d_4}{d_2}, y + a_1 x + a_1 \frac{d_4}{d_2} + a_3)$, $d_2 = a_1^2 + a_2, d_4 = a_4 - a_1 a_3$, sends the general equation to $E_q(a, b) : y^2 = x^3 + ax + b, \Delta = -a^3 b$. This is non-supersingular. (5) $char(F) = 3, a_1^2 = -a_2$, $(x, y) \mapsto (x, y + a_1 x + a_3)$, sends the general equation to $E_q(a, b) : y^2 = x^3 + ax + b, \Delta = -a^3$. This is supersingular.

Suppose $E = E_{a,b}(K), char(K) \neq 2, 3$. Let $x_1 = \mu^2 x$ and $y_1 = \mu^3 y$ then $(x_1, y_1) \in E_{a',b'}(K)$ with $a' = \mu^4 a$ and $b' = \mu^6 b$. Define the **j-invariant**: $j(E) = 1728 \frac{4a^3}{4a^3+27b^2}$. Theorem: If $j(E_1) = j(E_2)$ then $\exists \mu \in \overline{K}, \mu \neq 0 : a_2 = \mu^4 a_1, b_2 = \mu^6 b_2$. A homomorphism $\alpha : E \rightarrow E$ is an endomorphism if $\alpha$ is a rational map. $E[n] = \{P \in E(\overline{K}) : nP = \infty\}$. **Theorem:** (1) If $char(K) \neq 2$ $E[2] = \mathbb{Z}_2 \oplus \mathbb{Z}_2$; if $char(K) = 2$ $E[2] = \mathbb{Z}_2$ or 0. (2) If $char(K) \nmid n$ or is 0, $E[n] = \mathbb{Z}_n \oplus \mathbb{Z}_n$. (3) If $char(K) = p \mid n, n = p^r n'$ then $E[n] = \mathbb{Z}_n \oplus \mathbb{Z}_{n'}$ or $E[n] = \mathbb{Z}_{n'} \oplus \mathbb{Z}_{n'}$. Proof uses division polynomials. Let $E$ be an elliptic curve over $F_q$. Then $E(F_q) = \mathbb{Z}_n$ or $\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2}$ with $n_1 \mid n_2$. Proof uses the above theorem.

**Hasse:** Let $E_q$ be an elliptic curve then $q + 1 - 2\sqrt{q} \leq \#E_q \leq q + 1 + 2\sqrt{q}$, $\#E_q = q + 1 - t$, $t$ is the Frobenius trace. Theorem: $q = p^m$, $\exists E_q : \#E_q = q + 1 - t$ iff (i) $t \neq 0(p), t^2 \leq 4q$; or, (ii) $m = 1(2)$ and either (a) $t = 0$ or (b) $t^2 = 2q, p = 2$, or (c) $t^2 = 3q, p = 3$; or, (iii) $m = 0(2)$ and either (a) $t^2 = 4q$ or (b) $t^2 = q, p \neq 1(3)$ or (c) $t = 0, p \neq 1(4)$. $E_{p^m}$ is supersingular iff $p \mid t$. $E_q = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2}$ and $n_2 \mid n_1 \mid (q - 1)$. Proof of Hasse: Let $\psi$ be the Frobenius map. $\#E_p = |ker([1] - \psi)|$. First note that $deg([1]) = 1$ (in fact, $deg([n]) = n^2$) $deg(\psi) = p$. Also note that $deg(a + b) - deg(a) - deg(b) = B(a, b)$ is bilinear. $0 \leq deg([t] + [2]\psi) = t^2 - 4p - 2tB[1, -\psi] = 4p - t^2$; so $(deg([1] - \psi) - deg([1]) - deg(\psi))^2 \leq 4p$ but the first term is $\#E(F_p)$.

**Functions on Elliptic Curves:** If $E(a, b)$ is non-singular, $E$ is irreducible and we can embed $k[x, y]/(E)$ in the field of fractions $K(E)$ with $\frac{s}{t} \cong_E \frac{u}{v}$ iff $sv - ut = 0 \pmod{E}$ and we can define a map from $K(R) \rightarrow \overline{K} \cup \{\infty\}$.

If $K = F_{p^m}, p \neq 2, 3$ and $E_K^{(1)}(a, b) \cong E_K^{(2)}(\overline{a}, \overline{b})$ iff $\exists u \in K^*$ such that $u^4 \overline{a} = a$ and $u^6 \overline{b} = b$ under the map $(x, y) \mapsto (u^2 x, u^3 y)$. $j$**-invariant:** $j(E) = 1728 \frac{4a^3}{4a^3+27b^2}$. Then $j$ is invariant under the transformation above (i.e. - two curves related by the transformation have the same $j$ value) and, conversely, two curves with the same $j$ value are related in this way (and are thus isomorphic in the elliptic curve defined over the algebraic closure). The number of equivalence classes of elliptic curves over $K$ is $2q + 6$, $2q + 2$, $2q + 4$, $2q$ according to $q = 1, 5, 7, 11 \pmod{12}$. If $K = F_{2^m}$ and $E_K(a, b) : y^2 + xy = x^3 + ax^2 + b$ then $E_K^{(1)}(a, b) \cong E_K^{(2)}(\overline{a}, \overline{b})$ iff $b = \overline{b}, Tr(a) = Tr(\overline{a})$ and if so $\exists s : \overline{a} = s^2 + s + a$ under the map $(x, y) \mapsto (x, y + sx)$. Projective coordinates: $(X_1, Y_1, Z_1) \sim (X_2, Y_2, Z_2)$, $X_1 = \lambda^c X_2$, $Y_1 = \lambda^d Y_2$, $Z_1 = \lambda Z_2$, $c, d \in \mathbb{Z}^{>0}$. Jacobian projective coordinates: $\infty = (1 : 1 : 0)$ and $-(X : Y : Z) = (X : -Y : Z)$. Standard projective coordinates: $\infty = (0 : 1 : 0)$ and $-(X : Y : Z) = (X : -Y : Z)$.

Note: the decision ECDLP problem is in $NP \cap co-NP$. Attacks (1) Exhaustive Search - to avoid, make sure $\#E_q = nh$, $n$ a large prime $> 2^{160}$, $h$, small; (2) Pohlig-Hellman/Pollard-$\rho$ use Pohlig to reduce from $n = p_1^{e_1} ... p_t^{e_t}$ to $p$, since this step is easy, want $p$ large, Pollard costs $O(\sqrt{p})$ [For Pollard, "random" function is $f(X) = X + a_j P + b_j Q \pmod{p}$.]; (3) Isomorphism attack; (4) MOV for anomalous curves - to avoid make sure $q = p^m$ and $p \nmid \#E_q$; (5) Weil-Tate pairing - to avoid make sure $n \nmid (q^k - 1), k \leq C$ and that the DLP problem for $F_{q^C}$ is intractable; (6) Weil descent - to avoid, if $q = 2^m$, make sure $m$ is prime. Index calculus attack is unlikely because the lifting required from $E_q(a, b)$ to $E_{\mathbb{Q}}(\overline{a}, \overline{b})$ is unknown and the number of points of small height in elliptic curves over $\mathbb{Q}$ is small. Let $E(K)$ be an elliptic curve $m \in \mathbb{Z}, P = (x, y)$, $\exists \psi_m(x, y), \omega_m(x, y), \theta_m(x, y) \in K[x, y]$ such that $[m]P = (\frac{\theta_m(x,y)}{(\psi_m(x,y))^2}, \frac{\omega_m(x,y)}{(\psi_m(x,y))^3})$. $E[n] = \{P \in E(\overline{K} : [n]P = \infty\}$. $\psi_1 = 0$, $\psi_1 = 2$, $\psi_2 = 2y$, $\psi_3 = 3x^4 + 6ax^2 + 12bx - a^2$, $\psi_4 = 4y(x^6 + 5x^4 + 20bx^3 - 5ax^2 - 4abx - 8b^2 - a^3)$, $\psi_{2m+1} = \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3$ $\psi_{2m} = (2y)^{-1}\psi_m(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2)$, $\theta_m = x\psi_m^2 - \psi_{m+1}\psi_{m-1}$, $\omega_m = (4y)^{-1}(\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2)$.

An **endomorphism** is a homomorphic map between and an elliptic curve and itself that is expressible as a **rational function** i.e.- If $\alpha$ is an endomorphism and $P = (x, y), \alpha(X + Y) = \alpha(X) + \alpha(Y), \alpha(x, y) = (r_1(x, y), r_2(x, y))$. Because $y^2 = x^3 + ax + b$, we may assume $\alpha(x, y) = (r_1(x), yr_2(x))$; if $r_1(x) = \frac{p(x)}{q(x)}$, the degree of endomorphism is $max(deg(p(x)), deg(q(x)))$. This endomorphism $\alpha$ is a **separable endomorphism** if $r_1'(x) \neq 0$. If $\alpha \neq 0$ is a separable endomorphism of $E$, $deg(\alpha) = \#ker(\alpha)$ otherwise $deg(\alpha) > \#ker(\alpha)$.

The endomorphism $[n]P \mapsto Q$ has degree $n^2$. If $char(K) \nmid n$ then $E[n] = \mathbb{Z}_n \oplus \mathbb{Z}_n$. If $E[n] \subseteq E(\mathbb{K})$ then $\mu_n \in K$. Given $E_q(a,b), n \geq 1$, (1) $ker(\phi_q^n - 1) = \#E_{q^n}(a,b)$ and $\phi_q^n - 1$ is separable $\#E_{q^n}(a,b) = deg(\phi_q^n - 1)$. If $\alpha$ is separable, then $deg(\alpha) = \#ker(\alpha)$. For all isogony's $\psi$, there is a dual transformation, $\phi$, such that $\phi\psi = [2]$. Let $E(F_q)$ be an elliptic curve $E(F_q) \approx \mathbb{Z}_n$ or $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2}, n_1 \mid n_2$. The Frobenius endomorphism of degree $q$ and is not separable. $\#E_q(a,b) = 1 + \sum_{x \in F_q}(1 + (\frac{x^3+ax+b}{F_q}))$. If $E_q(a,b) = \mathbb{Z}_n \oplus \mathbb{Z}_n$ then $q = n^2 + 1$ or $q = n^2 \pm n + 1$ or $q = (n \pm 1)^2$.

Given $E(K)$, $P \in E(\overline{K})$, define $D = \sum_j a_j[P_j], a_j \in \mathbb{Z}$ and $deg(D) = \sum_j a_j$. $Div^0(E)$ are the **divisors** of degree 0. If $f$ is a function on $E$, $div(f) = \sum_P ord_P(f)[P] \in div(E)$. If $D$ is a divisor of $E$ with $deg(D) = 0$, $\exists f$ on $E$: $div(f) = D$ iff $sum(D) = \infty$. $D = \sum_P n_P P$ is the divisor of an elliptic curve function on $E$ iff (1) $\sum_P n_P = 0$ and (2) $\oplus_{P \in E}[n_P]P = 0$. $f \circ n(P) = f(nP)$. If $T \in E[n], \exists T' \in E[n^2] : nT' = T$ and $g = f \circ n, div(g) = \sum_{R \in E[n]}[T' + R] - [R]$. $g(P+S)^n = g(P)^n$ so $(\frac{g(P+S)}{g(P)})^n = 1$. Define the **Weil pairing** as $e_n(S,T) = \frac{g(P+S)}{g(P)}$.

**Counting points by Baby-step Giant-step** is $(O(q^{\frac{1}{4}+\epsilon}))$. Set $N = \#E_q$ then $q + 1 - 2\sqrt{q} \leq N \leq q + 1 + 2\sqrt{q}$; if $[m]P = \infty$ then $N = m$, probably. Put $m = q + 1 - 2\sqrt{q} + k, l = \lceil \sqrt{4\sqrt{q}} \rceil, k = al + b$, then $[m]P = [c]P + [a]S + [b]P, c = q + 1 - 2\sqrt{q}, S = [l]P$ or $[c]P + [a]S = -[b]P$. Baby step computes LHS and stores it. Giant step computes RHS and does a lookup. **Schoof:** Let $\varphi$ be the Frobenius automorphism $\varphi(x,y) = (x^q, y^q)$. Schoof calculates $t \pmod{l}$ for a set of primes $l \in \mathcal{P}$ with $\prod_{l \in \mathcal{P}} l > 4\sqrt{q}$ and then reconstructs $t$ using CRT finally returning $q + 1 - t$. Here's how: (1) For $l = 2$, $t = 0 \pmod{l}$ iff $(x^3 + ax + b, x^q - x) \neq 1$. (2) if $l$ is odd, set $q_l = q \pmod{l}, |q_l| < \frac{l}{2}$; find $(x', y') = \varphi(x,y)^2 + q_l(x,y) \pmod{\psi_l(x,y)})$; for $j = 1, 2, \ldots \frac{l-1}{2}$: (i) Compute $(x_j, y_j) = j(x,y)$; (ii) if $x' - x_j^q = 0 \pmod{\psi_l}$, go to iii, if not, try next $j$, if all such $j$'s have been tried, go to (iv); (iii) Compute $y', y_j$, if $\frac{y'-y_j}{y} = 0 \pmod{\psi_l}$ then $t = j \pmod{l}$ otherwise $t = -j \pmod{l}$; (iv) Let $w^2 = q \pmod{l}$, if no such $w$ exists, $t = 0 \pmod{l}$; (v) if $(x^q - x_w, \psi_l) = 1$ then $t = 0 \pmod{l}$, otherwise, set $g = numerator(\frac{y^q - y_w}{y}, \psi_l)$, if $g \neq 1$, $t = 2w \pmod{l}$ otherwise $t = -2w \pmod{l}$.

**Lenstra Elliptic Curve Factoring Method:**

1. $(n, 6) = 1, n \neq m^r$.
2. Choose random $b, x_1, y_1$ between 1 and $n$.
3. $c = {y_1}^2 + {x_1}^3 - bx_1 \pmod{n}$.
4. $(n, 4b^3 + 27c^2) = 1$.
5. $k = lcm(1, 2, \ldots, K)$.
6. Compute $KP = (\frac{a_k}{{d_k}^2}, \frac{b_k}{{d_k}^3})$
7. $D = (d_k, n)$ If $D = 1$, go to 5 and bump $K$ or go to 2 and select new curve.

### 1.3.3 Elliptic curves and Fermat

**Regular point:** unique tangent. **Singular point:** no tangent. **Non-singular curve:** no singular points. Two curves $C, D$ are projectively equivalent if there is a projective transformation $\phi$ with $\phi(C) = D$.
Every nonsingular cubic is equivalent to a curve which in affine coordinates is $y^2 = 4x^3 - g_2 - g_3 = 0$. This is the **Weierstauss Normal Form.** Note: To prove show that every non-singular curve has an inflexion point (triple tangent). Map inflexion to $(0, 0, 1)$.

**Elliptic Functions from Trigonometry:** $S(x) = \int \frac{dx}{\sqrt{1-x^2}}$. Let $\frac{dx}{du} = c(u)$, $s(u)^2 + c(u)^2 = 1$. $s'(u) = c(u)$, $c'(u) = -S(u)$, $s(-u) = -s(u)$ and $c(-u) = c(u)$. $s(x+y) = s(x)c(y) + s(y)c(x)$ and $c(x+y) = c(x)c(y) - s(y)s(x)$. $\Omega : R \rightarrow S^1$ ($S^1$ is the 1-sphere - circle) by $u \mapsto (c(u), s(u))$ is a morphism: $\Omega(x+y) = \Omega(x) \oplus \Omega(y)$. $\Omega$ has a non-trivial kernel $K$ since $S^1$ is compact but $R$ isn't. $K = 2\pi\mathbb{Z}$. These functions are periodic, satisfy the given derivatives, parameterize $S^1$ under the indicated morphism and provide the integration property.

By analogy, set $F(k,v) = \int \frac{dz}{\sqrt{(1-z^2)(1-k^2z^2)}}$ and define $sn$ by $F(k, sn(u)) = u$. $cn(u) = \sqrt{1-sn^2(u)}$, $dn(u) = \sqrt{1-k^2sn^2(u)}$. $sn, cn, dn$ are doubly periodic with periods $\omega_1, \omega_2$. $\mathcal{L}_{\omega_1,\omega_2} = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$. Now set $\wp(z)) = \frac{1}{z^2} + \sum_{\mathcal{L}-\{0\}} \frac{1}{(z-l)^2} - \frac{1}{l^2}$ then $\wp'(z)) = -2\frac{1}{z^2} + \sum_{\mathcal{L}-\{0\}} \frac{1}{(z-l)^3}$. $\wp$ is meromorphic and doubly periodic on $\mathcal{L}$. Further, if we set $g_2 = \sum_{\mathcal{L}-0} \frac{1}{l^4}$ and $g_3 = \sum_{\mathcal{L}-0} \frac{1}{l^6}$, $\wp'(z)^3 = 4\wp(z)^2 - g_2\wp(z) - g_3$. This leads to: Let $C$ be an elliptic curve in Weierstrauss Normal Form and $\wp$ be the corresponding Weierstrauss function then $(\wp(z), \wp'(z)) \in \mathbb{C}, \forall z$ and $(\wp(u), \wp'(u)) \oplus (\wp(v), \wp'(v)) = (\wp(u+v), \wp'(u+v))$. Motivation: Want to parameterize solutions by finding $y(t)^2 = x(t)^2 + ax + b$.

Define the **Moebius transformation** $g(z) = \frac{az+b}{cz+d}$ over $\mathbb{C}$. The group of Moebius transformations is denoted by $\mathcal{M}$ and is are conformal. The **modular group** $SL_2$ is the subset of $\mathcal{M}$ with $ad - bc = 1$ with the obvious identification and is generated by $\tau \mapsto \tau + 1, \tau \mapsto -\frac{1}{\tau}$. Note fundamental region. Set $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$; these correspond to $S(z) = \frac{-1}{z}$ and $T(z) = z + 1$. Define $H = \{z : Im(z) > 0\}$ and $D = \{z : -\frac{1}{2} \leq Re(z) \leq 0, |z| = 1 \vee -\frac{1}{2} \leq Re(z) < \frac{1}{2}, |z| > 1\}$. $\mathcal{M}$ maps $H$ into itself and $D$ is a fundamental domain for $SL_2$.

**Reimann surfaces:** Glue two copies of $C$ to get $\sqrt{z}$. For $N \in \mathbb{Z}, N > 0$ define $\Gamma_0(N) \subseteq SL_2$ with $N|c$. $\Gamma_0(N)$ acts on $H$ and $H/\Gamma_0(N) \equiv X_0(N)\backslash K$ where $K$ are the cusps. $X_0(N)$ is compact and the members are the modular functions of level $N$.

**Semi-Stable:** For all primes $l > 3$, $l|Disc$ and only two of the roots are equal $\pmod{l}$. Frey curve: $C^F_{a,b} \stackrel{\text{def}}{=} y^2 = x(x-a^p)(x-b^p)$. If $b$ is even and $a = -1 \pmod{4}$. Frey curve is semi-stable.

Denote $E_{A,B,C,D}(\mathbb{Q}) \stackrel{\text{def}}{=} y^2 = Ax^3 + Bx^2 + CX + D, A, B, C, D \in \mathbb{Q}$. Define $b_p$ to be the number of solutions to $E_{A,B,C,D}(\mathbb{Q}) = 0$. $E$ is **modular** if $\exists$ eigenfunction, $f(z) = \sum_n a_n e^{2\pi i n z}$. $E/\mathbb{Q}$ is modular if $\exists f$ and eigenfunction with $a_p = p + 1 - b_p$ for all but finitely many $p$.

**Taniyama-Shimura Conjecture:** Every elliptic curve is modular. Alternate T-S: $E(A, B, C, D)$. $\exists$ modular functions $f(z), g(z)$ such that $g(z)^2 = Af(z)^3 + Bf(z)^2 + Cf(z) + D$.

Define the **conductor** $Cond_{a,b,c} = \prod_{p|abc} p$. Two elliptic curves are isomorphic iff their **j-invariants** are equal. The j-invariant of $C^F_{a,b} = 2^8 \frac{(a^{2p}+b^{2p}+a^pb^p)^3}{a^{2p}b^{2p}c^{2p}}$. If $F(\frac{az+b}{cz+d}) = (cz+d)^2F(z)$, $F$ is a modular form of weight 2.

**Proof of Fermat's Last Theorem:** Suppose it's false and that $a^p + b^p = c^p$ is a counterexample. Let $C^F_{a,b}$ be the Frey curve. $Disc(C^F_{a,b}) = a^{2p}b^{2p}c^{2p}$ so $C^F_{a,b}$ is semi-stable. Wiles proved every semi-stable elliptic curve is modular so $C^F_{a,b}$ is modular and has a cusp form of weight 2 and level $N$ where $N$ is the conductor. If $l$ is an odd prime and $l|N$, by Serre, we can obtain a new $F$ of weight 2 of level $N/l$. By induction, keep doing this until $N = 2$. The dimension of the space of cusps is equal to the genus of compact Reimann surface $X_0(N)$. But $Genus(X_0(2)) = 0$, so there is no such cusp forms of weight 2, level 2. This contradiction establishes the theorem. Incidentially, the restriction of semi-stability in Wiles Theorem has been removed.

## 1.4 Analysis, Geometry and Topology

### 1.4.1 Geometry and Topology

$[\vec{a}, \vec{b}, \vec{c}] = \vec{a} \cdot (\vec{b} \times \vec{c})$. Plane $\Pi$, perpendicular to unit vector $\vec{n}$ and containing $\vec{a}$: $\vec{x} \cdot \vec{n} = \vec{a} \cdot \vec{n} = d$. Distance from $\vec{y}$ to $\Pi$ is $|d - \vec{y} \cdot \vec{n}|$. $\vec{x} \times \vec{y} = (x_2y_3 - y_2x_3)\vec{i} + (x_3y_1 - y_3x_1)\vec{j} + (x_1y_2 - y_1x_2)\vec{k}$. Denote $[\vec{a}, \vec{b}]$ as the line from $\vec{a}$ to $\vec{b}$; $[\vec{x_0}, \vec{x_0} + \vec{a}] = \{\vec{x} : (\vec{x} - \vec{x_0}) \times \vec{a} = 0\}$. So the line that includes $\vec{x_0}$ and $\vec{x_1}$ is $\{\vec{x} : (\vec{x} - \vec{x_0}) \times (\vec{x_1} - \vec{x_0}) = 0\}$. Denote $[\vec{a}, \vec{b}, \vec{c}]$ as the plane containing $\vec{a}$, $\vec{b}$ and $\vec{c}$. $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$. Let $\theta$ be the angle (measured counterclockwise) between $[0, u]$ and $[0, v]$ then $\Delta(u, v) = u_1v_2 - u_2v_1 = |u||v|sin(\theta)$. $\Delta(u, v, w) = [u, v, w] = det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}$. Distance between $(\vec{x} - \vec{x_0}) \times \vec{a} = 0$ and $(\vec{x} - \vec{x_1}) \times \vec{b} = 0$ is

$\frac{(\vec{x_0}-\vec{x_1})\cdot(\vec{a}-\vec{b})}{||\vec{a}\times\vec{b}||}$.

**Moebius Transformations:** $\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}$. $\mathcal{M} = \{\tau_{a,b,c,d}(z) : \tau_{a,b,c,d}(z) = \frac{az+b}{cz+d}\}$. If $\tau_{a,b,c,d}(z) = \tau_{\alpha,\beta,\gamma,\delta}(z)$, $\exists\lambda \in \mathbb{C}$ such that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \lambda \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$. If $\tau \in \mathcal{M}$, $\tau : \mathbb{C}_\infty \to \mathbb{C}_\infty$ and $\tau$ is a product of maps of the following type: $z \mapsto az$, $z \mapsto z+b$ and $z \mapsto \frac{1}{z}$. For all ordered points, $< z_1, z_2, z_3 >, < w_1, w_2, w_3 >$ in $\mathbb{C}_\infty$, there is a unique $\tau \in \mathcal{M}$ such that $\tau(z_i) = w_i$. For ordered points, $< z_1, z_2, z_3, z_4 >, < w_1, w_2, w_3, w_4 >$ in $\mathbb{C}_\infty$, there is a $\tau \in \mathcal{M}$ such that $\tau(z_i) = w_i$ iff the cross-ratio of $[z_1, z_2, z_3, z_4]$ equals the cross ratio of $[w_1, w_2, w_3, w_4]$. $\Phi : GL_2(\mathbb{C}) \to \mathcal{M}$ is a surjective homomorphism given by $\Phi(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = \frac{az+b}{cz+d}$; the kernel of the homomorphism is $\lambda I, \lambda \in \mathbb{C}$. The restriction of $\Phi$ to $SL_2(\mathbb{C})$ is also a surjection with kernel $\pm I$.

$cos(a) = sin(b)sin(c)cos(A) + cos(b)cos(c)$, $\frac{sin(a)}{sin(A)} = \frac{sin(b)}{sin(B)}$.

**Circumcenter:** common intersection of the 3 perpendicular bisectors of each side of a triangle. **Incenter:** common intersection of the 3 angle bisectors of each side of a triangle. **Orthocenter:** Intersection of the altitudes. Angle bisector divides opposite side in proportion to adjacent sides.

**Pick's Theorem:** Let $B$ be a polygon which contains $n_i$ interior lattice points and $n_b$ lattice points on its boundary. $A(B) = n_i + \frac{n_b-2}{2}$. **Flex:** A non-singular point intersecting P with multiplicity 3. Every irreducible cubic in the plane has a singular point or a flex. $H = det([F_{xx}, F_{yx}, F_{zx}]^T, ...)$. Flex or singular if $H = 0$.

Projective points as one dimensional subspaces. Projective lines are 1 dimensional. $n_{p\ on\ l} n_l = n_{l\ on;\ p} n_p$.

**Fundamental of Projective Geometry:** Given 3 distinct collinear points on each of two distinct lines there is a projective transform that maps the two sets of points in the specified order.

**Cross Ratio** of four points: $r = \frac{(x_1y_3-x_3y_1)(x_2y_4-x_4y_2)}{(x_1y_4-x_4y_1)(x_2y_3-x_3y_2)}$.

**Desargues:** If $ABC$ and $A'B'C'$ are perspective from a point $X$, then $AB \cap A'B' = P$, $AC \cap A'C' = Q$, $BC \cap B'C' = R$ are collinear. **Pappus:** If $ABC$ is on $L$ and $A'B'C$' is on $L'$, then $AB' \cap A'B = P$, $AC' \cap A'C = Q$, $CB' \cap C'B = R$ are collinear.

**Ptolemy's Theorem:** Let $ABCD$ be a cyclic quadrilateral (vertices lie on a circle). Then $AB \cdot CD + AD \cdot BC = AC \cdot BD$. **Pascal:** Suppose a hexagon is inscribed in a conic section, and opposite pairs of sides are extended until they meet in 5 points. Then if 4 of those points lie on a common line, the last point will be on that line, too.

**Menelaus:** If points $X, Y, Z$ on $BC, CA, AB$ (suitably extended) are collinear $\frac{AZ}{ZB}\frac{BX}{XC}\frac{CY}{YA} = 1$. Similarly, $ABC$ with $X$ opposite $A$. $AX, BY, CZ$ are concurrent iff $\frac{AZ}{ZB}\frac{BX}{XC}\frac{CY}{YA} = 1$

**Spherical Geometry:** Let $PQR$ be a spherical triangle with subtended angles $p, q, r$ on a sphere of radius $R$. The area of $PQR$ is $R^2(p + q + r - \pi)$. Proof: Let $P', Q', R'$ be the antipodal points of $P, Q, R$ respectively and $C_P, C_Q, C_R$ be the great circles containing $PP'$, $QQ'$ and $RR'$ respectively. Let $\Delta_C$ be the common area of the three great circles in the hemisphere containing $P, Q, R$ which forms the spherical triangle. If $\Lambda(C_P, C_Q)$ is the lune formed by the intersection of the great circles, set $\Delta_1 = \Lambda(C_P, C_Q) - \Delta_C$, $\Delta_2 = \Lambda(C_P, C_R) - \Delta_C$, $\Delta_3 = \Lambda(C_R, C_Q) - \Delta_C$, and let $\Delta'_C$, $\Delta'_1$, $\Delta'_2$, and $\Delta'_3$ be the corresponding antipodal areas. $\Delta_C + \Delta_1 + \Delta_2 + \Delta_3 = \Delta'_C + \Delta'_1 + \Delta'_2 + \Delta'_3$, and $\Delta_C + \Delta_1 + \Delta_2 + \Delta_3 + \Delta'_C + \Delta'_1 + \Delta'_2 + \Delta'_3 = 4\pi R^2$ ("EQ 1"), so $\Delta_C + \Delta_1 + \Delta_2 + \Delta_3 = 2\pi R^2$. Further, $\Delta_C + \Delta_1 = 2R^2p$, $\Delta_C + \Delta_2 = 2R^2r$, and $\Delta_C + \Delta_3 = 2R^2q$ so $3\Delta_C + \Delta_1 + \Delta_2 + \Delta_3 = 2R^2(p+q+r)$, subtracting EQ 1 from this and dividing by 2 gives the desired result.

**Euler's Formula**: $V - E + F = \chi$. For a sphere, $\chi = 2$. Let $n_i$: number of vertices with valence $i$, $2e \geq 3F$, $\sum in_i = 2E$. Let $\mathcal{U}(\vec{x}) = \frac{\vec{x}}{|\vec{x}|}$. Curve length: $s(t) = \int_{t_0}^{t} |\gamma'(t)|dt$. $\vec{T}(t) = \mathcal{U}(\gamma'(t))$, $\vec{N}(t) = \mathcal{U}(\gamma''(t)) - < \gamma''(t), \vec{T}(t) > \vec{T}(t)$. Alternatively, $\vec{T}(s) = \mathcal{U}(\gamma'(s))$, $\vec{N}(s) = \mathcal{U}(\vec{T}'(s))$. $\kappa(t) = \frac{<\vec{T}'(t),\vec{N}(t)>}{\gamma'(t)}$,

$\vec{B}(t) = \vec{T}(t) \times \vec{N}(t)$, $\tau(t) = \frac{<\vec{N}'(t),\vec{B}(t)>}{|\gamma'(t)|}$. First Fundamental Form: If $E = \vec{x_u} \cdot \vec{x_u}$, $F = \vec{x_u} \cdot \vec{x_v}$ and $G = \vec{x_v} \cdot \vec{x_v}$ then $I(du, dv) = Edu^2 + 2Fdudv + Gdv^2$. If $\vec{N} = \frac{\vec{x_u} \times \vec{x_v}}{|\vec{x_u} \times \vec{x_v}|}$ then $L = -\vec{x_u} \cdot \vec{N_u}$, $M = -\frac{1}{2}(\vec{x_u} \cdot \vec{N_v} + \vec{x_v} \cdot \vec{N_u})$, $N = -\vec{x_v} \cdot \vec{N_v}$ and $II(du, dv) = Ldu^2 + 2Mdudv + Ndv^2$. $\kappa_n = \frac{II}{I}$. $\kappa$ is a principal curvature iff $det \begin{pmatrix} L - \kappa E & M - \kappa F \\ M - \kappa F & N - \kappa G \end{pmatrix} = 0$. **Gaussian curvature:** If $k_1$ and $k_2$ are the maximum and minimum values of the curvature at a point on a surface, the Gaussian curvature is $K = k_1 k_2$; $\chi = 2 - 2g$ is the Euler characteristic, where $g$ is the genus. The **genus** of a connected, orientable surface is an integer representing the maximum number of cuttings along closed simple curves without rendering the resultant manifold disconnected and it is equal to the number of handles on it. **Gauss-Bonnet:** If X is a compact, hypersurface in $\mathbb{R}^{k+1}$, then $\int_X K = Vol(S^k)\frac{\chi(X)}{2}$.

Let $G(u, v)$ be a homogeneous polynomial and $(u_0, v_0) \in \mathbb{P}^1_K$, $\exists k \geq 0$ and $H(u, v)$ with $H(u_0, v_0) \neq 0$ : $G(u, v) = (v_0 u - u_0 v)^k H(u, v)$. Any line in $\mathbb{P}^2_k$ can be parameterized by $(x, y, z) = (a_0 u + b_0 v, a_1 u + b_1 v, a_2 u + b_2 v)$. $L$ intersects $C$ to order $n$ at $P = (x_0 : y_0 : z_0)$ if $\overline{C}(u, v) = (v_0 u - u_0 v)^n H(u, v)$ in the foregoing theorem; denote this as $ord_{L,P}(C) = n$, $ord_{L,P}(C) = \infty$ if $\overline{C}$ is identically 0. If $L_1, L_2$ are lines, $ord_{L_1,P}(P) = 1$ or $\infty$. If $C$ is a curve defined by $C(x, y, z) = 0$, $C$ is non singular at $P$ if $(C_x, C_y, C_z) \neq 0$ in which case the tangent line is $C_x X + C_y Y + C_z Z = 0$. If $C$ is non-singular at $P$ there is a line in $\mathbb{P}^2_K$ that intersect $C$ to order at least 2.

**Eight Point Theorem:** Suppose $C$ is a curve in $\mathbb{P}^2_K$ defined by homogeneous cubic polynomial $C(x, y, z) = 0$. Let $l_1, l_2, l_3$ and $m_1, m_2, m_3$ be lines in $\mathbb{P}^2_K$ with $l_i \neq m_j, \forall i, j$ and $P_{ij} = l_i \cap m_j$. Suppose further that $C$ is not singular at $P_{ij}, \forall i, j \neq 3, 3$. Then $P_{33} \in C$. This is proved in a series of lemmas. Lemma 1: Let $P_{i1} = (u_i : v_i)$ and $m_j : a_j x + b_j y + c_j z = 0$, $\overline{m_j}(u_i, v_i) = 0$ and $\overline{m_j}$ vanishes only at $P_{ij}$. $\overline{m_1}(u, v)\overline{m_2}(u, v)\overline{m_3}(u, v)$ is a homogeneous cubic polynomial. Lemma 2: If $R(u, v), S(u, v)$ are homogeneous of degree 3 and is not identically 0 and they both vanish at $(u_i : v_i)$ then $\exists \alpha \in K, \alpha \neq 0$: $R = \alpha S$. Lemma 3: $\overline{C} = \alpha \overline{m_1}(u, v)\overline{m_2}(u, v)\overline{m_3}(u, v)$ and $\overline{C} = \alpha \overline{l_1}(u, v)\overline{l_2}(u, v)\overline{l_3}(u, v)$. Lemma 4: $l_i \mid (C - \alpha m_1(u, v)m_2(u, v)m_3(u, v))$, $m_j \mid (C - \beta l_1(u, v)l_2(u, v)l_3(u, v))$ and if $D = C - \alpha m_1(u, v)m_2(u, v)m_3(u, v) - \beta l_1(u, v)l_2(u, v)l_3(u, v)$, then $l_i m_j \mid D$. Lemma 5: $D = l_1 m_1 l(u, v)$ and $l(P_{22}) = l(P_{23}) = l(P_{32}) = 0$, so $D = 0$. To conclude the proof of the eight point theorem, observe, since $D = 0$, $C = \alpha m_1(u, v)m_2(u, v)m_3(u, v) + \beta l_1(u, v)l_2(u, v)l_3(u, v)$ and $l_3(P_{33}) = m_3(P_{33}) = 0$ thus $C(P_{33}) = 0$.

The eight point theorem proves associativety of elliptic curve addition. Let $P, Q, R$ be points on $C$ and consider $l_1 = \overline{P, Q}$, $l_2 = \overline{\infty, Q + R}$, $l_3 = \overline{R, P + Q}$, $m_1 = \overline{Q, R}$, $m_2 = \overline{\infty, P + Q}$, $m_3 = \overline{P, R + Q}$. $l_1 \cap m_1 = Q$, $l_1 \cap m_2 = -(P + Q)$, $l_1 \cap m_3 = P$, $l_2 \cap m_1 = -(Q + R)$, $l_2 \cap m_2 = \infty$, $l_2 \cap m_3 = Q + R$, $l_3 \cap m_1 = R$, $l_3 \cap m_2 = (P + Q)$, $l_3 \cap m_3 = X$. $X$ is $-((P + Q) + R)$ (from the definition of $l_3$) and $-(P + (Q + R))$ (from the definition of $m_3$) by the definition of addition. Now apply the eight point theorem to get the result.

The eight point theorem also proves Pascal's Theorem: Let $ABCDEF$ be a hexagon inscribed in a conic section whose equation is $Q(x, y, z) = 0$. If $X = \overline{AB} \cap \overline{DE}$, $Y = \overline{BC} \cap \overline{EF}$, $Z = \overline{CD} \cap \overline{FA}$, then $X, Y, Z$ are collinear. Proof: Put $l_1 = \overline{EF}$, $l_2 = \overline{AB}$, $l_3 = \overline{CD}$, $m_1 = \overline{BC}$, $m_2 = \overline{DE}$, $m_3 = \overline{FA}$, $C(x, y, z) = Q(x, y, z)l(x, y, z)$ and apply the theorem. It also proves Pappus's Theorem: Let $l, m$ be two distinct lines $A, B, C$ on $l$ and $A', B', C'$ on $m$ none of which are on $l \cap m$. If $X = \overline{AB'} \cap \overline{A'B}$, $Y = \overline{BC'} \cap \overline{B'C}$, $Z = \overline{CA'} \cap \overline{C'A}$, then $X, Y, Z$ are collinear. Proof: Use Pascal with hexagon $AB'CA'BC'$.

### 1.4.2 Complex Analysis

If $w = f(x + iy) = u(x, y) + iv(x, y)$ is analytic in a region $\mathcal{R}$ then $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$ and $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$.

**Cauchy:** If $f(z)$ is analytic in a region $\mathcal{R}$ and its boundary $\mathcal{C}$ then $\int_{\mathcal{C}} f(z)dz = 0$. Morrera: If $f(z)$ is continuous in a simply connected region $\mathcal{R}$ and $\int_{\mathcal{C}} f(z)dz = 0$ around every simple closed curve $\mathcal{C}$ in $\mathcal{R}$, then $f(z)$ is analytic in $\mathcal{R}$.

If $f(z)$ is **analytic** inside and on a circle $\mathcal{C}$ of radius $r$ and center at $z = a$ then $|f^{(n)}(a)| \leq \frac{Mn!}{r^n}$ where $|f(z)| \leq M$ on $\mathcal{C}$ in $\mathcal{R}$. If an analytic function is bounded in the plane it is constant.

If $f(z)$ is analytic inside and on a circle $\mathcal{C}$ of radius $r$ and center at $z = a$ then $f(z) = \frac{1}{2\pi}\int_0^{2\pi} f(a + re^{i\theta})d\theta$.

**Meromorphic:** Analytic everywhere in the plane except at a finite number of poles. **Entire:** Analytic everywhere in the complex claim. If $f(z)$ is analytic inside and on a closed curve $\mathcal{C}$ except at a finite number of pole then $\frac{1}{2\pi i}\int_{\mathcal{C}} \frac{f'(z)}{f(z)} = N - P$ where $N$ and $P$ are, respectively, the number of zeros and poles of $f(z)$ inside $\mathcal{C}$. **Rouche's theorem:** If $f(z), g(z)$ are analytic in and on a simple closed curve $C$ and $|f(z)| > |g(z)|$ on $C$ then $f(z)$ and $f(z) + g(z)$ have the same number of zeros in $C$.

**Cauchy Integral Formula:** If $f(z)$ is analytic inside and on a closed curve $\mathcal{C}$ and $a$ is any point inside $\mathcal{C}$ then $f^{(n)}(a) = \frac{1}{2\pi i}\int_{\mathcal{C}} \frac{f(z)}{(z-a)^{n+1}}$. **Laurent:** If $f(z)$ is analytic inside an annular region $\mathcal{A} = \{a \leq z - z_0 \leq b\}$ then $f(z) = \sum_{n=-\infty}^{\infty} c_n(z-z_0)^n$. In that case, $c_{-1} = Lim_{z \rightarrow a} \frac{1}{(k-1)!}\frac{d^{k-1}}{dz^{k-1}}((z-a)^k f(z))$. **Residue Theorem:** $\int_{\mathcal{C}} f(z)dz = 2\pi i(a_{-1} + b_{-1} + ...)$.

### 1.4.3 Real Analysis and Manifolds

**Taylor:** $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(c)}{n!}(x-a)^n$ for some $c : a < c < x$. Proof: Set $F(t) = f(t) + \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k$ and let $E_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$. Note $F(x) - F(a) = E_n(x)$ and $F'(t) = \frac{f^{(n+1)}(t)}{n!}(x-t)^n$. Put $G(t) = (x-t)^n$ and $H(t) = G(t)[F(x) - F(a)] - F(t)[G[x] - G(a)]$. $H(a) = H(x)$ so $\exists c : a < c < x$ with $H'(c) = 0$. So $E_n(x) = F(x) - F(a) = \frac{F'(c)}{G'(c)}[G(x) - G(a)]$. Substituting gives the desired result.

$X$ separates if $\exists A, B, A \neq X, \emptyset, \exists B \neq X, \emptyset$ both open with $A \cup B = X$ and $A \cap B = \emptyset$. $X$ is connected iff there is no separation. Suppose $f : X \rightarrow Y$ is **continuous.** If $X$ is compact so is $f(X)$. If $X$ is connected, so is $f(X)$. If $X$ compact, **connected** set in $\mathbb{R}$ then $X = [a, b]$.

Let $\Phi : Q \rightarrow P$ be a **homotopy** of $\varphi_0$ into $\varphi_1$ as closed curves and let $y \notin \Phi(Q)$. Then the **winding number** $W(\varphi_r, y)$ is constant for $0 \leq r \leq 1$. Let $\varphi$ be a closed curve $\varphi : [a, b] \rightarrow P$ and suppose $y_0, y_1$ can be joined by a curve which does not intersect $\varphi$, then $W(\varphi, y_0) = W(\varphi, y_1)$. Let $f : D \rightarrow P$ be a mapping of the disk onto the plane and let $C = \partial D$ and let $y \notin f(C)$; if the winding number of $f|C$ about $y$ is not zero, then $y \in f(D)$ such that $f(x) = y$. Let $f : D \rightarrow P$ be a mapping of a disk onto a plane, $P$ and $C = \partial D$ that fixes all of $C$ then $D \subseteq f(D)$. No mapping of a disk onto its boundary fixes each point of the boundary. If $f$ is a mapping of a disk onto itself, it has a fixed point.

$\int_{-\pi}^{\pi} cos(mx)cos(nx)dx = \int_{-\pi}^{\pi} sin(mx)sin(nx)dx = \delta_{mn}\pi$. **Bernoulli:** $\phi_n'(x) = \phi_{n-1}(x), \phi_0(x) = 1, \int_0^1 \phi_n(x)dx = 1$. $\Gamma(x) = \int_0^{\infty} u^{x-1}e^{-u}du$. $\int\int_R(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial x}) = \int_C pdx + qdy$.

**Fixed Point Theorem:** Let $E$ be a complete metric space and $f : E \rightarrow E$. Suppose $\exists k < 1 : \forall p, q \in E, ||f(p) - f(q)|| \leq k||p - q||$. Then there is a unique $P \in E : f(P) = P$. Proof: Let $p_{n+1} = f(p_n)$. $||f(p_{n+1}) - f(p_n)|| \leq k||p_n - p_{n-1}|| \leq k^n||p_1 - p_0||$. This is a Cauchy sequence and converges. Set $p = lim_{n\rightarrow\infty}p_n$, $f(p) = p$. Uniqueness: if $q$ is another such point: $||f(p) - f(q)|| = ||p - q|| \leq k||p - q||$ so $||p - q|| = 0$.

**Simple Implicit Function Theorem:** Let $f$ be a real valued continuous function on an open set $E \subset \mathbb{R}^2, (a, b) \in E$ with continuous partial $\frac{\partial f}{\partial y}(a, b) \neq 0$. There are open sets $U, V$ with $a \in U, b \in V$ and a continuous function $\varphi : U \rightarrow V$ such that $f(x, \varphi(x)) = 0, x \in U$. Proof: Define $F(x, y) = y - f(x, y)(\frac{\partial f}{\partial y})^{-1}$. $F(a, b) = b$, $\frac{\partial F}{\partial y}(a, b) = 0$ and $F(x, y) = y$ iff $f(x, y) = 0$. Pick $r$ small enough so that in the ball $B_r(a, b) : |\frac{\partial F}{\partial y}| < \frac{1}{2}$. Choose $k : 0 < k < r$ then choose $h : 0 < h < \sqrt{r^2 - k^2}$ such that $|F(x, b) - b| < \frac{k}{2}$ when $|x - a| < h$. Put $U = (a - h, a + h), V = (b - k, b + k)$. Fix $x \in U$ and $|y - b| \leq k$ and suppose $||(x, y) - (a, b)||^2 < h^2 + k^2 < r^2$ and $||(x, y') - (a, b)||^2 < h^2 + k^2 < r^2$. $\exists y'' : |F(x, y) - F(x, y')| \leq \frac{\partial F}{\partial y}(x, y'')|y - y'| \leq \frac{1}{2}|y - y'|$ and $|F(x, y) - b| \leq |F(x, y) - F(x, b)| + |F(x, b) - b| < k$. Apply Fixed Point Theorem to get $\overline{y} = f(x, \overline{y})$. This is unique. Define $\varphi(x) = \overline{y}$. A simple argument shows $\varphi$ is continuous.

**Simple Inverse Function:** Let $g$ be a real valued function on an open set $E \subset \mathbb{R}$ and suppose $g'$ exists and is continuous in $E$ and $g'(b) \neq 0$. There are open sets $U, V \subset \mathbb{R}$ with $b \in V : g_{|V}$ is 1-1 and $g^{-1} : U \rightarrow V$ is differentiable. Proof: Put $f(x, y) = x - g(y)$ and apply the Implicit function theorem.

**Existence of solution to ordinary differential equation:** Let $f$ be a continuous real valued function in an open set $E \subset \mathbb{R}^2$ containing $(a,b)$ and suppose $\exists M : |f(x,y) - f(x,z)| < M|y-z|, (x,y),(x,z) \in E$ then $\exists h > 0$ and $\varphi : (a-h, a+h) \to (b-M, b+M) : \varphi'(x) = f(x, \varphi(x))$ on $(a-h, a+h)$ and $\varphi(a) = b$. Proof: This is equivalent to $\varphi'(x) = \int_a^x f(t, \varphi(t))dt + b$. Suppose $\psi$ is a function and define $F : \psi \mapsto \int_a^x f(t, \psi(t))dt + b$. $F$ maps the complete metric space of functions on a closed interval of $E$ itself. A fixed point in this metric space would satisfy the theorem; we show such a fixed point exists. Choose $N > |f(a,b)|, \exists r : ||(x,y) - (a,b)|| < r \to |f(x,y)| < N$. Choose $h > 0 : h < \frac{r}{2N}, h < \frac{1}{2}, hM < 1$ and consider the complete metric space of continuous functions on $[a-h, a+h]$ denoted by $\mathcal{C}([a-h, a+h])$; define $R = \{(x,y) \in E : |a-x| \leq h, |y-b| \leq Nh\}$ and $B = \{\psi : [a-h, a+h] \to [b-Nh, b+Nh]\}$, finally, Let $B_{Nh}(b)$ be the ball in $\mathcal{C}([a-h, a+h])$ of functions within $Nh$ of the constant function $b$. For $\psi, \omega \in B_{Nh}(b)$ note that $|\psi(x) - b| < Nh$ and $f(t, \psi(t)) < N$ so $|F\psi(x) - b| < Nh$. For $\psi, \omega \in B_{Nh}(b) : |F\psi(x) - F\omega(x)| \leq hM||\psi - \omega||$. This satisfies the conditions of the fixed point theorem and the fixed point satisfies the conclusion of the theorem.

**Implicit Function Theorem:** Let $a \in E^m \subset \mathbb{R}^m$ and $b \in E^n \subset \mathbb{R}^n$ with $(a,b) \subset E^{m+n}$, and open set. Suppose $f_1(a,b) = \ldots = f_n(a,b) = 0$ and $\frac{\partial f_i}{\partial y_j}$ exist and are continuous in $E^{m+n}$ and $det(\frac{\partial f_i}{\partial y_j}(a,b)) \neq 0$ then $\exists U^{open} \subset E^m, a \in U, V^{open} \subset E^n, b \in V, \varphi : U \to V$ such that $f_i(x, \varphi(x)) = 0$ for $i = 1, 2, \ldots, n$. Proof: Define $x = \vec{x} = (x_1, \ldots, x_m)$, $y = \vec{y} = (y_1, \ldots, y_n)$ and $F = \vec{F} = (F_1(\vec{x}, \vec{y}), \ldots, F_n(\vec{x}, \vec{y}))$. Define $F_i(x,y) = y_i - \sum_j c_{ij} f_j(x,y)$ with each partial continuous. (1) The $F_i$ are continuously differentiable; (2) $F_i(a,b) = b$; (3) $\frac{\partial F_i}{\partial y_j}(a,b) = 0$; (4) $f_i(x,y) = 0$ iff $F_i(x,y) = y_i$. For 3 to hold $(c_{ij})$ must be the inverse of the Jacobian. For 4 to hold, the determinant of the Jacobian must be $\neq 0$. Choose $r > 0$ such that for $(x,y) \in B_r(a,b) \subset E^{m+n}$, $|\frac{\partial F_i}{\partial y_j}| < \frac{1}{2n^2}$ and $det(\frac{\partial F_i}{\partial y_j}) \neq 0$. Choose $k : 0 < k < r$ and choose $h$ so that $0 < h < \sqrt{r^2 - k^2}$ and $||F(x,b) - b|| < \frac{k}{2}$ if $||x - a|| < h$. Fix $x \in U$ with $||(x,y) - (a,b)|| < r$. If $y' \in E^n, ||y' - b|| \leq k, \exists y'' : F(x,y) - F(x,y') = (y - y') \cdot (\frac{\partial F_1(x,y'')}{\partial y_1}, \ldots, \frac{\partial F_n(x,y'')}{\partial y_n}) \leq \frac{1}{2n^2}(|y_1 - y_1'| + \ldots + |y_n - y_n'|) \leq \frac{1}{2n}||y - y'||$. So $||F(x,y) - F(x,y')|| < k$ and the fixed point theorem applies.

**Extended Inverse Function Theorem:** $f_i(x,y) = x_i - g_i(y), a = g(b)$. Same deal.

**Inverse Function Theorem:** Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and $|det(f'(a)| \neq 0$. $\exists V^{open}, W^{open}, f^{-1}, a \in V, f(a) \in W$ with $f^{-1} : W \to V$ and $f^{-1}(f(x) = x$. Further $f'^{-1}(y) = \frac{1}{f'(f^{-1}(y))}$. Notes: Let $\lambda = D(f(a))$. May assume $\lambda = I$. Can show $|x_1 - x_2| \leq |f(x_1) - f(x_2)|$.

**Implicit Function Theorem:** If $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable in an open set containing $(a,b), f(a,b) = 0$ with $M = (D_{n+j}(f^i(a)))$ with $1 \leq i,j \leq m$. If $det(M) \neq 0, \exists A^{open} \subseteq \mathbb{R}^n$ and $B^{open} \subseteq \mathbb{R}^m, a \in A, b \in B : \forall x \in A$ there is a unique $g(x) \in B, f(x, g(x)) = 0$. Further, $g$ is differentiable. Notes: Look at $F(x,y) = (x, f(x,y))$ and apply Inverse Function Theorem.

**Partitions of unity:** $A^{open} \subseteq \mathbb{R}^n$ and $\mathcal{O}$ and open cover of $A$. $\exists \Phi \in \mathbb{C}^\infty$ such that $\forall \varphi \in \Phi$: (1) $0 \leq \varphi(x) \leq 1$ and $\forall x \in A$, (2) $\forall x, \varphi(x) = 0$ for all but finitely many $\varphi \in \Phi$, (3) $\sum_{\varphi \in \Phi} \varphi(x) = 1$. (4) $\forall \varphi \in \Phi, \exists U^{open} \in \mathcal{O} : \phi(x) = 0$ for $x \notin \overline{U}$ where $\overline{U}$ is some closed subset of $U$.

Direct proof of inverse function theorem. Suppose $f$ is a $\mathcal{C}'$ mapping $f : E \to \mathbb{R}^n$, $a \in E^{open} \subseteq \mathbb{R}^n$ with $f'(a)$ invertible and $f(a) = b$, then (a) $\exists U^{open}, V^{open} \subseteq \mathbb{R}^n : a \in U, b \in V$ such that $f$ is 1-1 on $U$; $f(U) = V$. (b) If $g = f^{-1}$ then $g \in \mathcal{C}'(V)$. Proof of a: Put $f'(a) = A$ and choose $\lambda : 2\lambda||A^{-1}|| = 1$, set $U = B_\lambda(a) \subseteq E$: $||f'(x) - A|| < \lambda, \forall x \in U$. Set $\varphi_y(x) = x + A^{-1}(y - f(x)), \forall y \in \mathbb{R}^n$. $||\varphi_y'(x)|| = ||A^{-1}(A - f'(x))|| < \frac{1}{2}$. $||\varphi_y(x_1) - \varphi_y(x_2)|| < \frac{1}{2}, \forall x_1, x_2 \in U$ [Equation 1] by the mean value theorem. $\varphi_y$ is a contraction map so it has a unique fixed point $x : y = f(x)$. Put $V = f(U)$ and suppose $y_0 \in V$, there is a $x_0 \in U : y_0 = f(x_0)$. Pick $r > 0 : \overline{B_r(x_0)} \subseteq U$. Fix $y : |y - y_0| < \lambda r$. For $x \in \overline{B_r(x_0)}$, $|\varphi(x_0) - x_0| \leq |\varphi(x) - \varphi(x_0)| + |\varphi(x_0) - x_0| < \frac{1}{2}|x - x_0| + \frac{r}{2} \leq r$ so $\varphi(x) \in \overline{B_r(x_0)}$ and again $\varphi_y$ is a contraction map. Its fixed point $x$ satisfies $f(x) = y, y \in \overline{B_r(x_0)} \subseteq f(U) = V$, so $V$ is open. Proof of b: Pick $y \in V$, $y + k \in V, \exists x, x + h \in U : y = f(x), y + k = f(x + h)$. Now $\varphi(x + h) - \varphi(x) = h + A^{-1}(f(x+h) - f(x)) = h - A^{-1}(f(x+h) - f(x)) = h - A^{-1}k \leq \frac{1}{2}h$ by equation 1, so $||A^{-1}k|| \geq \frac{||h|||}{2}$ and $||h|| \leq 2||A^{-1}k|| = \lambda^{-1}||k||$. $f'(x)$ has an inverse $T$ and $g(y + k) - g(y) - Tk =$

$h - Tk = -T[f(x+h) - f(x) - f'(x)h]$ and $\frac{||g(y+k)-g(y)-Tk||}{||k||} = \frac{||T||}{\lambda}\frac{||f(x+h)-f(x)-f'(x)h||}{||h||}$. Now $h \to 0$ as $k \to 0$. Since the right hand side goes to 0, the left hand side goes to 0 and we get $g'(y) = T$.

**Fubini's Theorem:** $\int\int_{I^2} f(x,y)dydx = \int_0^1(\int_0^1 f(x,y)dy)dx$. In a simply connected region of the plane, $S$ for $a \leq x \leq b$ bounded by $b_1(x) \leq y \leq b_2(x)$, $\int\int_S f(x,y)dydx = \int_a^b(\int_{b_1(x)}^{b_2(x)} f(x,y)dy)dx$. **Change of variables:** Let $A \subseteq \mathbb{R}^n$ be an open set, $g : A \to R$ continuously differentiable and $det(g'(x)) \neq 0, \forall x \in A$. If $f : g(A) \to R$ is integrable then $\int_{g(A)} f = \int_A f \circ g|\det(g')|$.

Let $\mathcal{T}^k(V) = \{T : V \to \mathbb{R}\}, V \subseteq \mathbb{R}^n$ where $\forall i$: $T(v_1, ..., v_{i-1}, u+w, v_{i+1}, ..., v_k) = T(v_1, ..., v_{i-1}, u, v_{i+1}, ..., v_k) + T(v_1, ..., v_{i-1}, w, v_{i+1}, ..., v_k)$ and $T(v_1, ..., v_{i-1}, au, v_{i+1}, ..., v_k) = aT(v_1, ..., v_{i-1}, u, v_{i+1}, ..., v_k)$. $\mathcal{T}^n(V)$ are called the $n-$tensors $V$. If $f : V \to W$ with $V, W \subseteq \mathbb{R}^n$ then $f^* : \mathcal{T}^n(W) \to \mathcal{T}^n(V)$ by $f^*(T(v_1, ..., v_n)) = T(f(v_1), ..., f(v_n))$. If $T \in \mathcal{T}^k, S \in \mathcal{T}^s$ define $T \otimes S = T(x_1)S(x_2)$. $\mathcal{T}^1(V)$ is just the dual space $V^*$. If $e_1, ..., e_n$ is a basis for $V$ and $\varphi_j \in V^*$ such that $\varphi_j(e_i) = \delta_{ij}$ then the set of all $k - fold$ tensor products $\varphi_{i_1} \otimes \varphi_{i_2} \otimes ... \otimes \varphi_{i_k}$ is a basis $\mathcal{T}^k(V)$ which thus has dimension $n^k$.

**Alternating forms:** $\Lambda^k(V) = \{T \in \mathcal{T}^k(V)\}$ such that $T(...v...w...) = -T(...w...v...)$. $\forall T \in \mathcal{T}^n(V), Alt(T) = \frac{1}{k!}\sum_\sigma sgn(\sigma)T(v_{\sigma(1)}, ..., v_{\sigma(n))} \in \Lambda^k(V)$. If $\omega \in \Lambda^k(V)$, $Alt(\omega) = \omega$. If $\omega, \eta \in \Lambda^k, \Lambda^l$, $\omega \wedge \eta = \frac{(l+k)!}{k!l!}Alt(\omega \otimes \eta)$. $\wedge$ is multilinear and $\omega \wedge \eta = (-1)^{kl}\eta \wedge \omega$; $f^*(\omega \wedge \eta = f^*(\omega) \wedge f^*(\eta)$. $(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta) = \frac{(k+l+m)}{k!l!m!}Alt(\omega \otimes \eta \otimes \theta)$. If $\omega = \sum w_{i_1,...i_k}dx^{i_1} \wedge ... \wedge x^{i_k}$ then $d\omega = \sum dw_{i_1...i_k} \wedge dx^{i_1} \wedge ... \wedge x^{i_k}$. $dim(\phi_{i_1} \wedge ... \wedge \phi_{i_k}) = \binom{n}{k}$. orientation: $[e_1, ..., e_n]$. **Volume elements:** $w_i = \sum_j a_{ij}v_j$ then $\omega(w_1, ..., w_n) = det(a_{ij})\omega(v_1, ..., v_n)$ for $\omega \in \Lambda^k$.

**Forms:** Let $p, v \in \mathbb{R}^n$, define the tangent space of $\mathbb{R}^n$ at $p$, ${\mathbb{R}^n}_p$, as the $(p, v)$ with $(p, v) + (p, w) = (p, v + w)$ and $(p, av) = a(p, v)$.

**Vector field:** $F(p) = F^1(p)(e_1)p + ... + F^n(p)(e_n)p$ with the usual rules $(F + G)(p) = F(p) + G(p)$ $(f \cdot g)(p) = f(p) \cdot g(p)$. $\nabla = \sum D_i \cdot e_i$.

$\omega(p) \in \Lambda^k({\mathbb{R}^n}_p)$: If $\varphi_i(p)$ is the dual basis for $(e_1)_p, (e_2)_p, ..., (e_n)_p$ then $\omega(p) = \sum \omega_{i_1,...i_k}\varphi^{i_1} \wedge ... \wedge \varphi^{i_k}$ is a differential form and $df(p)(v_p) = Df(p)(v)$. $df = \sum_i^n D_i f dx^i$.

If $f : \mathbb{R}^n \to \mathbb{R}^m$, $f_* : {\mathbb{R}^n}_p \to {\mathbb{R}^m}_p$ by $f_*(v_p) = (Df(p)(v))_{f(p)}$. Thus $f_* : \Lambda^k({\mathbb{R}^m}_{f(p)}) \to \Lambda^k({\mathbb{R}^n}_p)$. So if $\omega$ is a $k-$form on $\mathbb{R}^m$, $f^*\omega(p) = f^*(\omega(p))$ is a $k-$form on $\mathbb{R}^n$. $f^*(dx^i) = \sum_j D_j f^i \cdot dx^j$, $f^*(\omega_1 + \omega_2) = f^*(\omega_1) + f^*(\omega_2)$, $f^*(g \cdot \omega) = g \circ ff^*\omega$ and $f^*(\omega + \eta) = f^*\omega + f^*\eta$. If $f : \mathbb{R}^n \to R$, $Df(p) \in \Lambda^1(\mathbb{R}^n)$. $df(p)(v_p) = Df(p)(v)$. $f_*(v_p) = (Df(p)(v))_{f(p)}$. $f : \mathbb{R}^n \to \mathbb{R}^m$, $f_* : {\mathbb{R}^n}_p \to {\mathbb{R}^m}_{f(p)}$. $f_* : \Lambda^k({\mathbb{R}^m}_{f(p)}) \to \Lambda^k({\mathbb{R}^n}_p)$. $f^*(dx^i) = \sum_{j=1}^n D_j f^i dx^j$. $f^*(g \circ \omega) = g \circ f \circ f^*\omega$.

$d^2\omega = 0$, **closed form:** $d\omega = 0$, **exact form:** $\exists \eta : d\eta = \omega$. **Poincare:** If $A^{open} \subseteq \mathbb{R}^n$ is a star shaped region then every closed form in $A$ is exact. $\partial I^n = \sum_{i=1}^n \sum_{\alpha=0,1}(-1)^{i+\alpha}{I^n}_{(i,\alpha)}$ where ${I^n}_{(i,\alpha)} = I^n(x^1, ..., , x^{i-1}, \alpha, x^{i+1}, ...x^n)$. Note that $\partial^2 I^n = 0$. If $A^{open} \subseteq \mathbb{R}^n$ and $g : A \to \mathbb{R}^p$ is differentiable and $g'(x)$ has rank $p$ whenever $g(x) = 0$ then $g^{-1}(0)$ is an $n-p$ dimensional manifold. Diffeomorphism, $k-$dimensional manifold. An $n$-dimensional differentiable manifold is called **orientable** if it has a differential form $\omega$ of degree $n$ which is nonzero at every point on the manifold.

**Stokes:** If $M$ is a compact oriented $k-$dimensional manifold with boundary and $\omega$ is a $k - 1$ form on $M$ then $\int_c d\omega = \int_{\partial c}\omega$.

**Classical Integral Theorems:** Let $\vec{x} = (x_1, x_2, \ldots, x_n)$. **Lagrange:** Maximize $F(\vec{x})$ subject to $\phi_1(\vec{x}) = 0, \phi_2(\vec{x}) = 0, \ldots, \phi_m(\vec{x}) = 0$; form $G(\vec{x}) = F(\vec{x}) + \lambda_1\phi_1(\vec{x}) + \lambda_2\phi_2(\vec{x}) + \ldots + \lambda_m\phi_m(\vec{x})$ and solve $\frac{\partial G}{\partial x_j} = 0$. Suppose $\mathcal{R} \subseteq \mathbb{R}^n$ $\mathcal{R}' \subseteq \mathbb{R}^n$ and $f : \mathcal{R} \to \mathcal{R}'$ is continuously differentiable then $\int_\mathcal{R} F(\vec{x})d\vec{x} = \int_{\mathcal{R}'} F(f(\vec{u}))|J_f(\vec{u})|d\vec{u}$ where $J_f(\vec{u}) = |det(f')|$. **Green:** If $C$ surrounds $\mathcal{R}$, a simply connected region of the plane then $\int_C Pdx + Qdy = \int_\mathcal{R}(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y})dxdy$. **Gauss:** If $\mathcal{S}$ is a surface enclosing a convex region $\mathcal{V}$ and $\vec{F}$ is continuously differentiable then $\int_\mathcal{V} \nabla \cdot \vec{F}(\vec{x})d\vec{x} = \int_{\vec{\mathcal{S}}} \vec{F}(\vec{x}) \cdot dS$. **Stokes:** If $\mathcal{S}$ with boundary $\mathcal{C}$ and $\vec{F}$ is continuously differentiable then $\int_\mathcal{S} \nabla \times \vec{F}(\vec{x}) \cdot dS = \int_{\vec{\mathcal{C}}} \vec{F}(\vec{x}) \cdot d\vec{l}$. **Fourier:** $F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(u)e^{iux}du$ and

$f(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(x) e^{-iux} dx$.

**Calculus of variations:** Let $I = \int_{x_1}^{x_2} L(x, y, y')dx$ and $f(x)$ be the function that minimizes $I$ ($\delta I = 0$), then $-\frac{d}{dx}\frac{\partial L}{\partial y'} + \frac{\partial L}{\partial y} = 0$.

## 1.5 Probability

### 1.5.1 General Probability

$\mu_X = E(X)$, ${\sigma_X}^2 = Var[X] = E[(X - E[X])^2]$, **Covariance:** $\mu_{XY} = E((X - \mu_X)(Y - \mu_Y))$. **Correlation:** $\rho(X,Y) = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma(X)\sigma(Y)}$. **Moment generating function:** $G(e^t) = \sum_{k \geq 0} \Pr[X = k]e^{tk} = E[e^{tX}]$. Moment Generating Function for Poisson distribution ($f(x) = e^{-\lambda x}$) is $\phi(t) = E(e^{tx}) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \frac{\lambda}{\lambda - t}$. $E(X^2) = \frac{d}{dt}\phi(t) = \frac{2}{\lambda^2}$. $Var(X) = \frac{1}{\lambda^2}$.

**Stirling's approximation:** $n! \approx \sqrt{2\pi n}(\frac{n}{e})^n$. Proof: $M_n = ln(n!) = \sum_{i=1}^n ln(i)$. $\int_1^n ln(x) < M_n < \int_1^{n+1} ln(x)$. So $nln(n) - n < M_n < (n+1)ln(n+1) - n$. Set $d_n = ln(n) - (n + \frac{1}{2})ln(n) - n$. $d_n - d_{n+1} = (n + \frac{1}{2}ln(\frac{n+1}{n}) - 1$. Writing $\frac{n+1}{n} = \frac{1+\frac{1}{2n+1}}{1-\frac{1}{2n+1}}$ and expanding the log, and comparing to the geometric series in $2n+1$, we find $d_n$ converges to, say, $C$. So, $n! \approx e^C n^{n+\frac{1}{2}}e^{-n}$. To find $e^C$ use Wallis' formula: $lim_{n\to\infty}\frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}} = \sqrt{\pi}$. To get this, show $\int_0^{\frac{\pi}{2}} sin^n(x) = \frac{n-1}{n}\int_0^{\frac{\pi}{2}} sin^{n-2}(x)$.

**Bayes:** $P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum P(A|B_j)P(B_j)}$. **Normal Distribution:** $N(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $Z = \frac{(X-np)}{\sqrt{npq}}$. **Binomial Distribution:** $B(N,n,p) = \binom{N}{n}p^n(1-p)^{N-n}$, $E(B) = Np, \sigma^2 = Np(1-p)$. **Poisson Distribution:** $P(x) = e^{-\lambda}\frac{\lambda^x}{x!}$, $\mu = \lambda, \sigma^2 = \lambda$, probability of count in time $\Delta t$ is $\lambda\Delta t$.

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-(\frac{(x-\mu_1)^2}{\sigma_1^2}+(2\rho)\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2})/(2\sqrt{1-\rho^2})}$$

$\rho$ is the cross correlation between $x$ and $y$.

**Central Limit Theorem:** If $X_i$ are independent, identically distributed random variables and $S_n = X_1 + \ldots + X_n$, then $\lim_{n\to\infty} P(a \leq \frac{(S_n - np)}{\sigma\sqrt{n}} \leq b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-(u^2/2)}$. Proof: $E(S_n) = n\mu$, $\sigma^2 = Var(S_n) = n\sigma_{X_i}$. Define $S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}$. So $E[e^{tS_n^*}] = E[e^{\frac{t(X_1-\mu)}{\sigma\sqrt{n}}}e^{\frac{t(X_2-\mu)}{\sigma\sqrt{n}}}\ldots e^{\frac{t(X_n-\mu)}{\sigma\sqrt{n}}}] = E[e^{\frac{t(X_1-\mu)}{\sigma\sqrt{n}}}]^n$. Expanding the exponential in the taylor series, we get $E[e^{tS_n^*}] = E[1 + \frac{t(X-\mu)}{1!\sqrt{n}\sigma} + \frac{(t(X-\mu))^2}{2!(\sqrt{n}\sigma)^2} + \ldots] = e^{\frac{-t^2}{2}}$. This is the same moment generating function as the normal distribution, so were done.

$\chi^2 = \frac{(Y_2 - np_2)^2}{(np_2)} + ... + \frac{(Y_{12} - np_{12})^2}{(np_{12})}$, $P(\chi^2 \leq x) = \frac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})}\int_0^x u^{\frac{\nu}{2}-1}e^{-\frac{u}{2}}du$.

**Markov:** Let $Y$ be a random variable assuming only non-negative values, and with expected value $E[Y]$ convergent. Then for any $t > 0$, $\Pr[Y \geq t] \leq \frac{E[Y]}{t}$.

**Chebyshev:** Let $Y$ be a random variable with expected value $\mu = E[Y]$ and variance, $Var(Y)$. Then for any $t > 0$, $\Pr[|Y - \mu| \geq t] \leq \frac{Var(Y)}{t^2}$.

**Chernoff:** Let $T_1, T_2, \ldots, T_N$ be mutually independent Bernoulli variables $T = \sum_i^N T_i$. Then $\forall c \geq 0$, $Pr(T \geq cE(T)) \leq e^{\alpha E(T)}$ where $\alpha = ln(c) + \frac{1}{c} - 1$.

**Wald:** Let Q be a random variable that takes on only non-negative integer values such that $E(Q) < \infty$. Let $R_1, R_2, \ldots$ be a sequence of random variables with the same distribution and let $T = R_1 + R_2 + \ldots + R_Q$. Suppose $R_k$ is independent of the event that it is included in the sum, that is $\forall k \geq 1$, $R_k$ is independent of an indicator variable for the event $Q \geq k$ then $E(T) = E(Q)E(R_1)$.

**Occupancy:** Let $X_i$ be an indicator for a ball falling into $i$. $E(X_i) = 1$. Let $Z_i$ be the probability that the bin is empty. $E(Z_i) = \frac{n}{e}$. Let $p_m(r,n)$ be the probability of finding $r$ balls in $n$ cells with exactly $m$ empty cells. $p_m(r,n) = \binom{n}{m}(1 - \frac{m}{n})^r p_0(r, n-m)$. Further, $p_0(r,n) = \sum_{i=0}^n (-1)^i\binom{n}{i}(1 - \frac{i}{n})^r$.

**Lovasz Local Lemma:** Let $G = (V,E)$ be a dependency graph for events $e_1, e_2, ..., e_n$ in a probability space. Suppose $\exists x_i \in [0,1]$ for $1 \leq i \leq n$, such that $Pr[e_i] \leq x_i\Pi_{(i,j)\in E}(1 - x_j)$. Then $Pr[\cap\overline{e_i}] \geq \Pi_i^n(1 - x_i)$.

If $\{p_i\}$ and $\{q_i\}$ are probability distributions and $G(q_1, q_2, \ldots, q_n) = -\sum p_i ln(q_i)$. Then $G$ is minimum when $p_i = q_i$.

### 1.5.2 Statistical Inference and Hidden Markov Models

Let $Y = Pred(L)$, $\sigma^2(Y,L) = E((Y-L)^2)$. Value of predictor: $W(Y,L) = \frac{\sigma^2(Y,L) - E(L-Y)^2}{\sigma^2(Y,L)}$. $0 = W(E(L), L) \leq W(Y,L) \leq W(L,L) = 1$. $E((X-t)^2)$ is minimized $t = E(Y)$. Let $cov(X,Y) = E(XY) = E(X)E(Y)$. Best linear predictor: $Y = aX + b$, $a = \frac{cov(X,Y)}{cov(X,X)}$ (and solve for $b$). Worth of best predictor (using mean square error) is $\rho(X,Y)^2 = \frac{cov(X,Y)^2}{cov(X,X)cov(Y,Y)}$. Posterior models. $P(|Y-\mu| \geq t) \leq \frac{var(Y)}{t^2}$.

Let $S = \{1, 2, 3, \ldots, n\}$ be the $n$ possible states of a hidden markov process with $T$ transitions and $T+1$ outputs. Notation: Denote $\vec{X}^{(L)} = \prod_{i=0}^{L} X$ and $\vec{x} \in \vec{X}^{(L)}$ with $\vec{x} = (x_0, x_1, \ldots, x_L)$; we denote $\vec{x}_i = x_i$. Suppose the output vector of the process is $\vec{O} \in (\mathbb{Z}_m)^{(T)}$. Finally, suppose the following distributions are given: initial state distribution - $\pi(i), i \in \mathbb{Z}_m$; output distribution - $q_{ij} = q(j|i) = Pr(O_t = j|\vec{S}_t = i), \forall t$; state transition distribution: $p_{ij} = P(j|i) = Pr(\vec{S}_t = j|\vec{S}_{t-1}) = i], \forall t$.

- Problem 1 Given $O = O_0, O_1, O_2, ..., O_T$, $\lambda = (P, q, \pi)$, how do we compute $Pr(O|\lambda)$ efficiently?
- Problem 2 Given $\vec{O} = O_0, O_1, O_2, ..., O_T$ and $\lambda$, how do we choose an $\vec{q}$ which is optimal?
- Problem 3 How do we adjust the model parameters $\lambda = (P, q, \pi)$, to optimize $Pr(\vec{O}|\lambda)$, given the observed sequence: $\vec{O}$?

Problem 1: Assuming the foregoing, the probability of the output $\vec{O}$ is:

$$Pr[\vec{O}|\lambda] = \sum_{\vec{s} \in \vec{S}^{(T)}} \pi(\vec{s}_0) q(O_0|\vec{s}_0) \prod_{i=1}^{T} P(\vec{s}_i|\vec{s}_{i-1}) \prod_{i=1}^{T} q(O_i|\vec{s}_i)$$

The following recursion greatly improves the calculation cost. Let $\alpha_0(i) = \pi(i)q(O_0|i), \forall i$ and $\alpha_t(i) = (\sum_{j=1}^{k} \alpha_{t-1}(j) P(S_t = i|S_{t-1} = j))q(O_t|i), \forall i$. This is called the "forward recursion". Then $\alpha_t(i) = \sum_{\vec{s} \in \vec{S}^t, \vec{s}_t = i} \pi(\vec{s}_0) q(O_0|\vec{s}_0) \prod_{j=1}^{t} P(\vec{s}_j|\vec{s}_{j-1}) \prod_{j=1}^{t} q(O_j|\vec{s}_j)$, the probability of the observation of the sequence up to time $t$ given $\vec{s}_t = i$. $Pr(\vec{O}|\lambda) = \sum_{i=1}^{n} \alpha_T(i)$; computing $\{\alpha_T(i)\}$ takes $O(n^2(T+1))$ rather than $O(2(T+1)n^{T+1})$. This solves problem 1.

Problem 2: Slightly abusing the notation from above define $\beta_t(i) = Pr(O_{t+1}, \ldots, O_T|S_t = i, \lambda)$. The "backwards recursion" is: $\beta_T(i) = 1, \forall i$, $\beta_t(i) = \sum_{j=1}^{n} P(S_t = i|S_{t+1} = j)\beta_{t+1}(j)q(O_{t+1}|j)$. Now define $\gamma_t(j) = P(s_t = 1|\vec{O}, \lambda)$ so $\gamma_t(j) = \frac{\alpha_t(i)\beta_t(j)}{P(\vec{O}|\lambda)}$. The most likely state at time $t$ is the one that maximizes $\gamma_t(i)$.

Problem 3: Define $\gamma_t(i,j) = P(S_t = i, S_{t+1} = j|\vec{O}, \lambda)$ so $\gamma_t(i,j) = \frac{\alpha_t(i) P(S_t = j|S_{t+1} = i) q(O_{t+1}|j) \beta_{t+1}(j)}{P(\vec{O}|\lambda)}$ and $\gamma_t(i) = \sum_{j=1}^{n} \gamma_t(i,j)$. $\gamma_t(i,j)$ is the probability of being in state $i$ at $t$ and transitioning to state $j$.
Now, suppose the model, $\lambda = (\pi, P, q)$, is unknown, the MLE of the model, given observations $\vec{O}$ is determined by:

- $0 = \frac{\partial}{\partial \pi(i)}[Pr(\vec{O} = (O_0, \ldots, O_T)) - \lambda_1(\sum_{k=0}^{m-1} \pi(k) - 1)]$.
- $0 = \frac{\partial}{\partial P(j|i)}[Pr(\vec{O} = (O_0, \ldots, O_T)) - \lambda_2(\sum_{k=0}^{m-1} P(k|i) - 1)]$.
- $0 = \frac{\partial}{\partial q(j|i)}[Pr(\vec{O} = (O_0, \ldots, O_T)) - \lambda_3(\sum_{k=0}^{m-1} q(k|i) - 1)]$.

Solving gives the following **re-estimation formulas:**

- $\hat{\pi}(i) = \gamma_0(i) = \frac{\alpha_0(i)\beta_0(i)}{\sum_{k=1}^{n} \alpha_0(k)\beta_0(k)}$, $\sum \pi(i) = 1$.
- $\hat{P}(j|i) = \frac{\sum_{t=0}^{T-1} \gamma_t(i,j)}{\sum_{t=0}^{T-1} \gamma_t(i)} = \frac{\sum_{t=0}^{T-1} \alpha_t(i) q(O_{t+1}|j) P(j|i) \beta_t(j)}{\sum_{t=0}^{T} \alpha_t(i)\beta_t(i)}$, $\sum_j P(j|i) = 1$.
- $\hat{q}(j|i) = \frac{\sum_{t \in \{0,1,\ldots,T-1\}, O_t = j} \gamma_t(i)}{\sum_{t=0}^{T-1} \gamma_t(i)} = \frac{\sum_{t=0, O_t = j}^{T-1} \alpha_t(i)\beta_t(i)}{\sum_{t=1}^{T} \alpha_t(i)\beta_t(i)}$, $\sum_j q(j|i) = 1$.

Baum showed that if $Q(\lambda, \overline{\lambda}) = \sum_{s \in S} P_{\lambda}(O,s) log(P_{\overline{\lambda}}(O,s)$ and $Q(\lambda, \overline{\lambda}) > Q(\lambda, \lambda)$ then $P_{\overline{\lambda}}(O,s) > P_{\lambda}(O,s)$. Optimizing $Q$ instead of $P$ gives the Baum EM algorithm. Note that optimizing using dynamic programming may give a different result: $\delta_0(i) = \pi(i)q(i|O_0)$, $\delta_t(i) = max_{j \in \{1,\ldots,n\}}(\delta_{t-1}(j)p_{ji}q_{iO_t})$ since it optimizes the overall path. You can deal with underflow by taking logs or (in the HMM case) scaling in a way that maintains the re-estimation result.

**EM as Gaussian mixture problem:** $p(\vec{x}) = \sum_{k=1}^{K} N(\vec{x}|\vec{\mu_k}, \vec{\Sigma_k})$, let $\vec{z}$ be a $K$ dimensional random variable from the sample space all of whose components are 0 but a single one which is 1 (i.e.- $z_k = 1$) under the Gaussian model $(\pi_k, \mu_k, \Sigma_k)$. $p(\vec{x}|z_k = 1) = N(\vec{x}|\vec{\mu_k}, \vec{\Sigma_k})$, $p(z_k = 1) = \pi_k$ and $p(\vec{x}) = p(\vec{x}|\vec{z})p(\vec{z})$. $\pi_k$ is the prior estimate of $z_k = 1$ and $\gamma(z_k)$ is the posterior estimate. $\gamma(z_k) = p(z_k = 1|\vec{x}) = \frac{p(z_k=1)p(\vec{x}|z_k=1)}{\sum_j p(z_j=1)p(\vec{x}|z_j=1)}$. For mixing, let $< \vec{x_1}, \ldots, \vec{x_N} >$ be a sample. The log likelihood is $p(\vec{x}|\vec{\pi}, \vec{\mu}, \vec{\Sigma}) = \sum_{n=1}^{N} ln(\sum_{k=1}^{K} \pi_k N(\vec{x_n}|\mu_k, \Sigma_k))$ and EM maximizes this. Maximizing equations come from taking derivatives with respect to $\mu_k$ and setting them to 0 — $0 = -\sum_{n=1}^{N} \frac{\pi_k N(\vec{x_n}|\mu_k, \Sigma_k}{\sum_j \pi_j N(\vec{x_j}|\mu_j, \Sigma_j)} \cdot \Sigma_k(\vec{x_n} - \mu_k)$. The term in the denominator is $\gamma(z_{nk})$, $N_k = \sum_{n=1}^{N} \gamma(z_{n,k})$ and $\mu_k = \frac{1}{N_k} \sum_{n=1}^{N} \gamma(z_{n,k})$. Taking the derivatives with respect to $\Sigma_k$ give the remaining equations (Note: $\mu_k = \frac{N_k}{N}$). An alternative (Bayesian) view is to regard $\vec{z}$ as latent, $\Theta$ as the model parameters and $ln(p(\vec{X}|\Theta)) = ln(\sum_z p(\vec{X}|\vec{Z}, \Theta))$. We use this to estimate the likelihood from $\Theta^{old}$ for general $\Theta$: $\mathcal{Q}(\Theta, \Theta^{old}) = \sum_z p(Z|X, \Theta^{old}) ln(p(X, Z|\Theta))$; the "M" step corresponds to finding $\Theta^{new} = arg\ max_{\Theta}(\mathcal{Q}(\Theta, \Theta^{old})$.

**Principal Component Analysis:** Suppose $x_1, x_2, \ldots, x_N \in \mathbb{R}^D$ and we project this space onto $< u_1, u_2, \ldots, u_M >$ where $u_k \in \mathbb{R}^D$ and $u_i u_i^T = 1$. For example, for $M = 1$, the variance of the projection is $\frac{1}{N} \sum_{n=1}^{N} (u_1^T x_n - u_1^T \overline{x}) = u_1^T S u_1$ where $\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ and $S$ is the co-variance matrix. Finding the first principal component requires us to to maximize $u_1^T S u_1$ subject to $u_1^t u_1 = 1$. Using Lagrange multipliers, this is equivalent to maximizing $f(u_1) = u_1^T S u_1 + \lambda_1(1 - u_1^T u_1)$. Taking derivative, we get $S(u_1) = \lambda_1 u_1$ with $\lambda_1$ the largest eigenvalue of $S$. Can also find $\lambda_1$ with EM. For general $M$, $u_i^T u_j = \delta_{ij}$, $\vec{x_n} = \sum_{i=1}^{D} \alpha_{ni} u_i$, $\alpha_{nj} = (x_n^T u_j)$, $x_n = \sum_{i=1}^{D} (x_i^T u_i \cdot u_i)$ and we want to minimize $J = \frac{1}{N} \sum_{n=1}^{N} ||x_n - \overline{x}||^2$ which reduces to an eigenvalue problem.

### 1.5.3 Information and Coding Theory

**Shannon conditions for entropy:** (a) continuous in probability, (b) monotonically increasing in number of messages, additive with respect to refinement: $H(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}) = H(\frac{1}{2}, \frac{1}{2}) + \frac{1}{2} H(\frac{1}{2}, \frac{1}{2})$. Number of bits of information obtained in observing event that occurs with probability $p$ is $lg(p)$. $H(P) = \sum -p_i lg(p_i)$, $lg(|X|) \geq H(X) \geq 0$. $I(X, Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y)$. $H(X, Y) \leq H(X) + H(Y)$. $H(U|V) = 0$ iff $U = g(V)$.

$D(p||q) = \sum_x p(x) lg(\frac{p(x)}{q(x)}) \geq 0$. Markov chain denoted by $X \rightarrow Y \rightarrow Z$. If $X \rightarrow Y \rightarrow Z$ then $I(X;Y) \leq I(X;Z)$. Let $T(X)$ be any statistic and $F =< f_\theta(x) >$ and $X$ a sample from $F$ then $I(\theta; T(X)) \leq I(\theta; X)$. $T$ is a **sufficient statistic** if equality holds. $T(X)$ is a minimal sufficient statistic relative to $F$ if it is a statistic of every other sufficient statistic $U(X)$. $\theta \rightarrow T(X) \rightarrow U(X) \rightarrow X$. A stochastic process $X =< X_1, X_2, \ldots >$ is **stationary** if the joint distribution of any subsequence is invariant with respect to time shifts. **Entropy** of a stochastic process is $H(X) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \ldots, X_n)$. For a stationary Markov chain, the entropy rate is given by $H(X) = H(X_2|X_1)$. If $X$ is a stationary markov chain then so is the process $< Y_i = \phi(X_i) >$ and $H(Y_n|Y_{n-1}, \ldots, Y_1, X_1) \leq H(Y) \leq H(Y_n|Y_{n-1}, \ldots, Y_1)$ equality holds by taking the limit across the inequalities.

$H_\delta(X) = lg(min\{|T| : T \subseteq A_X, Pr(x \in T) \geq (1 - \delta)\}$. **Asymptotic Equipartition:** $n$, independent identically distributed random variables $X_i$, if $X^n = (X_1, X_2, \ldots, X_n)$ is almost certain to belong to $B \subseteq A_X^n$ having about $2^{NH}$ members, each with probability "close" to $2^{-NH}$. This is equivalent to **Shannon's Source coding Theorem:** The $n$ r.v.'s can be encoded by $NH$ bits with negligible information loss. To show this, show for any $\delta$ there's an $n$ such that $H_\delta(X^{(n)}) \approx NH$. Hint: Define $Y = \frac{1}{n} lg(\frac{1}{p(x)})$. Let $T_{n,\beta} = \{y \in A_X^n : [\frac{1}{n} lg(\frac{1}{p(x)}) - H]^2 < \beta^2\}$.

**Channel Capacity:** $C = max_{P(x)}(H(I|J) - H(I))$. For a DMC, BSC with error rate $p$, this implies $C_{BSC}(p) = 1 + plg(p) + qlg(q)$. So for BSC $R = 1 - H(P)$.

Detect $t$ errors $d(C) \geq t+1$. Correct $t$ errors $d(C) \geq t+1$. Perfect code: $M(\sum_k^t \binom{n}{k}(q-1)^k) = q^n$.

**Shannon Source Coding:** If a memoryless source has entropy $H$ then any uniquely decipherable code over an alphabet $\Sigma$ with $D$ symbols must have length $\geq \frac{H}{lg(D)}$. Further, $\exists$ a uniquely decipherable code with average length $\leq 1 + \frac{H}{lg(D)}$.

**Shannon's Theorem Channel Coding:** If $0 \leq R \leq 1 + plg(p) + qlg(q)$, $M_n = 2^{\lceil Rn \rceil}$, then $P^*(M_n, n, p) \rightarrow 0$ as $n \rightarrow \infty$. Notation: Each codeword has $n$ bits. Let $P_i$ be the probability of making an error in decoding if $x_i$ is transmitted. Then $P_C = \frac{1}{M}\sum_i P_i$ is the probability of making a decoding error if a randomly chosen codeword is transmitted and every codeword is equiprobable. $P^*(M_n, n, p) = min_C(P_C)$, with $BlockLength(C) = n$, $R = \frac{lg(|C|)}{n}$ and $M_n = 2^{\lfloor Rn \rfloor}$. Proof: Define the following terms: $f(u,v) = 0$, if $d(u,x) > \rho$ and $f(u,v) = 1$, if $d(u,x) \leq \rho$, $g_i(y) = 1 - f(y, x_i) + \sum_{i \neq j} f(y, x_i)$. Then $P_i = \sum_y P(y|x_i) g_i(y) = \sum_y P(y|x_i)[1 - f(y, x_i)] + \sum_y \sum_{i \neq j} P(y|x_i) f(y, x_i)$. So, $P_C = min_C[\frac{1}{M}\sum_i(\sum_y \sum_y P(y|x_i)[1 - f(y, x_i)] + \sum_y \sum_{i \neq j} P(y|x_i) f(y, x_i))]$. Now, taking expectations over all eligible $C$ and using the fact that at least one particular $C$ must have $P_C \leq$ the expected value of $P_C$ over all $C$, we get $P_C \leq [\frac{1}{M}\sum_i \sum_y E(P(y|x_i)[1 - f(y, x_i)]) + \sum_y \sum_{i \neq j} E(P(y|x_i))E(f(y, x_i))]$. Now, let $N_e$ be the number of received bits in error in a string of length $n$, then $E(N_e) = np$ and $Var(N_e) = \sqrt{npq}$. Set $b = \sqrt{\frac{npq}{\frac{\epsilon}{2}}}$ then $P(n_e > np + b) \leq \frac{\epsilon}{2}$ by Chebychev. If $B_\rho(x)$ is the set of words of distance $\leq \rho$. So, we get $P_C \leq \frac{\epsilon}{2} + M^{-1}\sum_i \sum_y \sum_{i \neq j} E(P(y|x_i))E(f(y, x_i)) \leq \frac{\epsilon}{2} + (M-1)2^{-n}|B_\rho)|$. Now $\rho = pn$ and $B_\rho(x) = \sum_{i \leq \rho}\binom{n}{i}$. But $1 = [\lambda + (1-\lambda)]^n = \sum_{k=0}^{pn}\binom{n}{k} \leq \lambda^{pn}(1-\lambda)^{n(1-p)}\sum_{k=0}^{pn}\binom{n}{k}$. So, $2^{-nH(p)} \geq \sum_{k=0}^{pn}\binom{n}{k}$. Putting this back in the equation for $P_C$ we get $P_C \leq \frac{\epsilon}{2} + (M-1)2^{-n(1+H(p))} \leq 2^{n(R-1-H(p))}$ which goes to 0 if $R < 1 + H(p)$.

$(n, M, d)$ **codes:** $M$ is number of codewords, $d$ is minimum distance, $n$ is dimension. An $[n, k, d]$ **linear code** is an $k-$subspace of an $n-$ space over $F$ with minimum distance $d$. **Standard form for generator** is $G = (I_k|A)$ with $k$ message bits, $n$ codeword bits. Codeword $c = mG$ and $d = min_{u \neq 0, u \in C}\{wt(u)\}$. **Parity check matrix**, $H$, of a code is the generator of its dual code. $C^\perp = \{x : (x, y))) = 0, \forall y \in C\}$. Note that $GH = 0$. If $C$ is a code, $C^\perp$ is a code (the **dual code**). $H = (-A^T, I_{n-k})$, $GH^T = 0$. Consider a table with the codewords forming the first row, subsequent rows add error $e$ until all $2^n$ blocks are in the table. Each row is a coset and the element of minimum weight in each row is called the coset leader. To decode received word $r = c + e$: (1) compute syndrome $s(r) = rH^T$, (2) find coset leader with $s(r)$ and locate the codeword, $c_0$ in that column, (3) decode as $r - c_0$.

Define $V(n, r) = \sum_{j=1}^r \binom{n}{j}$. **Hamming Bound:** $|C| \leq \frac{2^n}{V(n,e)}$. **Sphere Packing Bound:** If $d = 2e + 1$, $A_q(n, d)\sum_{k=0}^e \binom{n}{k}(q-1)^k \leq q^n$. **GSV Bound:** $A(n, d) \geq \frac{2^n}{V(n, d-1)}$, where $A(n, d)$ is the largest code with minimum distance $d$.

A **Hamming code** is a $[n, k, d]$ linear code with $n = 2^m - 1$, $k = 2^m - 1 - m$ and $d = 3$. To decode, if $r = c + e$ is received (1) calculate $s(r) = rH^T$, (2) find $j$ which is the column of $H$ with syndrome $s(r)$, correct position $j$. The $[7, 4]$ code has encoding matrix

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

with check equations $y_1 + y_3 + y_5 + y_6 = 0$, $y_2 + y_3 + y_6 + y_7 = 0$, $y_4 + y_5 + y_6 + y_7 = 0$. For Hamming, $n = 2^m - 1$, $m$ parity checks identify error position. Motivation for BCH is to use another $m$ parity checks which identify $f(j) = j^3$ positions. Rows of Hadamard matrix $HH^T = nI$ forms a $(n, 2n, \frac{n}{2})$ code. Let $A_i$ be the number of codewords of weight $i$ for a code $C$, then $A(z) = \sum_i A_i z^i$ is the weight enumerator.

A **cyclic code,** $C$, has the property that $(c_1, c_2, \ldots, c_n) \in C \rightarrow (c_n, c_1, \ldots, c_{n-1}) \in C$. Denoting $U_n(x) = x^n - 1$ we have the following theorem: $C$ is a cyclic code of length $n$ iff its generator $g(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1} \mid U_n(x)$ where codewords $c(x)$ have the form $m(x)g(x)$. Further, if $U_n(x) = h(x)g(x)$, $c(x) \in C$ iff $h(x)c(x) = 0 \pmod{U_n(x)}$. Example: $g(x) = 1 + x^2 + x^3$ generates $(7, 4)$ code. $g(x)m(x) = c(x)$, $a = (1010), a(x) = 1 + x^2$; $g(x)a(x) = c(x) = x^5 + x^4 + x^3 + 1$, $c = (1001110)$. In shift register implementations, bits come out of 0-degree term, recurrence is shifted into high-degree. Cyclic codes ideals in $\mathbb{Z}_2/(x^n - 1)$. Codewords are multiples of the generator polynomial $g(x)$. Let $\alpha$ be a primitive element of $GF(2^m)$. $[n = 2^m - 1, k = n - m, d = 3]$ hamming code has parity check $H = (1, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{2^m-2})$. If $g(x)$ is the generator for $\alpha$, generator matrix is

$$C = \begin{pmatrix} g(x) & 0 & 0 \\ 0 & xg(x) & 0 \\ 0 & 0 & x^2 g(x) \\ \ldots & & \end{pmatrix}$$

For BCH with $[n = 2^m - 1, k = n - 2m, d \geq 5]$, $g(x) = M^{(1)}(x)M^{(3)}(x)$ where $M^{(3)}(x)$, is the minimum polynomial for $\alpha^3$.

**BCH codes:** If $g(x)|x^n - 1$, the ideal generated by $g(x)$ is a cyclic code. If $g(x)$ factors into linear factors in $GF(2^n)$ with roots $A = \{\alpha_1, \ldots, \alpha_r\}$, the set $C$ defined by $f(x) \in C$ iff $f(\alpha) = 0, \forall \alpha \in A$ is a cyclic code. For BCH, pick $g(x) = m_1(x)m_2(x) \ldots m_r(x)$ of degree $d$ with each factor irreducible. Let $n - d$ message bits be the high order coefficients $C_I(x)$ of an $n - 1$ degree polynomial whose remaining terms are $C_R(x)$ with $C_I(x) = g(x)q(x) + C_R(x)$. For a 2-ECC, pick $g(x) = m_1(x)m_2(x)$ with $m_1(x)$ the irreducible monic polynomial for a primitive $n$th root of 1, $\alpha$ and $m_2(x)$ the irreducible monic polynomial for $\alpha^3$. Alternatively, suppose $g(x)$ is a cyclic code and $\alpha$ is a primitive $n$th root of $g(x)$ and $g(\alpha^l) = g(\alpha^{l+1}) = \ldots = g(\alpha^{l+\delta}) = 0$ then $d \geq \delta + 2$ and the resulting BCH code has weight $d$. Decoding BCH for $r = c + e$: (1) compute $(s_1, s_2) = rH^T$, (2) if $s_1 = 0$, no error, (3) if $s_1 \neq 0$ put $\frac{s_2}{s_1} = \alpha^{j-1}$, error is in position $j$ (of $p \neq 2, e_j = \frac{s_1}{\alpha^{(j-1)(k+1)}}$), (3) $c = r - e$.

**Reed-Solomon** code is BCH code over $F_q$ with $n = q - 1$. Let $\alpha$ be a primitive root of 1 and choose $d : 1 \leq d < n$ with $g(x) = (x - \alpha)(x - \alpha^2) \ldots (x - \alpha^{d-1})$. The BCH code generated by $g(x)$ is a Reed Solomon code (an MDS code too).

Building codes and **Reed Muller:** If $C_1 : (n, M_1, d_1)$ and $C_2 : (n, M_2, d_2)$, $C_3 = C_1 * C_2$ denotes the code where codewords in $C_3$ are $(u, u + v), u \in C_1, v \in C_2$. It is a $(2n, M_1M_2, min(2d_1, d_2))$ code. $RM(0, m) = \{0, 1\}$, $RM(r+1, m+1) = RM(r+1, m) * R(r, m)$. $R(r, m)$ is a $(n_r, M_r, d_r)$ code, with $n_r = 2^m$, $d_r = 2^{m-r}$ and $M_r = 2^a$, $a = 1 + \binom{m}{1} + \ldots + \binom{m}{r}$. $R(r, m)$ has parameters $[n = 2^m, k = 1 + \binom{m}{1} + \ldots + \binom{m}{r}, d = 2^{m-r}]$, it consists of boolean functions whose polynomials are of degree $\leq m$. $RM(r, m)^\perp = RM(m - r - 1, m)$.

$R = \frac{1-H_2(p)}{1-H_2(p_e)}$ (4,7) code. $U = \frac{H(K)}{D}$, $2^{RN}$ messages $2^{rN}$ meaningful ones, $2^{H(K)}$ keys. $2^{H(K)} - 1$ keys have probability, q, of spurious decryption $R - r = D$. $F$= number of false ones. $F = (2^{H(K)} - 1)q = 2^{(H(K)-D)N}$. The correct key maps cipher into meaningful class always. False keys map cipher into meaningful/meaningless randomly. After how many message is the expected number of spurious keys which map all the samples into meaningful less than 1? Shannon: $M_C$: total message length, $M$: meaningful part, $p$: probability of error. $pM_C = k$, $2^{M_C - M} \geq \binom{M_C}{k}$.

**Hadamard Code:** Let $h_{ij} = (-1)^{a_0b_0 + \ldots + a_4b_4}$, where $a$ and $b$ index the rows and columns respectively. This gives a $32 \times 32$ entry matrix, $H$. Let generators be $G = [H| - H]^T$. For each of the $0 \leq i < 2^6$ possible messages, send the row corresponding to $i$. To decode, for the 32 bit received word, $r$, compute $d_i = r \cdot R_i$, where $R_i$ is the 32 bit row $i$. If there are no errors, the correct row will have $d_i = 32$ and all other rows will have $d_i = 0$. If one error, $d_i = 30$, etc.

**Golay Code** $\mathcal{G}_{24}$ is a $[24, 12, 8]$ linear code. $G = [I_{12}|C_0|N] = [I|B]$ where $C_0 = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0)^T$ and $N$ is formed by circulating $(1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 0)$ 11 times and appending an row of 11 1's. The first row of $N$ corresponds to the quadratic residues $\pmod{11}$. Note that $wt(r_1 + r_2) = wt(r_1) + wt(r_2) - 2[r_1 \cdot r_2]$, all codewords have weight divisible by 4 and $d(C) = 8$. $\mathcal{G}_{24} = \mathcal{G}_{24}^\perp$. To decode Golay, write $G = [I_{12}|B]$ and $B^T = (b_1, b_2, \ldots, b_{12})$ with $b_i$ a column vector. Suppose $r = c + e$ is received and $wt(e) \leq 3$. Put $s = rG^T$ and compute $sB$, $s + c_i^T$, $1 \leq i \leq 24$ and $sB + b_j^T$, $1 \leq j \leq 12$. If $wt(sB) \leq 3$, there is a non-zero entry in the $k$-th position of $sB$ if the $k + 12$-th position of $e$ is non-zero. If $wt(s) \leq 3$ a non-zero entry in $s$ at position $k$ corresponds to a non-zero entry in position $k$ of $e$. If $wt(s + c_j^T) \leq 2$, for some $j$, $13 \leq j \leq 24$ then $e_j = 1$ and non-zero entries of $s + e_j^T$ are in the same positions as non-zero entries of $e$. If $wt(sB + b_j^T) \leq 2$, for some $j$, $1 \leq j \leq 12$ then $e_j = 1$ and non-zero entries of $sB + b_j^T$ at position $k$ correspond to non-zero entries of $e_{k+12}$.

**Leech:** Let $R(C)$ be the row space of $C$ over $GF(2)$. Define the $\Gamma$ to be the collection of $(v_1, v_2, \ldots, v_{24}) = v \in \mathbb{Z}^{24}$ such that (1) $\sum_{i=1}^{24} v_i = 4m$, (2) $v_i = m \pmod 4$, if $c_i = 0$, (3) $v_i = m + 2 \pmod 4$ if $c_i = 1$.

**Rogers Bound:** $RB(n) = \frac{\sqrt{(n+1)}(n!)^2 \pi^{\frac{n}{2}}}{2^{\frac{3n}{2}} \Gamma(\frac{n}{2}+1)} f_n(n)$, $F_{n+1}(\alpha) = \frac{2}{\pi} \int_{\frac{arcsec(n)}{2}}^{\alpha} F_{n-1}(\beta) d\theta$, $sec(2\beta) = sec(2\theta) - 2$, $F_1(\alpha) = F_0(\alpha) = 1$, $f_n(sec(2\alpha)) = F_n(\alpha)$. $RB(3) = .7404$. $A_1 = 0$, $A_{2n} = \begin{pmatrix} A_n & A_n \\ A_n & \overline{A_n} \end{pmatrix}$.

$L_8$: $v \in L_8$ iff $v \in \mathbb{Z}_8$ and $v_i = a_i \pmod 2$ or $v_i = \overline{a}_i \pmod 2$. $L_8 \rightarrow \Lambda_8$: $v \in \Gamma_8$ iff $v \in L_8$ and $\sum_{i=1}^{2} 4v_i = 4m$. Contact number: 4320, radius: $\sqrt{2}$.

$L_{24}$: Sphere centers are equal $\pmod 2$ to $R(C)$ and $\sum_i v_i = 0 \pmod 4$.

| **Shape** | **Number** |
|---:|---:|
| $0^{16}, (-1)^8$ | 759 |
| $0^{16}, (-1)^6, 1^2$ | 21252 |
| $0^{16}, (-1)^4, 1^4$ | 53130 |
| $0^{16}, (-1)^2, 1^6$ | 21252 |
| $0^{16}, 1^8$ | 759 |
| $0^{22}, (-2)^2$ | 276 |
| $0^{22}, -2, 2$ | 552 |
| $0^{22}, 2^2$ | 276 |
| **Total** | 98256 |

Density of $L_{24}$: $\frac{2^{24}}{2 \times 2^{12}} = 2^{-11}$, first factor of 2 in denominator is from condition that the sum of the coordinates $= 0 \pmod 4$. Packing density: .0009647.

$\Gamma_{24}$: Express coordinates in $L_{24}$ in binary and retain the ones that satisfy the following conditions (a) the 24 1's bits are either all 0 or all 1, (b) the 2's bits form a row in $R(C)$, (c) 4's bits rows have even parity for points with 1's bits that are all 0 and odd otherwise. Equivalently, suppose $\vec{c} \in R(C)$ and for $m \in \mathbb{Z}$, define $\vec{c}(m) = \{v \in \mathbb{Z}^{24} : \sum_i v_i = 4m, c_i = 0 \rightarrow v_i = m \pmod 4, c_i = 1 \rightarrow v_i = m + 2 \pmod 4\}$, $\Lambda = \Lambda_{24} = \cup_m \vec{c}(m)$. Contact number: 98256 (even parity) + 98304 (odd) = 196,560. Density: .001929. Shapes: $(0^{16}, (\pm 2)^8)$, $(0^{22}, (\pm 4)^2)$, $((\pm 1)^{23}, (\pm 3))$. Each vertex is adjacent to 4600 others. Example: $(4, 4, 0, \ldots, 0)$ is adjacent to $(4, 0, \ldots, 0)$ - there are 88 of these, $(2, 2, \ldots, 0)$ - there are $77 \times 2^7$ of these and $(1, 3, \ldots, 0)$ - there are 2048 of these.

Definition: **Conway's group .O** is the set of rotations in $\mathbb{R}^{24}$ fixing $O$ pointwise and $\Lambda$ setwise.

Notation: $v_S = \sum_{i \in S} v_i$. The set $G\Lambda = \{2v_K, K \in R(C)\} \cup \{v_\Omega - 4_\infty\}$ generates $\Lambda$. If $v, w \in G\Lambda$, then $v \cdot v = 16n$ and $v \cdot w = 0 \pmod 8$. $\Lambda_n = \{x \in \Lambda, x \cdot x = 16n\}$. $\Lambda_1 = \emptyset$, $\Lambda_2$ consists of $\Lambda_2^2$ of shape $(0^{16}, (\pm 2)^8)$ - there are 97152 of these, $\Lambda_2^3$ of shape $((\pm 1)^{23}, (\pm 3)^1)$ - there are 98,304 of these, $\Lambda_2^4$ of shape $(0^{22}, (\pm 4)^2)$ - there are 1104 of these.

Structure in $.O$. $\Omega = PL(23)$, $\alpha : x \mapsto x + 1$, $\beta : x \mapsto 2x$, $\gamma : x \mapsto \frac{-1}{x}$, $\delta : x \mapsto 9x^3, x \notin Q$ and $\delta : x \mapsto \frac{x^3}{9}, x \in Q$. $PSL(23) =< \alpha, \gamma >$, $M_{24} =< \alpha, \gamma, \delta >$. If $\pi \in S_\Omega$, define $(v_i)^\pi = v_{\pi(i)}$. $\epsilon_S(v_i) = -v_i, i \in S$ and $\epsilon_S(v_i) = v_i, i \notin S$.

Preliminary results: If $S \in R(C)$, $\epsilon_S \in .O$. $E =< \epsilon_S >_{S \in R(C)}$, $M = M_{24}$. $N = EM$. If $\lambda \in .O$ and $\lambda$ fixes $v_i$ (some $i$) then $\lambda \in N$. If $\lambda \in N$ then $\lambda(\Lambda_2^4) = \Lambda_2^4$.

Main result: If $H > N$, $H$ is transitive on $\Lambda_2$ and $H = .O$. Proof: (1) $\Lambda_2^2$, $\Lambda_2^3$, $\Lambda_2^4$ are all $N$-orbits. A counting argument shows that the union of two of them can't be an $H$ orbit (otherwise, $p \mid |.O|$ for $p > 23$). Now define $\Lambda_2(x) = \{y : y \perp x\}$. (2) $H_x$ is transitive on $\Gamma_2(x)$ (hard). Orbit of $\Lambda_2^4$ under $N_x$ is $\{y\} \cup \{\lambda(y)\}, \lambda \in N_x$. There are 926. Since $M_{24}$ is $5-$transitive $|H_x : H_{x,y}| = 926$ and $|.O| = |H| = 196560 \cdot |H_x|$; further, $H_x$ is transitive on $\Lambda_2(x) = \{y : y \perp x\}$. An orbit of $H_x$ has 93150 elements so $|H_x| = (93150)|H_{x,y}|$ and $H_{x,y} = E_{10}M_{22}$. This gives the order of $H$ and shows $H = .O$.

The simple groups: **“.1”** $= .O/Z(.O)$. **“.2”** $= \{x \in .O, x$ stabilizes 2 points $v, w \in \Lambda_2 : |v-w| = 4\sqrt{2}\}$. **“.3”** $= \{x \in .O, x$ stabilizes 2 points $v, w \in \Lambda_2 : |v-w| = 4\sqrt{3}\}$.

**Reed-Solomon construction:** Fix $n$ elements, $< \alpha_1, ..., \alpha_n >$, $|F| \geq n$, $E(m) =< M\alpha_1, ..., M\alpha_n >$, $d(E(m_1, m_2)) \leq n + k - 1$.

# Chapter 2

# Computer Science

## 2.1 Basics

$f \in O(g) \leftrightarrow g \in \Omega(f) \leftrightarrow L_{x \to \infty} \frac{f(x)}{g(x)} < \infty$. $f \in o(g) \leftrightarrow g \in \omega(f) \leftrightarrow L_{x \to \infty} \frac{f(x)}{g(x)} = 0$. $G_1 \subset (G, E)$ is a strongly connected component iff $x, y \in G_1$ means there is a directed path $x \to y$ and a directed path $y \to x$.

**Recurrences:** Suppose $T(n) = aT(n/b) + f(n)$. If $f(n) = O(n^{log_b(a)-\epsilon})$ then $T(n) = \Theta(n^{log_b(a)})$. If $f(n) = \Theta(n^{log_b(a)})$ then $T(n) = \Theta(n^{log_b(a)} lg(n))$. If $f(n) = \Omega(n^{log_b(a)+\epsilon})$ and $af(n/b) \leq cf(n), c < 1$ then $T(n) = \Theta(f(n))$.

Adding an $m$ bit number and $n$ bit number takes $O(max(m, n))$ time and $O(m + n)$ space. Multiplying an $m$ bit number and $n$ bit number takes $O(mn)$ time and $O(m + n)$ space. The extended gcd of an $m$ bit number and $n$ bit number takes $O(mn)$ time and $O(m + n)$ space. $A^E \pmod{M}$ where $M$ is an $m$ bit number and $E$ is an $n$ bit number takes $O(nm^2)$ time. Rotation is linear in $\oplus$ but not in $+$. $GCD(u, v)$ average running time: $O((1 + \frac{max(u,v)}{(u,v)}) lg(min(u, v)))$.

```
heapify(A,i) {
    l:= LEFT(i); r:= RIGHT(i);
    if (l <= heapsize[A] and A[l]>A[i])
        M:= l;
    else
        M:= i;
    if (r <= heapsize[A] and A[r] > A[M])
        M:= r;
    if (M != i)  {
        swap (A[i], A[M]);
        heapify (A, M);
        }
    }

heapsort(A) {
    n= length[A];
    for(i=n;i>1;i--) {
        swap(A[1], A[i]);
        n--;
        heapify(A,1);
        }
    }

heapsort(A) {
    // stored in A[1...n]
    for(i=2;i<=n;i++)
        SiftUp(i);
    for(i=n;i>1;i--) {
        swap(A[1], A[i]);
        ShiftDown(i-1);
        }
    }
```

Finding the **shortest path** between x and y in $G = (V, E)$ where $l(e) > 0$ is the weight of $e \in E$ is $O(elg(n))$. d(v) contains an overestimate of the shortest path from s to v. $prev(v)$ contains the previous element in the shortest path from s to v. (Ford-Bellman version works for negative weights.)

```
shortestpath(V,E,s)  {
    for (v in V) {
     d(v):= infinity;
     prev(v):= empty-set;
        }
    H:= empty-set;
    d(s)= 0;
    mark(s);
    while (H is not empty) {
        h= deletemin(H);
        for e=(v, w) in E, w unmarked) {
            if(d(w)>(d(v)+l(e))) {
```

```
        d(w)=d(v)+l(e);             }
        prev(w)= v;               }
        insert(w, H);           }
        }
```

**Union-find:** Link(x, y): make x and y kids of a common parent. Parent node points to itself. $m$ UNION-FIND operations on $n$ elements is $O((m+n)lg(n))$.

```
makeset(x) {
  p(x)= x;
  rank(x)= 0;
  }

find(x) {
  if(x != p(x)) p(x)=  find(p(x));
  return(p(x));
 }
```

```
link(x, y) {
  if(rank(x)>rank(y)) swap(x, y);
  if(rank(x)==rank(y)) rank(y)++;
  p(x)= y;
  return(y);
  }

union(x,y) {
  link(find(x), find(y));
  }
```

**2-3 Trees:** Interior node has smallest key of 2nd and 3rd descendant. Insert: Do membership test stop at terminal position; id 2 kids, add one, if not, split into two, $(n, n')$. Add $n'$ using insert. Delete: If two kids left, done. Otherwise, try to move node of a siblings under common parent; if you can't, transfer this node to a sibling. If this leaves a singleton, in the parent, recurse the transfer on parent.

$G = (V, E)$. Each edge has a weight. Blue: $V = X \cup (V - X)$, no blue edges between $X$ and $V - X$. Pick and edge of min wt between them. Color it Blue. Red: Find a cycle with no red edge. Pick an edge of max wt, color it red. Apply blue and red in any order as long as possible. Blue edges form a MST.

**Floating Point Numbers:** $f \times b^{e-q}$ is represented as $(e, f)$.

Given $\epsilon > 0$ there is a multiplication algorithm such that the number of elementary operation $T(n)$ needed to multiply two $n$-bit numbers satisfies $T(n) < c(\epsilon)n^{1+\epsilon}$. Strassen: $T(n) = O(nlg(n))$.

**NP Completeness:** $P \subseteq N$. If $A \leq B$ [1] and $B \in P$ then $A \in P$. $L \in NPC$ if and only if $L \in NP$, $A \in NP \rightarrow A \leq L$. Classical computation theory classifies problems by a "certain" solution on all instances. Later we will encounter problems which can be solved in polynomial time "up to an arbitrary error, $\epsilon$" and call the class $RP$ for "randomized polynomial." $P \subseteq RP \subseteq NP$.

P: MST. Given a weighted graph, G, and a weight, K $\exists$ a tree, NP: TSP. Given a weighted graph, G, and a weight, K $\exists$ a cycle, C, that connects all nodes of G with weight $\leq K$.
P: Circuit value. NP: Circuit SAT.
P: 2-SAT: Use $\phi = (a_1 \vee b_1) \wedge ... \wedge (a_n \vee b_n)$ to form graph with nodes $a_i, b_i, \overline{a_i}, \overline{b_i}$ insert edges $\overline{a_i} \rightarrow b_i$ and $\overline{b_i} \rightarrow a_i$. Find strongly connected components. If no strongly connected component contains a variable and its negation, it is satisfiable; otherwise not. So 2-SAT is not NP hard. NP: 3-SAT. Note in disjunctive normal form SAT is easy but translating is hard.
P: matching. NP: 3D matching.
P: Linear Programming. NP: Integer Programming.

**Ford-Fulkerson:** Augmenting path p is a simple path from s to t that increases the flow.

```
Initialize flow, f to 0;
while (there is an augmenting path, p)
    augment flow along p;
return f;
```

**Undecidable:** Suppose $Term(P, X)$ is a boolean function which takes a program, $P$, and an input $X$. $Term(P, X)$ returns true iff $P$ terminates on $X$. $Term(P, X)$ returns false iff $P$ does not terminate on $X$. Theorem. $Term(P, X)$ does not exist. Suppose it did. Set

[1] $A \leq B$ means problem $A$ can be transformed to problem $B$ in polynomial time; this is called a reduction *from $A$ to $B$*.

```
diag(P,X) {
  if $Term(P,P)$==true
    loop
    }
```

$diag(diag)$ terminates iff it doesn't terminate. Contradiction.

**Stable Matching** (up to $n^2$ rounds). (1) Boy goes to favorite girl on list. (2) Girl tells highest choice "maybe", tells everyone else No. (3) Boy crosses off girls that have said no. (4) terminate in the round when every girl has told one boy "maybe", convert "maybe" to yes.

**Linear Programming Standard Form:** Maximize $x = C^T X$, subject to $AX = B$, $X \geq 0$. Problem: There may be exponentially many corners. (Reason: introduce to $n$ constraint inequalities $m$ slack variables; the corner points occur when $m$ variables are 0. There are $\binom{m+n}{m}$ ways to select the variables to be set to 0.) Simplex idea: move along growing paths instead of trying all corners randomly. Dual, minimize $x = B^T W$, subject to $A^T W = C$, $W \geq 0$.
Notation: basic variables $\neq 0$, basic variables $= 0$. A is an $m \times n$ matrix, with $m$ variables (including slack) and $m$ constraints. Tableau has basic variables and their values in 2 first columns. Top row is all variables as labels middle is matrix (A). Rightmost column is constants (B). Bottom row is $C - C^T X$ in terms of the non-basic variables.

1. Locate most negative coefficient in bottom row, call column containing it $x_j$.
2. Compute $\frac{B_i}{A_{ij}}$. The smallest one, denoted $k$, is the pivot.
3. Convert pivot to 1 and eliminate all coefficients in the same column.
4. Replace $x_k$ row by $x_j$.
5. repeat until no negative numbers in bottom row.

**NP Complete:** SAT, $k$-SAT ($k > 2$), k-clique, Vertex Cover, Independent set, Subset Sum, Partition, Bin Packing, Hamilton circuit. **Clique/SAT reduction:** Each occurrence of a variable is a vertex, edges between vertices if their occurrence in the clauses have same complementarity. $k$ is number of clauses. SAT/$k$-sat reduction: $l_1 \vee l_2 \vee \ldots \vee l_n \rightarrow l_1 \vee l_2 \vee x_1 \wedge \overline{x_1} \vee l_3 \vee x_2 \wedge \ldots \overline{x_{n-3}} \vee l_{n-1} \vee \ldots \vee l_n$. Phase transition for SAT: $\frac{clauses}{variables} \approx 4.3$. 3-SAT → MQ. Replace + with $\vee$, $\cdot$ with $\wedge$ , 1 with true, 0 with false. If $c_i = x_{i_1} \vee x_{i_2} \vee x_{i_3}$ add $x_{i_1} + x_{i_2} + x_{i_3} x_{i_4}$ and $x_{i_1} \cdot x_{i_2} + x_{i_2} \cdot x_{i_3} + x_{i_1} \cdot x_{i_3} = x_{i_5}$ and $x_{i_4} + x_{i_5} + x_{i_4} \cdot x_{i_5} = 1$.

**Hard core bit:** Let $f$ be a one-way function from $\{0,1\}^n$ to $\{0,1\}^n$, $x \in \{0,1\}^n$ , $r \in \{0,1\}^n$ , and let $G$ be a function that takes $\{0,1\}^n$ to $\{0,1\}^{n+1}$ by $G(x,r) = f(x), r, < x, r >$. Let $P$ be a prediction function. Goldreich-Levin: If there is an algorithm $A$ such that $|Prob_r[A(f(x),r) =< x, r >] - \frac{1}{2}| \geq \epsilon$ then there is an algorithm $I$ that produces a list $L$ of size $\leq \frac{1}{\epsilon^2}$ with $x$ in $L$, (2) $I$ runs in time polynomial in $n$ and $\frac{1}{\epsilon}$ and doesn't compute $f$. **Negligible:** smaller that inverse of any polynomial. Witness: $w : \Sigma^* \rightarrow P(\Gamma^*)$. **Decision problem:** $A_w \subseteq \Sigma^*$, $A_w = \{x \in \Sigma * | w(x) \neq 0\}$. Example: $x \in \Sigma^*$ is an encoding of a Boolean Form. $y \in \Gamma*$ is an encoding of a truth assignment. $\#P$ is class of witnesses, $w$, such that: (i) there is a P-time algorithm to decide if $x \in w(x)$ and (ii) $\exists k \in N$ such that $\forall y \in w(x), |y| \leq |x|^k$. $w \in \#P \rightarrow A_w \in NP$ and $A \in NP \rightarrow \exists w, A = A_w$.
Counting perfect matchings of a bipartite graph is $\#P$ complete.

**Finite State Machine:** Finite alphabet, $A$, finite states, $S$, two functions: $\delta : S \times A \rightarrow S$ and $\gamma : S \times A \rightarrow A$. Finite State Automata is FSM without output.

Language $L$ is a subset of $A^*$. **Regular expression**, $R$ over alphabet, $A$ with letters $a \in A$: (1) $\epsilon \in R$, (2) $a \in A$, (3) $r^* \in R$ if $r \in T$, (4) $r_1 r_2 \in R$ if $r_1, r_2 \in R$, (4) $r_1 \vee r_2 \in R$ if $r_1, r_2 \in R$. Language associated with a regular expression: (1) $L(\epsilon) = \{\epsilon\}$, (2) $L(a) = \{a\}$, (3) $L(r^*) L(r)^*$, (3) $L(r_1 r_2) = L(r_1) L(r_2)$, (4) $L(r_1 \vee r_2) = L(r_1) \cup L(r_2)$. $L$ is a **regular language** if $\exists r \in R$ with $L = L(r)$. **Phrase structured Grammar**, $G$, consists of (1) Vocabulary $V$, (2) terminals (denoted by lower case letters) $T \subseteq V$, (3) variables or non-terminals $V \setminus T$ (denoted by upper case letters), (4) a designated non-terminal $S$, called the start symbol, (5) a finite set $P$ of productions: $\alpha \rightarrow \beta$. $w \Rightarrow w'$ iff $\exists u, v, w = u \alpha v$ and $w' = u \beta v$.

Grammar types defined by production rule limitation: (1) Type 0: no limitations, (2) Type 1: production rules of the form $\alpha \rightarrow \beta$, $|\alpha| \leq |\beta|$ or $\alpha \rightarrow \epsilon$, (3) Type 2: production rules of the form $A \rightarrow \beta$, (4) Type 3: production rules of the form $A \rightarrow a$ or $A \rightarrow aB$, (5) **context free:** production rules of the form $A \rightarrow \beta$, (6) **context sensitive:** production rules of the form $\alpha A \alpha' \rightarrow \alpha\beta\alpha'$, (7) regular: production rules of the form $A \rightarrow a$, $A \rightarrow aB$ or $S \rightarrow \epsilon$. Backus-Naur form for type 2 context free grammar: (i) ::= replaces $\rightarrow$, (ii) non-terminals enclosed in brackets <> and (iii) all productions with the same non-terminal LHS are combined into a single RHS. Example: $<$ *sentence* $>$ ::= $<$ *noun phrase* $><$ *verb phrase* $>$, $<$ *noun phrase* $>$ ::= $<$ *noun* $>$ | $<$ *article* $><$ *noun* $>$, $<$ *noun* $>$::= boy.

A language $L$ can be generated by a type 3 (regular) grammar iff there is a finite automaton $M$ that accepts $L$. Pushdown automata (with infinite stack) recognize $L$ iff $L$ is context free. $L$ is recognized by a linear bounded automata (tape linearly bounded in length of input) iff $L$ is context sensitive.

**Minimizing State machines:** Two states, $s_i$, $s_j$, are 0 equivalent if the states have the same output for every input. States are $k+1$ equivalent if they have the same outputs for any input and their successor states are $k$ equivalent. Minimization procedure: Define $\pi_0$ as all states that are 0 equivalent. Do until no further refinement happens: sub-partition $\pi_k$ into $\pi_{k+1}$ into subblocks are $k+1$ equivalent. This terminates. When it does, merge equivalent states.

**Pumping Lemma:** Let $L$ be a finite state grammar accepted by a finite state machine, $M$, with $n$ states. If $\alpha$ is a string accepted by $M$ of length at least $n$, then $\alpha = u||v||w$ where $u||v^i||w$ is also in $L$.

Turing machines are FSMs with a bi-directionally infinite tape with a finite number of pre-marked squares and an additional transition function $\sigma : S \times A \rightarrow \{L, R, HALT\}$.

**Huffman algorithm:** Label each node with frequency. As long as more than one node is present, take the two nodes with the lowest frequency and combine them into a single node with the two combinants as children. New node has combined frequency. Left subnode has lower of two frequencies, right the higher. Read code by traversing from root. Left traversal at parent is 0, right, 1.
Resulting code is prefix free. Further $H(X) \leq l(x) \leq H(X) + 1$.

### 2.1.1 Concurrency

```
ECMA Consistency
    1. Reads and writes cannot move before volatile read.
    2. Reads and writes cannot move after volatile write.

CompareExchange(ref int loc, int value, int comp) {
    Monitor.Enter;
    ret= loc;
    if(ret==comp) loc= value;
    Monitor.Exit;
    return ret;
}

class SpinLock {
    volatile int isEntered=0;   // 1 if lock acquired
    int Enter() {
        while(CompareExchange(isEntered,1,0)!=0);
        }
    Exit() {
        isEntered= 0;
    }
}

Memory Consistency Rules
    1. Behavior of Thread in isolation is unaffected
    2. Reads cannot move before lock
    3. Writes cannot move after lock
```

```
DPLL(C,A) {
// C: clauses, A: literal assignments
// Termination:
// empty clause: unsatisfiable
// empty set of clauses: satisfiable
    if(A is empty)
        return SATISFIED;
    if(A has an empty clause)
        return UNSATISFIABLE;
// unit clause is a clause with one literal
    if unit clause (l) occurs in A
        return DPLL (assign(l,C), A + l));
    if l occurs with same polarity throughout
        return DPLL (assign(l,C), A + l));
    l= choose-literal(A);
    return DPLL (assign(l,C), A + l)) OR
           DPLL (assign(not l,C), A + not l));
    }
Note: If A, B, C are p-free,
(A | p) & (B|!p) &C) is inconsistent iff (A|B)&C is.
```

```
Chase(C,x) {
    set x to t;
    delete all clauses containing x from C;
    delete all occurences of !x from clauses in C;
    if (empty clause)
        return UNSATISFIABLE;
    if (unit clause l)
        return Chase(l,t);
    if (C is empty)
        return SATISFIED;
```

```
Priority Queue (arrays start at 1 here)

ExtractMax(A) {
    if(heapsize(A)<1)
        return error;
    max= A[1];
    A[1]= A[heapsize(A)];
    heapsize(A)=heapsize(A)-1;
    Heapify(A,1);
    return max;
}
```

```
Insert(A,k) {
    heapsize(A)=heapsize(A)+1;
    i= heapsize(A);
    while(i>1 & A[parent(i)]<k) {
        A[i]= A[parent(i)];
        i= parent(i);
        }
    A[i]= k;
    }
```

```
Select(A,k) {
    // select kth element from A[1,...n-1]
    if(k==0) return min(A);
    // For randomized, choose x in A at random
    x= SideSelect(A);
    Set B= < y in A: y <=x>
    Set C= < y: y>x >
    if(k<|B|) return Select(B,k)
    return Select(C,|B|-k);
    }
```

```
SideSelect(A,k) {
    for(i<=0<=n=INT(size(A)/5))
        Sort successive 5 elements
        // A[5i]<=A[5i+1]<=A[5i+2]<=A[5i+3]<=A[5i+4]
    R= < A[5i+2] > , 0<=i<=n
    x= SideSelect(R,Size(R)/2);
    // note x <= 3*INT((n-5)/10) elements.
    }
// Note E(T(n))= E(T(sn))+n,  x ~ 3/4
```

```
struct semaphore {
    int count;
    ProcessQueue queue;
    };

void P(semaphore s) {
    if(s.count>0) {
         (s.count)--;
    else
         s.queue.Insert(); // block
    }
```

```
void V(semaphore s) {
    if(s.queue.empty())
        (s.count)++;
    else
        s.queue.remove(); //schedule process
    }

shared semaphore s= 1;
     P(s);
     //critical section
     V(s);
```

```
Map()
Reduce()                                Readers/writers
Scan()     // || prefix                 linear sweep
Scatter()
Gather()
```

Architecture and current PCs: $P = C \times V^2 \times f$. Big **endian** word: $0, 1, 2, 3$ (descending byte address). Little endian word: $3, 2, 1, 0$ (descending byte address).

| **Optimization Level** | **Description** | **Level** |
|---|:---:|---:|
| High | Procedure inlining | 3 |
| Local | common subexpression | 1 |
| Local | constant propagation | 1 |
| Local | stack height reduction (expression tree) | 1 |
| Global | global common subexpression | 2 |
| Global | global constant propagation | 2 |
| Global | code motion | 2 |
| Global | induction variable elimination | 2 |
| Global | loop unrolling | 4 |
| Global | strip mining | 4 |
| Arch specific | strength reduction | 1 |
| Arch specific | pipeline scheduling | 1 |
| Arch specific | branch offset | 1 |

Effect on performance of Bubblesort (100K items). Base is 300MHz Sparc Ultra.

| **Optimization level** | **Relative performance** | **Clocks** | **Instructions** | **CPI** |
|---:|---:|---:|---:|---:|
| 0 | 1.00 | 158,615 | 114,938 | 1.38 |
| 1 | 2.37 | 66,990 | 37,470 | 1.79 |
| 2 | 2.38 | 66,521 | 39,993 | 1.66 |
| 3 | 2.41 | 65,747 | 44,993 | 1.46 |

SRAM: $.5 - 1ns$, $4,000\$/GB$. DRAM: $50 - 70ns$, $100\$/GB$. Disk: $10^7 ns$, $1\$/GB$. Dram address setup: 1 memory cycle, access time: 15 cycles, data transfer: 1 cycle. 4-way interleave plus multiword block gets time down to 20 cycles on average. Miss penalty to main: 500 cycles, to L2: 25 cycles. TLB: 512 entries. Miss: 100 cycles. Miss percentage; .5-1. Disk seek latency: 10 ms, rotational latency: 5 ms, transfer rate: 50 MB/s, MTTF: $10^6$ hours. Bus speed: system (800 MHz), NB (266 MHz), SB (33 MHz). Bandwidth:

| **Device** | **Bandwidth** |
|---:|---:|
| Memory | 3.2GB/sec |
| Disk | 150 MB/sec |
| AGP | 2.1 GB/sec |
| PCI | 132 MB/sec |
| NIC | 20 MB/sec |

**Dwarves:** Finite state machines, combinatorics, graphs, Structured/unstructured grids, dense matrix, sparse matrix, map-reduce, backtrack/branch-and-bound, $N$-body, FFT, Graphical models.

***LU*-factorization:** Let $A \neq 0$ be an $m \times n$ matrix. There are permutation matrices $P, Q$ such that $P^T AQ = LU$ where $L$ is lower triangular and $U$ is upper triangular. $QR$-factorization: Let $X \in \mathbb{C}^{n \times p}$ have rant $p$ then $x = QR$ where $Q$ is an orthogonal matrix and $R$ is an upper triangular matrix. $QR$-factorization via unitary operations is used in the least square approximation problem. Spectral decomposition: $U^H AU = diag(\lambda_1, \lambda_2, \ldots, \lambda_n)$. The eigenvalues of $X^H X$ are the sequence of singular values of $X$. For a $p \times q$ matrix, row major storage is $A[1, 1] = a[1]$, $A[1, 2] = a[2]$, …, $A[2, 1] = a[q + 1]$, etc., and in general, row major storage is $A[i, j] = a[(i - 1)q + j]$, column major storage is $A[1, 1] = a[1]$, $A[1, 2] = a[q + 1]$, …, $A[2, 1] = a[2]$, etc., and in general, column major storage is $A[i, j] = a[(j - 1)p + i]$. One step of Gaussian Elimination:

$$\begin{pmatrix} \alpha_{11} & \alpha_{12}^T \\ \alpha_{21} & A22 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ b_2 \end{pmatrix} \rightarrow \begin{pmatrix} \alpha_{11} & \alpha_{12}^T \\ 0 & A22 - \alpha_{11}^{-1}\alpha_{21}\alpha_{12}^T \end{pmatrix} = \begin{pmatrix} \beta_1 \\ b_2 - \alpha_{11}^{-1}\beta_1\alpha_{21} \end{pmatrix}$$

.

# Chapter 3

# Cryptography and Computer Security

## 3.1 Classical Systems

**Shannon Theory:** What is the amount of information in a number $n : 0 \leq n < 2^m$. Information learned about $Y$ by observing $X$ is $I(Y,X) = H(Y|X) - H(Y)$. Note $H(X|Y) = \sum p_X(x)H(Y|X = x)$ which is generally not equal to $\sum_{X,Y} p_Y(y|x)lg(p_Y(y|x))$. $H_E = lim_{N\to\infty} \frac{H(P^n)}{n}$. $H(K|C) = H(M|C) + H(K|M,C)$. **Perfect secrecy:** $Pr(M|C) = P(M)$. **Unicity Theorem:** Let $H$ be the entropy of the source (say English) and let $\Sigma$ be the alphabet. Let $K$ be the set of (equiprobable) keys, then $u = \frac{lg(|K|)}{(lg(|\Sigma|)-H)}$. $IC(f) = \frac{\sum(f_i(f_{i-1}))}{n(n-1)}$. $MC(f,f') = \frac{\sum f_i f_i'}{nn'}$.

**Vigeniere alphabet chaining:** If $\alpha$ is the mixed plaintext alphabet and $\beta$ is the mixed cipher alphabet underneath, rearranging with the plain alphabet into its normal form we get the tableaux:

| 1 | 2 | ... | n |
|---|---|---|---|
| $\beta(\alpha^{-1}(1))$ | $\beta(\alpha^{-1}(2))$ | ... | $\beta(\alpha^{-1}(n))$ |
| $\beta(\alpha^{-1}(1)+1)$ | $\beta(\alpha^{-1}(2)+1)$ | ... | $\beta(\alpha^{-1}(n)+1)$ |
| ... | ... | ... | ... |
| $\beta(\alpha^{-1}(1)+n-1)$ | $\beta(\alpha^{-1}(2))$ | ... | $\beta(\alpha^{-1}(n)+n-2)$ |

Note that the columns have the same sequence of characters as the original rows — if plain A corresponds to cipher F and if plain F corresponds to cipher W then the distance between plain A and plain F is the same as cipher F and cipher W in the original sequence.

**Heburn:** Five rotors, two ratchet controls. Key: $[i,j,k,m,n]$ and 2 ratchet stepping controls at right and left $(l,r)$. Rightmost $(R_5)$ rotor moved after every enciphered letter. Leftmost $(R_1)$ moved when fast rotor reached position specified by $r$. $a(m)$ character in line to $R_5$. When the leftmost rotor hit $l$ the middle $(R_3)$ rotor moved one position. Equation: $(p)KC^iR_1C^{-i}C^jR_2C^{-j}C^kR_3C^{-k}C^mR_4C^{-m}C^nR_5C^{-n}L = c$, $C$ is the cyclic in alphabetical order. Solution: $c(m) = a(m)C^{(m+p)}R_5C^{-(m+p)}L$, $d(m,p) = c(m)L^{-1}C^{(m+p)}R_5^{-1}C^{-(m+p)}$ then $d(m,p)R_5^{-1}C^{n-m}R_5 = d(n,p)$. Practical application relies on the IC for the monoalphabetic substitution (imagine all the input letters are the same). If $i = d(m,p)$, $j = d(n,p)$ and $k = n-m$. To remove noise, tally $s'[i,j,k] = \sum_m \sum_n s[i,m,k-m]s[m,j,n]$, this can be iterated.

**Enigma:** $K$: Keyboard. $P = (ABCDEFGHIJKLMNOPQRSTUVWXYZ)$. $N$: First Rotor. $M$: Second Rotor. $L$: Third Rotor. $U$: Reflector. Note: $U = U^{-1}$. $i,j,k$: Number of rotations of first, second and third rotors respectively. $c = (p)P^iNP^{-i}P^jM^{P-j}P^kLP^{-k}UP^kL^{-1}P^{-k}P^jM^{-1}P^{-j}P^iN^{-1}P^{-i}$. Later military models added plug-board or "Stecker "$(S)$:

$$c = (p)SP^iN^{P-i}P^jMP^{-j}P^kLP^{-k}UP^kL^{-1}P^{-k}P^jM^{-1}P^{-j}P^iN^{-1}P^{-i}S^{-1}.$$

Total key including rotor wiring (in bits): $67.1 + 3 \times 88.4 = 312.3$. **Method of Batons (no Stecker):** Let $N$ be the fast rotor and $Z$ the combined effect of the other apparatus, then, $N^{-1}ZN(p) = c$ at first letter; assuming other rotor doesnt turn, $P^{-i}N^{-1}P^iZP^{-i}NP^i(p) = c$ or $ZP^{-i}N(p(i))P^i = P^{-i}NP^ic(i)$. Rejewski: Let $Q = MLUL^{-1}M^{-1} = Q^{-1}$, the first 6 permutations (used to encrypt settings twice) are:

$$A = A^{-1} = SP^1NP^{-1}QP^1N^{-1}P^{-1}S^{-1}, B = B^{-1} = SP^2NP^{-2}QP^2N^{-1}P^{-2}S^{-1}$$

$$C = C^{-1} = SP^3NP^{-3}QP^3N^{-1}P^{-3}S^{-1}, D = D^{-1} = SP^4NP^{-4}QP^4N^{-1}P^{-4}S^{-1}$$
$$E = E^{-1} = SP^5NP^{-5}QP^5N^{-1}P^{-5}S^{-1}, F = F^{-1} = SP^6NP^{-6}QP^6N^{-1}P^{-6}S^{-1}$$

Their products and ciphertext $(c_1c_2c_3c_4c_5c_6)$ satisfy:

$$AD = SP^1NP^{-1}QP^1N^{-1}P^3NP^{-4}QP^4N^{-1}P^{-4}S^{-1}, (c_1)AD = c_4$$
$$BE = SP^2NP^{-2}QP^2N^{-1}P^3NP^{-5}QP^5N^{-1}P^{-5}S^{-1}, (c_2)BE = c_5$$
$$CF = SP^3NP^{-3}QP^3N^{-1}P^3NP^{-6}QP^6N^{-1}P^{-6}S^{-1}, (c_3)CF = c_6$$

So we can find $AD$, $BE$ and $CF$ after about 80 messages. To solve for rotors if $S$ is known. First note the following **Theorem**: If two permutations of the same degree consist of disjoint transpositions then their product contains an even number of cycles of the same length (and conversely) and cillies (guessed simple indicators like aaa) align cycles. Let $U = P^{-1}S^{-1}ASP = PNP^{-1}QPN^{-1}P^{-1}$, $V = P^{-2}S^{-1}BSP^2$, etc, then $VW = NP^{-1}N^{-1}(UV)NPN^{-1}$, $WX = NP^{-1}N^{-1}(VW)NPN^{-1}$, etc. which can be solved for $N$.

Assume we know all rotor wirings and the plaintext for some received ciphertext. We do not know plugboard, rotor order, ring and indicator.

```
Position   123456789012345678901234
Plain Text OBERKOMMANDODERWEHRMACHT
CipherText ZMGERFEWMLKMTAWXTSWVUINZ
```

Observe the loop $A[9] \rightarrow M[7] \rightarrow E[14] \rightarrow A$. $(E)M_7M_9M_{14} = E$, where $M_i$ is the effect of the machine at position $i$. British Bombe searched probable text for these loop isomorphisms. False alarms have probability $\frac{1}{26}$ for each independent loop tested.

## 3.2 Public Key Systems

**RSA:** $n = pq$, choose e, $ed = 1 \pmod{\phi(pq)}$, $e$ is often $2^{16} + 1$ for efficiency.

**DLP:** Given $g, h$ and $h = g^x$, find $x$. **DHP:** Given $g, a = g^x, b = g^y$, find $z = g^{xy}$. **DDH:** Given $g \in G, a = g^x, b = g^y, c = g^z$, determine if $z = xy$. $DDH \leq DHP \leq DLP$. **Theorem:** $FACTOR \leq SQRT \leq FACTOR$. If the RSA problem is hard, then RSA is secure under a chosen plaintext attack. If DHP is hard, El Gamal is secure under a chosen plaintext attack.

**Finding square roots (mod p):** We want $x : x^2 = a \pmod p$. First check $(\frac{a}{p}) = 1$. If $p = 3 \pmod 4$, $x = \frac{a(p+1)}{4} \pmod p$. If $p = 5 \pmod 8$, $b = \frac{a(p-1)}{4} = \pm 1 \pmod p$, then if $b = 1, x = \frac{a(p+3)}{8} \pmod p$, otherwise, if $b = -1, x = \frac{(2a)(4a)(p-5)}{8} \pmod p$. This leaves the hard case, $p = 1 \pmod 8$). The algorithm of *Tonelli and Shanks* solves this case (and the others). Again, we want $x : x^2 = a \pmod p$. Put $p-1 = 2^e q$, $q$, odd. Choose $n$: $(\frac{n}{p}) = -1$, $z = nq \pmod p$, $Q = \frac{(q-1)}{2}$. Put $y = z$; $r = e$; $x = aQ \pmod p$; $b = ax^2 \pmod p$; $x = ax \pmod p$. Now $R = 2^r - 1, ab = x^2, yR = -1, bR = 1$;. Do the following: loop:
if$(b = 1)$ return$(x)$;
Let $M = 2^m$. For smallest $m > 0 : bM = 1 \pmod p$
if$(m = r)$ return non-residue;
$t = y^{2^{r-m-1}} \pmod p$; $y = t^2 \pmod p$; $r = m$; $x = xt$; $b = by$; goto loop;

Factoring $n$ may be equivalent to computing $\phi(n)$ which is equivalent to finding $d$. **Strong primes:** $p - 1$ has a large prime factor $r$, $p + 1$ has a large prime factor $a$, $r - 1$ has a large prime factor $t$. **Miller-Rabin** has error probability $p = \frac{1}{4}$.

**El Gamal:** Let $g$ be a generator of $F_q^*$. A picks $a$ at random, this is A's secret. User picks $k$ at random and sends $(g^k, Pg^{ka})$. **El Gamal Signature:** $g$ is a primitive element $\mathbb{Z}_p^*$. $(p, g, y = g^x)$ are public, $x$ is secret. To sign $m$, pick $k$: $1 \leq k \leq p-2$ with $(k, p-1) = 1$. $sig_K(m, k) = (r, s)$, $r = g^k$, $s = k^{-1}(m - xr)$. $ver_k(m, r, s)$ is true iff $y^r r^s == g^m$. Note: $k$ must be different for each signature and $m$ must be a hash. Recommended parameters: $> 768$ bits. Existential forgery if hash isn't used in El Gamal: For key elements, $(< \mathbb{Z}_p >, g, a)$, pick $(u, v)$, $r = g^u g^v = g^{u+av}$. $s = -rv^{-1} \pmod{p-1}$, $M = su$. Note that $t = r^s y^r = g^{su}$.

**Diffie Hellman:** Let $g$ be a generator of $F_q^*$. A generates $a \in F_q^*$ at random and transmits $g^a$, B generates $b \in F_q^*$ at random and transmits $g^a$, they use $g^{ab}$ as key.

**Blinding and E-cash:** Let $M$ be a note or check. To blind, generate random $k$. Let $(e, d, n)$ be the bank's key and $H$, a hash. Send bank $r = H(M)k^e$. Bank sends back $r^d$, now multiply by $k^{-1}$. For fraud resistant protocol, do this for a bunch of $k_s$'s. Bank signs one of them.

**DSA:** Pick $p, q$, $2^{159} < q < 2^{160}$, $2^{511+64t} < p < 2^{512+64t}$, $0 \leq t \leq 8$ with $q|(p-1)$. Let $x$ be a primitive root $\pmod p$. Set $g = x^{\frac{p-1}{q}} > 1 \pmod q$. Finally, pick $a$ at random and set $A = g^a \pmod p$. $p, q, g, A$ are public, $a$ is secret. To sign $M$: generate random $k : k < q$. Set $r = g^k \pmod q$ and compute $s = k^{-1}(h(M)+xr) \pmod q$, where $h$ is a cryptographic hash. Signature is $(r, s)$. To verify: $u_1 = s^{-1}h(M) \pmod q$, $u_2 = s^{-1}r \pmod q$, $v = g^{u_1}g^{u_2} \pmod p \pmod q$. If $v = r$, it verifies. Unlike El Gamal signature, $s$ does not carry full information about $p$ (only $\pmod q$) and since $q$ is large, the Pohlig-Hellman attack is harder.

**Montgomery Arithmetic:** Suppose $(r, n) = 1$; think of $r = 2^k$, $2^k < n < 2^{k+1}$. $R = ab \pmod n$. $\overline{a} = ar \pmod n$. $rr' - nn' = 1$ MontPro$(\overline{a}, \overline{b})$: $t = \overline{a}\overline{b}$; $m = tn' \pmod r$; $u = \frac{mn+t}{r}$; if$(u > n)$ $u- = n$; return$(u)$; MontMult$(a, b, n)$: Compute $n'$ ; $\overline{a} = ar \pmod n$; $\overline{b} = br \pmod n$; $\overline{x} = MontPro(\overline{a}, \overline{b})$; $x = MontPro(\overline{x}, 1)$;
return$(x)$.

**NAF:** Let $k = \sum_{j=0}^{l} s_j 2^j, s_j \in \{0, 1\}$, NAF form is $k = \sum_{j=0}^{l+1} c_j 2^j, c_j \in \{-1, 0, 1\}$, conversion is achieved by following algorithm:

```
$c_0 = 0;$
$for(j = 0; j \leq l; j++)\{$
$c_{j+1} = \lfloor (k_j + k_{j+1} + c_j)/2 \rfloor;$
$s_j = k_j + c_j - 2c_{j+1};\}$
```

AMD-64 3Ghz dual core timings.

| Algorithm | KSize | T($\mu$-sec) | Cycles | Algorithm | KSize | T($\mu$-sec) | Cycles |
|---|---|---|---|---|---|---|---|
| ECDSA-SIGN | 256 | 4942 | 14,827,000 | ECDSA-VERIFY | 256 | 9,848 | 29,546,000 |
| ECDSA-SIGN | 384 | 13,000 | 38,860,000 | ECDSA-VERIFY | 384 | 25,900 | 77,639,000 |
| ECDSA-SIGN | 521 | 29,500 | 88,287,000 | ECDSA-VERIFY | 521 | 58,900 | 176,524,000 |

| Algorithm | KeySize | T($mu$-sec) | Cycles | Algorithm | KeySize | T($mu$-sec) | Cycles |
|---|---|---|---|---|---|---|---|
| DSA-SIG | 512 | 1,077 | 3,233,000 | DSA-VERIFY | 512 | 2,142 | 6,427,000 |
| DSA-SIG | 768 | 2,332 | 6,999,000 | DSA-VERIFY | 768 | 4,641 | 13,924,000 |
| DSA-SIG | 1024 | 4,027 | 12,083,000 | DSA-VERIFY | 1024 | 8,015 | 24,047,000 |

| Algorithm | KeySize | T($\mu$-sec) | Cycles | Algorithm | KeySize | T($\mu$-sec) | Cycles |
|---|---|---|---|---|---|---|---|
| RSA-SIGN | 1024 | 3,488 | 10,465,000 | RSA-VERIFY | 1024 | 168 | 505,000 |
| RSA-SIGN | 2048 | 22,905 | 68,717,000 | RSA-VERIFY | 2048 | 608 | 1,825,000 |
| RSA-SIGN | 3072 | 72,494 | 217,491,000 | RSA-VERIFY | 3072 | 1,340 | 4,021,000 |
| RSA-SIGN | 4096 | 168,548 | 505,664,000 | RSA-VERIFY | 4096 | 2,363 | 7,091,000 |

| Algorithm | KeySize | T(sec) | Algorithm | KeySize | T(sec) |
|---|---|---|---|---|---|
| RSA KeyGen | 1024 | .37 | ECC KeyGen | 160 | .0053 |
| RSA KeyGen | 2048 | 3.5 | ECC KeyGen | 224 | .0056 |
| RSA KeyGen | 3072 | 11.2 | ECC KeyGen | 256 | .0067 |

**McEliece Cryptosystem:** Bob chooses $G$, an $[n, k, d]$ linear code, $G_1 = SGP$ where $P$ is an $n \times n$ permutation matrix and $S$ is a $k \times k$ invertible matrix. To send a message to Bob, Alice adds an error, $e$, of weight $t$, $y = xG_1 + e$. To decrypt, (1) compute $y_1 = yP^{-1} = xSG + e_1$; (2) apply error decode to $y_1$ to get $x_1$; (3) compute $x_0 : x_0G = x_1$; (4) compute $x = x_0S^{-1}$. Want $d$ to be large. For example, use Goppa code

$(n = 2^m, d = 2t+1, k = n = mt)$: $m = 10, t = 50$ to get $[1024, 524, 101]$.

## 3.3 Symmetric Key Systems

**CBC:** $y_0 = IV$, $y_i = E_K(x_i + y_{i-1})$. **OFB:** $z_0 = IV$, $z_{i+1} = E_K(z_i)$, $y_i = x_i + z_i$. **CFB:** $y_0 = IV$, $z_i = E_K(y_{i-1})$, $y_i = x_i + z_i$. **CTR:** $z_i = E_K(Nounce||ctr)$, $y_i = x_i \oplus z_i$. **HMAC:** $(K, m) \mapsto h((K \oplus a)||h(K \oplus b)||m)$. **GCM:** $F = GF(2^{128}), p(x) = 2^{128} + x^7 + x^2 + x + 1$, $(z_0, y_0) = (IV, 0^{128})$, $(z_1, y_i) \mapsto (z_{i+1}, y_{i+1})$ by $z_{i+1} = \pi_i(z_i)$, if $\pi_i(x) = 0$, $z_i \oplus y_i$ otherwise and $y_{i+1} = y_i >> 1$ if $LSB(y_i) = 0$ otherwise $y_{i+1} = (y_i >> 1) \oplus R$, $R = [11100001||0^{120}$. Define $X \cdot Y = (z_{128}, y_{128})$. $inc_s(X) = MSB_{len(X)-s}(X)||[int(LSBs(X)) + 1 \pmod{2^s}]_s$.
$GHASH_H(X), len(X) = 128m$: $H = E_K(0^{128})$. $Y_0 = 0^{128}$, $Y_{i+1} = (Y_i \oplus X_{i+1}) \cdot H$. return $Y_m$.
$GCTR_K(ICB, X)$: If $X$ is the empty string, then return the empty string as $Y$. $n = \lceil (len(X)/128 \rceil$. Let $CB_1 = ICB$, $CB_i = inc_{32}(CB_{i-1}, i = 1 \ldots n$. $Y_i = X_i \oplus E_K(CB_i)$. $Y_n{}^* = X_i{}^* E_K(CB_i)$. return $Y$.
$GCM - AE_K(IV, P, A)$: $H = E_K(0^{128})$. If $len(IV) = 96$, $J_0 = IV||0^{31}||1$. If $len(IV) \neq 96$, let $s = 128\lceil len(IV)/128 \rceil - len(IV)$, and let $J_0 = GHASH_H(IV||0^{s+64}||len(IV)^{64})$. $C = GCTR_K(inc_{32}(J0), P)$. Let $n$ Define $S = GHASH_H(A||0^v||C||0^u||len(A)^{64}||len(C)^{64})$. $T = MSB_t(GCTR_K(J_0, S))$. return $(C, T)$.

**Recurrence for LFSR of length** $k$**:** $s_j = c_1 s_{j-1} + \ldots c_k s_{j-k}$. **Hamming weight:** $w_H(x) = \#\{n : x_n \neq 0\}$. **Modular weight:** $w_M(x) = |x'|$ where $x' = x \pmod{2^n}$ and $-2^{n-1} < x' \leq 2^{n-1}$. **NAF weight:** $w_{NAF}(x) = \#\{i < n : \alpha_i \neq 0\}$. $\Delta^{\oplus}(x, y)$, $\Delta^{+}(x, y)$, $\Delta^{\pm}(x, y)$ are the xor, modular and signed differences respectively. **Distortion for map** $\varphi$**:** $D(\varphi, d_1, d_2) = sup_{x \neq y} \frac{d_2(\varphi(x), \varphi(y))}{d_1(x,y)} sup_{x \neq y} \frac{d_1(x,y)}{d_2(\varphi(x), \varphi(y))}$. $f(x_1, \ldots, x_n)$ is $m$**-correlation immune** if $I(f(x_1, \ldots, x_n); x_{i_1}, \ldots, x_{i_m}) = 0$ for any choice of the $i_k$. This happens when the boolean spectrum of $F(w)$ is 0 when $w$ has weight $\leq m$. **Shrinking Generator:** Take two LFSR: $LFSR_1$ and $LFSR_2$ synchronously clocked. Use $LFSR_2(t)$ in stream when $LFSR_1(t) = 1$. Take $LFSR_i(t) = x_i(t)$ for $i = 1, 2, \ldots, n$ use $f(x_1(t), x_2(t), \ldots, x_n(t))$ where $f$ is non linear. For $k$ stage shift register design, where stage $i$ has $n_i$ bits of state, keysearch takes $2^{n_0 + n + 1 + \ldots + n_{k-1}}$ while correlation attack [Example: Geffe combiner: $f(x, y, z) = xy \oplus yz \oplus z$], then $f(x, y, z) = x$ with $p = \frac{3}{4}$] takes $2^{n_0} + 2^{n+1} + \ldots + 2^{n_{k-1}}$.

Let $M_{j,k}(x) = \begin{pmatrix} x_j & x_{j+1} & x_{j+2} & \cdots & x_{j+k-1} \\ x_{j+1} & x_{j+2} & x_{j+3} & \cdots & x_{j+k} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{j+k-1} & x_{j+k} & x_{j+k+1} & \cdots & x_{j+2k-2} \end{pmatrix}$. If $< x_i >$ is generated by an LFSR of length $N$ but not one shorter then $det(M_{j,N}(x)) = 1$ and $det(M_{j,n}(x)) = 0, n > N$. If $x_{n+m} = c_0 x_n + \ldots + c_{m-1} x_{n+m-1}$ and $c(x) = x^m + c_{m-1} x^{m-1} + \ldots + c_0$, the associated connection polynomial, is irreducible, then the sequence repeats at an interval of $k = 2^m - 1$.

**Connection polynomial** for $L_n(\vec{s})$ is $c(x) = 1 + c_1 x + \ldots + c_l x^l$ with $c(x) = 0$ if $L_n(\vec{s}) = 0$. Define $d_n$ as $n$th discrepancy, suppose $m$ is the position of change of length in minimal generating LFSR. $L_m(\vec{s}) \leq L_n(\vec{s})$ and $L_{m+1}(\vec{s}) = L_n(\vec{s})$. The recurrence is $c^{(n+1)}(x) = c^{(n)}(x) - d_n d_m{}^{-1} x^{n-m} c^{(m)}(x)$. The synthesis algorithm is $O(n^2)$. **Theorem:** A LFSR of length $k$ has maximal period $(= 2^k - 1)$ iff its connection polynomial is primitive. Proof: Let $G(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{m-1} x^{m-1} + \ldots$, $a_m = c_1 a_{m-1} + \ldots + c_m a_1$, etc. We get a recurrence yielding $\frac{K}{1 - c(x)}$, $f(x) = 1 - c(x)$. If sequence is $p$, $G(x) = (a_0 + a_1 x + \ldots + a_{m-1} x^{m-1}) + (a_0 + a_1 x + \ldots + a_{m-1} x^{m-1}) x^p + \ldots = \frac{(a_0 + a_1 x + \ldots + a_{m-1} x^{m-1})}{1 - x^p} = \frac{K}{(f(x))}$.

**Massey's Lemma:** If $L_n(\vec{s})$ generates $< s_0, \ldots, s_{n-1} >$ but not $< s_0, \ldots, s_n >$ then $L_{n+1}(\vec{s}) \geq max(L_n(\vec{s}), n + 1 - L_n(\vec{s}))$. Proof: Suppose $L$ generate $< s_0, s_1, \ldots, s_{n-1} >$ but not $< s_0, s_1, \ldots, s_n >$ and let $L'$ with $L'_{n+1}(\vec{s}) = l'$ then $l' \geq n + 1 - l$. Proof. If $l \geq n$, $l' \geq 1$ so it's true. If $l < n$, let $c_i$ be the coefficients of $L$ and $c'_i$, the coefficients of $L'$. $s_j + \sum_{i=1}^{l} c_i s_{j-i} = 0$ for $j = l, l+1, \ldots, n-1$ but not for $j = n$ and $s_j + \sum_{i=1}^{l'} c'_i s_{j-i} = 0$ for $j = l', l'+1, \ldots, n$ so $-\sum_{i=1}^{l} c_i s_{j-i} = \sum_{i=1}^{l} c_i \sum_{k=1}^{l'} c'_k s_{n-i-k}$ Switching the order of summation, the second sum is $s_n$ which is a contradiction.

**Berlekamp-Massey:** Given $s_1, s_2, \ldots, s_{n-1}$ output linear complexity $L$.

1. $C(x) = 1, L = 0, m = -1, b(x) = 1, n = 0$.

2. $d = S_n + \sum_{i=1}^{L} c_i s_{n-i}$.
3. If $(d == 1)$ $t(x) = c(x), c(x)+ = b(x)x^{n-m})$ if$(L \leq \frac{n}{2})$ $L = n + 1 - L, m = n, b(x) = t(x)$
4. $n = n + 1$;

```
RC4Init() {
    for (i=0; i<256; i++)
        s[i]= i;
    fill k[] with key repeating
         as necessary;
    j= 0;
    for(i = 0; i<256; i++) {
        j= (k[i]+s[i]+j) (mod 256);
        swap(s[i], s[j]);
        }
    i= 0;
    j= 0;
    }

byte Next() {
    i= (i+1) (mod 256);
    j= (j+s[i]) (mod 256);
    swap(s[i], s[j]);
    return(s[(s[i]+s[j]) (mod 256)]);
    }
```

Let $\Lambda(s^n)$ be the associated linear complexity of the sequence $< s_i >$ of length $n$ and $N_n(L)$ be the number of sequences of length $n$ with linear complexity $L$, then $N_n(L) = 2N_{n-1}(L) + N_{n-1}(n-L)$, if $n \geq L > \frac{n}{2}$; $N_n(L) = 2N_{n-1}(L)$, if $L = \frac{n}{2}$; and, $N_n(L) = N_{n-1}(L)$, if $\frac{n}{2} \geq L \geq 0$. So $N_n(L) = 2^{min(2n-2L, 2L-1)}$, if $n \geq L > 0$, $N_n(L) = 1$, if $n \geq L = 0$. $E(\lambda(s^n)) = \frac{n}{2} + \frac{4+R_2(n)}{18} - 2^{-n}(\frac{n}{3} + \frac{2}{9})$, $Var(\Lambda(s^n)) = \frac{86}{81}$.

**RC4 Weakness:** Let $S_i$ be the state at time $i$, $N = 2^n$ ($n = 8$, usually). Let $< z_i >$ be the output sequence. $P(z_2) = 0) = \frac{2}{N}$. Proof: Suppose $S_0[2] = 0$, $S_0[1] \neq 2$, $S_0[1] = X$, $S_0[X] = Y$. Round 1: $i = 1$, $X = S_0[1] + 0$. Exchange $S_0[1]$ and $S_0[Y]$. Round 2: $i = 2$, $j = X + S_1[2] = X$, Output $S_1[S_1[2] + S_1[X]] = S_1[X] = 0$. So $P(z_j = 0) \approx \frac{1}{N} + \frac{1}{N}(1 - \frac{1}{N}) \approx \frac{2}{N}$. So by Bayes, if $z_2 = 0$, we can extract byte of state with probability $\frac{1}{2}$.

**ANSI 9.17 random stream generator:** $I = E_k(D)$. $x_i = E_k(I \otimes s)$ and $s = E_k(x_i \otimes s)$. **FIPS 186 One Way Function (OWF):** $t$, $c$ 160 bits. Output $G(t, c)$ where $t = H_1 || H_2 \ldots || H_5$. Pad $c$ with 0s to get 512 bit block $X$. Break $X$ into 16 32 bits words $x_0, \ldots, x_{15}$ and set $m = 1$, apply iterative step of SHA-1.

```
Dual Elliptic Curve RNG
  s[0] in [0,1, ..., #E-1]
  output 240 bits
  for(i=1 to k {
    s[i]= x(s[i-1]P);
    r[i]= lsb[240] x(s[i]Q);
    }
  return(r[1] ... r[k]);

// State for Hash_DRBG
  V // seedlen bits
  C // seedlen bits
  reseedCtr

Hash_DRBG_Instantiate(entBitsIn, nonce,
                      extraEnt)
  seedBits= entBitsIn||nonce||extraEnt;
  seed= Hash_df(seedBits, seedlen);
  V= seed;
  C= Hash_df((0x00||V), seedlen);
  reseedCtr= 1;
  return;

Hash_DRBG_Reseed(entBitsIn, addInBits)
  seedBits= 0x01||V||entBitsIn||addInBits;
  seed= Hash_df(seedBits, seedlen);
  V= seed;
  C= Hash_df((0x00||V), seedlen);
  reseedCtr= 1;
  return;

Hash_DRBG_Generate(numReqBits, addInBits)
  if(reseedCtr>reseedInterval) then
       Reseed;
  if(addInBits!=NULL)
     w= Hash(0x02||V||addInBits);
     V=(V+w) mod 2**seedlen;
  returnedBits= Hashgen(numReqBits, V);
  H= Hash(0x03||V);
  V=(V+H+C+reseedCtr) mod 2**seedlen;
  reseedCtr= reseedCtr+1;
  return returnedBits;

Hashgen(numReqBits, V)
  m= reqNumBits/outlen;
  data= V;
  W= NULL;
  for i= 1 to m
    w= Hash(data);
    W= W||w;
    data= (data+1) mod 2**seedlen;
  returnedBits= Leftmost numReqBits
                bits of W;
  return returnedBits;

Hash_df(inBits, numRetBits):
  temp= NULL;
  m= numRetBits/outlen;
  counter= 8-bit representation of 1;
  for i= 1 to len do
    temp= temp|| Hash(counter||
            numRetBits||inBits);
    counter= counter+1;
  reqBits= Leftmost numRetBits of temp;
  return reqBits;

// State for CTR_DRBG
    V // outlen bits
    C // keylen bits
    reseedCtr
    nStrength
    fPrediction

CTR_DRBG_Update(provided_data, Key, V):
  temp= NULL;
  while(len(temp)<seedlen) do
    V=(V+1) mod 2**outlen;
    outBits= blockEncrypt(Key, V);
    temp= temp||ouput_block;
  temp= Leftmost seedlen bits of temp;
  temp= temp^provided_data;
  Key= Leftmost keylen bits of temp;
  V= Rightmost outlen bits of temp;
  return Key and V;

// Full Entropy
CTR_DRBG_Instantiate(entBitsIn, extraEnt):
  // Ensure that the length of
  // extraEnt is seedlen bits.
  temp= len(extraEnt);
  if(temp<seedlen))
    extraEnt= extraEnt||
                [seedlen-temp] bits of 0;
  seedBits= entBitsIn^extraEnt;
  Key= [keylen] bits of 0;
  V= [outlen] bits of 0;
  (Key, V)= Update (seedBits, Key, V);
  reseedCtr= 1;
  return;

// Derivation function required
CTR_DRBG_Instantiate(entBitsIn, extraEnt):
  seedBits= entBitsIn||nonce||extraEnt;
  seedBits= Block_Cipher_df(seedBits, seedlen);
  Key= 0 of[keylen];
  V= 0 of[outlen];
  (Key, V)= Update (seedBits, Key, V);
  reseedCtr= 1;
  return;
```

```
// Full entropy
CTR_DRBG_Reseed(entBitsIn, addInBits):
  temp= len(addInBits);
  if(temp<seedlen), then
      addInBits= addInBits||
           [seedlen-temp] bits of 0;
  seedBits= entBitsIn^addInBits.;
  (Key, V)= Update (seedBits, Key, V);
  reseedCtr= 1;
  return;

// Derivation Function Required
CTR_DRBG_Reseed(entBitsIn, addInBits):
  seedBits= entBitsIn||addInBits;
  seedBits= Block_Cipher_df(seedBits,
                            seedlen);
  (Key, V)= Update (seedBits, Key, V);
  reseedCtr= 1;
  return;

CTR_DRBG_Generate(numReqBits, addInBits):
    if reseedCtr>reseedInterval, then
            reseed;
    if(addInBits!=NULL)
        temp= len(addInBits);
    if(temp<seedlen))
        addInBits= addInBits||
                 [seedlen-temp] bits of 0;
        (Key, V)= Update (addInBits, Key, V);
    else
      addInBits= [seedlen] bits of 0;
   temp= NULL;
   while(len(temp)<numReqBits) do:
     V=(V+1) mod 2**outlen;
     outBits= blockEncrypt(Key, V);
     temp= temp||outBits;
   returnedBits= Leftmost numReqBits of temp;
   // Update for backtracking resistance.
   (Key, V)= Update(addInBits, Key, V);
   reseedCtr= reseedCtr+1;
   return returnedBits;

BCC(Key, data):
  CV= [outlen] bits of 0;
  n= len(data)/outlen;
  Split the data into n blocks of outlen bits
       forming block[1] to block[n];
  for i= 1 to n do
    inBlock= CV^block[i];
    CV= blockEncrypt(Key,inBlock);
  outBits= CV;
  Return outBits;

Block_Cipher_df(numRetBits, inBits)
  if(numRetBits>maxNumBits), then
    return ERROR;
  L= len(inBits)/8;
  N= numRetBits/8;
  S= L||N||inBits||0x80;
  // Pad S with zeros, if necessary.
  while(len(S) mod outlen) != 0
       S= S||0x00;
  temp= NULL;
  i= 0;
  K= Leftmost keylen bits
         of 0x00010203...1D1E1F.
  while len(temp)<keylen+outlen)
    IV= i||[outlen-len(i)] bits of 0;
    temp= temp||BCC(K,(IV||S));
    i= i+1;
  K= Leftmost keylen bits of temp;
  X= Next outlen bits of temp;
  temp= NULL;
  while len(temp)<numRetBits
    X= blockEncrypt(K, X);
    temp= temp||X;
  reqBits= Leftmost numRetBits of temp;
  return reqBits;
```

MGF property: Given no input and partial output, remaining output is unpredictable.

```
mgf1(mSeed, nLen)
1. if (mLen>2^32$, return error
2. T= ||;
3. uL= ceiling(mLen/hLen),
   // hLen is length of hash used
4. for(c=0; c<uL;c++)
       T= t|| h(mSeed || c);
5. output leading bits
```

```
PSS-Encode(M, emBits, salt, sLen)
// M- message
// emBits- bits of EM >= 8 hLen + 8 sLen + 9
1. emLen= ceil(enBits/8);
2. if (l(M)> largest message), return error;
3. mH= h(M)
4. if( emLen < hLen+sLen+2 ), return error;
5. M'= (0x00)^8 || mH || salt
6. H= h(M');
7. DB= (0x00)^(emLen-hLen-sLen-2)
       || 0x01 || salt
8. dbMask= mgf(H, emLen-hLen-1);
9. maskedDB= DB^dbMask;
10. Set leftmost 8*emLen-emBits to 0 in maskedDB
11. EM= maskedDB || H || 0xbc
12. return EM;
```

```
emsa-pkcs(M, emLen)
// emLen= l(EM)>= tLen+11
1. H= h(M)
2. T= hash-prefix || H ;  // tLen= l(T)
3. EM= 0x00 || 0x01 || (0xff)^(emLen-tLen-3) || T;
4. return EM;
```

**Blum-Blum-Shub:** Select $p, q$ each $= 3 \pmod 4$, $n = pq$, $s \in [1, n-1]$-seed, $(s, n) = 1$ $x_0 = s^2 \pmod n$ for(i=1 to l) $x_i = x_{i-1}^2 \pmod n$ $z_i = LSB(x_i)$ . Next bit test: Given $l$ bits, no polynomial time algorithm can predict the $l+1$st with probability $> \frac{1}{2} + \epsilon$.

```
RC6 input: A,B,C,D, r rounds, w-bit round keys in S[0...2r+3].
```

```
RC6() {
    B= B+S[0];
    D= D+S[1];
    for(i=1;i<=r;i++) {
        t= (B*(2B+1)) <<< lg(w);
        u= (D*(2D+1)) <<< lg(w);
        A= ((A^t)<<<u)+S[2i];
        C= ((C^u)<<t)+S[2i+1];
        (A, B, C, D) = (B,C,D,A);
        }
    A= A+S[2r+2];
    C= C+S[2r+3];
    }
```

```
// Key L[0 to k-1];
RoundKeys(L,S,k) {
    S[0]= 0xB7E15163al Elliptic Curve RNG
    s[0] in [0,1, ..., #E-1]
    output 240 bits
    for(i=1 to k {
        s[i]= x(s[i-1]P);
        r[i]= lsb[240] x(s[i]Q);
        }
    return(r[1] ... r[k]);
```

**OAEP:** Want to send $m$. Let $\rho(r)$ be a pseudo random number generator initialized with seed $r$. Calculate $a = \rho(r) \oplus m$, $b = r \oplus H(a)$. Send $E(a||b)$.

**Traitor tracing:** $y = \Pi_{i=1}^{2k} {h_i}^{\delta_i}$. $\delta$ is the representation vector with respect to the base $h$. Convex combinations of representations are also solutions. Generate $l \geq 2k+2$ private keys with security parameter $s$ to defend against coalition of size $k$. Choose $g$, a generator of $G_q$, $r_i$, $i = 1, 2, \ldots, 2k$ at random with $h_i = g^{r_i}$. Public key is $< y, h_1, h_2, \ldots, h_{2k} >$ where $y = \Pi_{i=1}^{2k} {h_i}^{\alpha_i}$. Private key is $\theta_i$ with $\theta_i \gamma^{(i)}$ a representation of $y$. $\Gamma = \{\gamma^{(1)}, \gamma^{(2)}, \ldots, \gamma^{(l)}\}$ are public. Each $\gamma^{(i)} = \sum_j \gamma_j$ is a codeword. $\theta_i = \frac{\sum_{j=1}^{2k} r_j \alpha_j}{\sum_{j=1}^{2k} r_j \gamma_j}$. Encrypt: pick

$a$ randomly $C =< My^a, {h_1}^a, \ldots, (h_{2k})^a >$. To decrypt $C =< C, H_1, \ldots, H_{2k} >$, compute $M = \frac{S}{U^{\theta_i}}$ where $U = \Pi_{i=1}^{2k} {H_i}^{\gamma_i}$. Tracing: Assume $q > max(l, 2k)$ examine $l-2k-1 \times 2k$ matrix $A$

$$A = \begin{pmatrix} 1 & 1 & 1 & ... & 1 \\ 1 & 2 & 3 & ... & l \\ 1^2 & 2^2 & 3^2 & ... & l^2 \\ ... & ... & ... & ... & ... \\ 1^{l-2k-1} & 2^{l-2k-1} & 3^{l-2k-1} & ... & l^{l-2k-1} \end{pmatrix}$$

Rowspace ↔ polynomials of degree $\leq l-2k-1$. Let $B$ be formed by the column vectors $b_1, b_2, \ldots, b_{2k}$, the basis of vectors satisfying $AX = 0(q)$. $\exists w$ of Hamming wt $\leq k$ with $vB = d$, null space of B $\leftrightarrow f$ with $deg(f) \leq l-2k-1$ and $v - w =< f(1), f(2), \ldots f(l) >$ in all but (at most) $k$ places. Use Berlekamp to find $f$ from $v$.

## 3.4 Public Key Analysis

Generally **primality testing** is $O(n^{clg(lg(n))})$ and **factoring** (Number Theory Sieve) is $O(n^{c(lg(n)^{\frac{1}{3}}(lg(lg(n))^{\frac{2}{3}}})$.

**Solovay-Strassen:** Choose $1 \leq a \leq (n-1)$. If $(\frac{a}{n}) = a^{\frac{n-1}{2}} \pmod{n}$ then $n$ is prime with probability $\frac{1}{2}$. Use the following to compute $(\frac{a}{n})$: (1) $(\frac{m_1 m_2}{n}) = (\frac{m_1}{n})(\frac{m_2}{n})$, (2) $(\frac{m}{n}) = -(\frac{n}{m})$, if $m = n = 1, 3 \pmod{4}$, $(\frac{m}{n}) = (\frac{n}{m})$, otherwise, (3) $(\frac{2}{n}) = -1$, if $n = 1, 7 \pmod{8}$, 1, if $n = 3, 5 \pmod{8}$, (4) $(\frac{2^k t}{n}) = (\frac{2}{n})^k(\frac{t}{n})$.

**Pockington:** Let $n > 1$ and $s \mid (n-1)$. Suppose for some $a$, (1) $a^{\frac{n-1}{2}} = 1 \pmod{n}$, and (2) $\forall q, q|s$, $(a^{\frac{n-1}{q}} - 1, n) = 1$. Then $p \mid n$. So if $s > \sqrt{n}$, $n$ is prime.

**Pollard $p-1$:** (for numbers with a factor, $p$, where $p-1$ factors into small primes). $n$ is $B$ smooth. $Q = \prod_{q|B} q^{\lfloor \frac{ln(n)}{ln(p)} \rfloor}$. $\forall a, a^Q = 1(p)$, $gcd(a^Q - 1, n) = d$. $Q = LCM_{p^i \leq B}(p^i)$.

**Application of Pollard-$\rho$ to discrete log:** Let $x_{i+1} = f(x_i)$ and $n$ be the order of the multiplicative group. $m = (\mu(1 + \lfloor \frac{\lambda}{\mu} \rfloor)$. For the discrete log problem, $h = g^x$, $x_m = x_{2m}$; tail of length $\lambda$, $\mu$ is length of cycle. $\lambda \leq m < \lambda + \mu$. Let $S_1, S_2, S_3$ partition the multiplicative set, $1 \notin S_2$ and define $x_{i+1} = f(x_i) = hx_i, x_i \in S_1$ $x_{i+1} = f(x_i) = x_i^2, x_i \in S_2$ $x_{i+1} = f(x_i) = gx_i, x_i \in S_3$; $a_{i+1} = a_i \pmod{n}, x_i \in S_1$ $a_{i+1} = 2a_i \pmod{n}, x_i \in S_2$ $a_{i+1} = a_i + 1 \pmod{n}, x_i \in S_3$; $b_{i+1} = b_i + 1 \pmod{n}, x_i \in S_1$ $b_{i+1} = 2b_i \pmod{n}, x_i \in S_2$ $b_{i+1} = b_i \pmod{n}, x_i \in S_3$ and consider 3-tuples $(x_i, a_i, b_i)$ with $(x_0, a_0, b_0) = (1, 0, 0)$. Then $log_g(x_i) = a_i + b_i log_g(h)$ is an invariant of the sequence. When $x_m = x_{2m}$, $a_m + xb_m = a_{2m} + xb_{2m}$ and $x = \frac{a_{2m} - a_m}{b_{2m} - b_m}$.

**Quadratic Sieve:** Factor Base is $\mathcal{B}_B = \{-1, 2, 3, \ldots p_l\}, p_l \leq B$. Define a sequence $a_i = ([\sqrt{n}] + i)^2 - n$. Set $b_i = (\sqrt{\lfloor n \rfloor}] + i)$, $b_i^2 - a_i = n$. For the $a_i$'s that factor over the base, find a bunch using linear algebra after taking the log. Then for these $a_{i_l}$'s, $\prod a_{i_l} = y^2 \pmod{n}$, where $y$ is a product of the corresponding $b_i$'s. Sieving finds $B-$smooth elements of sequence. **Sieving:** Fix sieving interval $-C \leq s \leq C$, compute $f(s) = (x + \lfloor \sqrt{n} \rfloor)^2 - n$, find $s : p \mid f(s)$ - i.e.- find roots of $f(x) = 0 \pmod{p}$, walk through sieving interval by steps of $p$ for others. Divide each $f(s)$ in sieving interval by the higher dividing power of each $p$, ones with 1 or $-1$ are smooth. Wiedemann algorithm for solving sparse linear equations is $L_n[\frac{1}{2}, 2v + o(1)]$. Sieving is $O(L_n[\frac{1}{2}, v + \frac{1}{4v} + o(1)]/p)$. Reason: Let $\psi(X, Y)$ be the number of $Y-$smooth numbers in $[1, X]$. $Pr(a \in [1, X]\ is\ Y - smooth) = \frac{\psi(X,Y)}{X}$; expected trials to find one: $\frac{X}{\psi(X,Y)}$ need about $\pi(Y)$ to get enough for a square and each takes $\pi(Y)$ work to test, so the total work is $W(X, Y) = \frac{\pi(Y)^2 X}{\psi(X,Y)}$. Minimum occurs when $Y = e^{\frac{1}{2}\sqrt{ln(X)ln(ln(X))}}$ and $X \approx n^{\frac{1}{2}+\epsilon}$. Try $n = 24961, 157$.

**Number Field Sieve:** $F = \{p : p \leq B\}$ want to find $a, \lambda : b = a + N\lambda$ and $b$ is $B$-smooth so $\prod p \in Fp^{a_p} = \prod_{p \in F} p^{b_p} \pmod{N}$. Procedure: (1) Fix $\lambda$, (2) let Array $A$ have $A + 1$ 0's, (3) $\forall p \in F$, add $lg(p)$ to all positions congurent to $-\lambda N \pmod{p}$ and (4) choose $a$ larger than some threshhold. Construct two monics of degree $d_1, d_2$: $f_1(m) = f_2(m) = 0 \pmod{N}$ using the number fields $K_1 = \mathbb{Q}(\theta_1)$ and $K_2 = \mathbb{Q}(\theta_2)$. We have two homomorphisms $\phi_i : \mathbb{Z}[\theta_i] \rightarrow \mathbb{Z}/N\mathbb{Z}$, with $\theta_1 \mapsto m$. Set $S = \{(a, b) \in \mathbb{Z}^2 : (a, b) = 1\}$ satisfying $\prod_S (a - b\theta_1) = \beta^2$ and $\prod_S (a - b\theta_2) = \gamma^2$. Then $\phi_1(\beta)^2 = \phi_2(\gamma)^2 \pmod{N}$ and $(\phi_1(\beta) - \phi_2(\gamma)) \mid N$. What's left is to find $S$, $\beta^2$, $f_1$ and $f_2$. An algebraic integer is smooth if the the ideal it generates is divisible only by small primes. Define $F_i(X,Y) = Y^{d_i} f_i(X/Y)$ then $N_{\mathbb{Q}[\theta_i)/\mathbb{Q}}(a - bi) = F_i(a,b)$. Use two factor bases $\mathcal{F}_i = \{(p, \theta_i - r), f_i(r) = 0 \pmod{p}\}$. $F_i(a,b) = \prod_{(p_j,r)\in\mathcal{F}_i} {p_j}^{s_j^{(i)}}$. Sieving: (1) fix $a$, (2) init sieve array $-B \le b \le B$, $S[b] = lg(F_1(a,b) \cdot F_2(a,b))$, (3) $\forall (p,r) \in \mathcal{F}_i$ subtract $lg(p)$ from every element: $a - rb = 0 \pmod{p}$, (4) the desired $b$'s are the ones: $S[b] \le Threshhold$. $\prod_{(a,b)\in S}(a - b\theta_i) = \mathcal{I}^2, \mathcal{I} \subseteq \mathbb{Z}[\theta_i]$. Now find enough relations such that $\prod_S (a - b\theta_1) = \beta^2$, etc.

Example: $N = 290^2 + 1$, $f_1(x) = x^2 + 1, f_2(x) = x - m, m = 290$. $f_1(m) = f_2(m) = 0 \pmod{N}$.

| $x$ | $y$ | $N(x-iy)$ | Factors | $x - my$ | Factors |
|---|---|---|---|---|---|
| $-38$ | $-1$ | 1445 | $5 \cdot 17^2$ | 252 | $2^2 \cdot 3^3 \cdot 7$ |
| $-22$ | $-19$ | 845 | $5 \cdot 13^2$ | 5488 | $2^4 \cdot 7^3$ |

$(-31 + i) = -(2+i)(4-i)^2$, $-22 + 19i) = -(2+i)(3-2i)^2$, $(-38+m)(-22+19m) = 2^6 3^2 7^4 = 1176^2 = (31-12i)^2$. $\phi_1(31-12i) = 31 - 12m = -3449$, $(-3449)^2 = (1176)^2$. $(N, -3449 + 1176) = 2273$, $(N, -3449 - 1176) = 37$.

**Pollard $\rho$:** $f(x) = x^2 + 1$. Compute $x_{i+1} = f(x_i)$. Look at $gcd((x_i - x_j), n)$ for factors of $n$. Floyd's trick: Compute $(x_i, x_{2i})$ from $(x_{i-1}, x_{2i-2})$, test $(x_{2i} - x_i, n)$. **Expected tail length:** $\sqrt{\frac{\pi n}{8}}$. **Expected loop:** $\sqrt{\frac{\pi n}{2}}$.

Define $L_n[u,v] = e^{v ln(n)^u ln(ln(n))^{1-u}}$. $L_n[0,v]$ is polynomial and $L_n[1,v]$ is exponential. Let $\psi(x,B)$ be the $B-$smooth numbers $\le x$. Let $\epsilon > 0$; if $x \ge 10$ and $w \le (ln(x))^{1-\epsilon}$, then $\psi(x, x^{\frac{1}{w}}) = xw^{-w+f(x,w)}$ and $\frac{f(x,w)}{w} \to 0$ for $w \to \infty$. Result: As $n \to \infty$, $\psi(n^a, L_n[u,v]) = n^a L_n[1-u, -(\frac{a}{v})(1-u) + o(1)$. For **QS:** $a \approx \frac{1}{2}$. **NFS** discrete log is $L_n[\frac{1}{3}, (\frac{64}{9})^{\frac{1}{3}}]$. **MPQF:** $O(e^{(\sqrt{ln(N) ln(ln(N))})})$. QS and NFS cross at 350 bits. Results below. Note: $1 MIP - yr = 3.1 \times 10^{13}$ instructions. $120000 Mip - years = 55 Opteron - 2.2 GHz - years$.

| | RSA-129 | RSA-130 | RSA-200 |
|---|---|---|---|
| Date | 4/1996 | 8/1999 | 5/2005 |
| Time (MIP-years) | 500 | 8,000 | 120,000 |
| Rows | $3.5 \times 10^6$ | $6.7 \times 10^6$ | $6.4 \times 10^7$ |
| Non Zero Members | $1.4 \times 10^8$ | $4.2 \times 10^8$ | $1.1 \times 10^{10}$ |
| NZ/R | 39 | 62 | 171 |
| Linear Algebra (hrs) | 68 | 224 | 2160 |

| RSA key | ECC key | Symmetric Key | ArithOps | SieveMem | LAMem |
|---|---|---|---|---|---|
| 428 | 110 | 51 | $5.5 \times 10^{17}$ | 2GB | 128MB |
| 512 | 119 | 56 | $1.7 \times 10^{19}$ | 64MB | 10GB |
| 768 | 144 | 69 | $1.1 \times 10^{23}$ | - | - |
| 1024 | 163 | 79 | $1.3 \times 10^{26}$ | 256MB | 100GB |
| 2048 | 222 | 109 | $1.5 \times 10^{35}$ | - | - |

Finding discrete logs using **Pohlig-Silver:** Let $g$ a generator for $F_q$. Find $x$ such that $g^x = y$. $q - 1 = {p_1}^{\alpha_1} \dots {p_k}^{\alpha_k}$. First, precompute: $r_{i,j} = g^{\frac{j(q-1)}{p_i}}$, for $j = 1, 2, \dots, p-1$. Want to find $x \pmod{p^{\alpha_i}}$, for each $p$ then use Chinese Remainder Theorem (CRT). $x = x_0 + x_1 p + x_2 p^2 + \dots + x_{\alpha-1} p^{\alpha-1}$. $y^{(q-1)/p} = g^{x(q-1)/p} = r_{p,x_0}$. This yields $x_0$. Next put $y_1 = \frac{y}{g^{x_0}}$. This reduces the discrete log over any group order to discrete log over $p$. This takes $O(\sum_{p||G|}(e(p)(lg(|G|) + \sqrt{p}))$ if we use Pollard.

Finding discrete logs using **Index Calculus**. $g$ a generator for $F_q$ with $q = p^n$. Find $x$ such that $g^x = y$. Precompute: Let $f(x)$ be an irreducible polynomial of degree $n$ over $F_p$. Let $B_m$ be the set of irreducible polynomials of degree $\le m$. Pick random $t$ and compute $c(x) = g(x)^t = c_0 \prod_{B_m} a(x)^{\alpha_{c,a}}$. $ind(c(x)) = ind(c_0) + \sum_{a \in B_m} \alpha_{c,a} ind(a(x)) = t \pmod{q-1}$. Now solve for the $ind(a(x))$. To compute $ind(y(x))$, pick random $t$ and compute $y(x)g(x)^t = \prod_{B_m} a(x)^{\alpha_{c,a}} \pmod{f(x)}$. This runs in $L_p[\frac{1}{2}, c + o(1)]$. In $E_F$, there is no good basis corresponding to primes.

**Square Roots:** Suppose $(\frac{a}{p}) = 1$, so that $a$ is a square and let $n$ be a quadratic non-residue $\pmod{p}$.

Want to find $x$: $x^2 = a \pmod{p}$. Set $p - 1 = 2^{\alpha}s$ and put $b = n^s \pmod{p}$ and $r = a^{\frac{s+1}{2}}$. Then $a^{-1}r^2$ is a $2^{\alpha-1}$ root of $1 \pmod{p}$. $b$ is a primitive $2^{\alpha-1}$ root of $1 \pmod{p}$ otherwise $n$ would not be a non-residue now use powers of $b$ to make the $x = b^j r$ a perfect square; to do this set $j = j_0 + 2j_1 + \ldots + j_{\alpha-2}2^{\alpha-2}$ and do the Pohlig routine.

**Shanks Baby/Giant:** $< g >= G$. Given $y = g^x$, find $x = log_g(y)$. Put $m = \sqrt{n}$, Compute $(j, g^j)$ for $j = 1, \ldots, m$ sorted by second coordinate. Set $t \leftarrow g^{-m}$, $s \leftarrow y$.
For(i=0 to m-1) { /* is $s$ second component?*/ if($s = g^j$) return($x = im + j$); $s \leftarrow st$}. Alternative: Solve $g^x = a \pmod{p}$. Pick $n : n^2 \geq (p-1)$ and compute $g^j \pmod{p}$ and $ag^{-nk} \pmod{p}$ for $0 \leq j, k \leq n$; match two lists giving $g^j = ag^{-nk} \pmod{p}$ or $g^{j+nk} = a \pmod{p}$.

**Boneh-Joux attack** on El Gamal/RSA with small messages and no preprocessing. Suppose we encrypt an $m$ bit message $M$ which is small then $M$ is often smooth — i.e. $M = M_1M_2$. If the El Gamal system is $< p, g, y = g^a >$ and either the order of $g$ is small (less than $\frac{p}{2^m}$) or $p - 1 = qs$ and the DL problem is tractable for subgroups of order $s$, much of the time ($\approx .18$) which solves the problem using about $2^{m/2}$ exponentiations. Here is the general problem: Let $z \in G_q \rightarrow \mathbb{Z}_p^*$, where $G_q$ is a subgroup of order $q$; if $\Delta < 2^m$ and $u = z\Delta \pmod{p}$ then given $u$, find $z$. Here is a meet in the middle shortcut. Suppose $\Delta = \Delta_1\Delta_2$, $\Delta_1 \leq 2^{m_1}, \Delta_2 \leq 2^{m_2}$, by tablizing $\Delta_1^q$ for possible $\Delta_1$'s and trying every possible $\Delta_2$ in $(\frac{u}{\Delta_2})^q = \Delta_1^q (\text{mod} p)$, we can find $\Delta = \Delta_1\Delta_2$ in $O(2^{m_1} + 2^{m_2})$ time and $2^{m_1}$ space. With $m_1 = m_2 = 32$ this can solve for a 64 bit session key with probability about .18.

Defense for Boneh-Joux: **OAEP (IND-CCA)** $c = E(m) = f(a = M \oplus G(r)||b = r \oplus H(a)$
REACT: $E(m, r||s) : (a = f(x, r), b = k \oplus m, c = H(m, x, a, b), k = G(x)$. For El Gamal: $a = Rand(1..q)$, $R = Rand(< g >)$, $A = g^a$, $A' = Rg^a$, $k = G(R)$, $B = E_k(m)$, $C = H(R, m, A, a', B)$.

| $n$ | $p$ | $H(B_n(p))$ | $n$ | $p$ | $H(B_n(p))$ |
|---|---|---|---|---|---|
| 2 | .5 | 2 | 3 | .5 | 3 |
| 2 | .60 | 1.94 | 3 | .60 | 2.91 |
| 2 | .75 | 1.62 | 3 | .75 | 2.43 |
| 2 | .80 | 1.44 | 3 | .80 | 2.16 |
| 2 | .90 | .93 | 3 | .90 | 1.4 |
| 2 | .95 | .57 | 3 | .95 | .85 |

| $\lambda$ | $H(P(\lambda))$ | $\lambda$ | $H(P(\lambda))$ |
|---|---|---|---|
| .5 | .91 | .60 | 1.00 |
| .75 | 1.14 | .80 | 1.18 |
| .90 | 1.27 | .95 | 1.31 |

**Shamir's attack on RSA with multiplication bug:** Assume the RSA implementation uses the CRT (which yields a speedup of 4) and let the public key be $n = pq$ with $p < q$. Suppose that $a \times b$ (two 32 bit quantities) is computed incorrectly on a computer with a word size of $w$ bits. We can pick $c = \lfloor \sqrt{n} \rfloor$ so $p < c < q$. Put $c = c_k 2^{wk} + c_{k-1}2^{w(k-1)} + \ldots + c_1 2^w + c_0$ and select $m$ such that $m = c_k 2^{wk} + c_{k-1}2^{w(k-1)} + \ldots + a2^w + b$. Assume we can have the flawed machine compute $m^d \pmod{n}$. In the CRT (since $p < m < q$ is likely), $m_1 = m \pmod{q}$ will be computed correctly but $m_2 = m \pmod{p}$ will be computed incorrectly. Since $a$ and $b$ are not likely to appear in the representation of $m_1$ but will appear in $m_2$, ${m_1}^2$ will be computed correctly but ${m_2}^2$ will be computed incorrectly. When the combined result $y = m^d \pmod{n}$ is computed, $y$ will likely be correct $\pmod{p}$ but incorrect $\pmod{q}$. Thus $p \mid y^e - m$ but $p \nmid y^e - m$ and $p = (y^e - m, n)$. Padding interferes with this attack.

**Weiner's attack:** $|\alpha - \frac{p}{q}| \leq \frac{1}{2q^2}$ with $d < \frac{1}{3}N^{\frac{1}{4}}$. Put $N = pq$, $q < p < 2q, ed = 1 \pmod{\phi}$. $|\frac{e}{\phi} - \frac{k}{d}| < \frac{1}{d\phi}$ with $ed - k\phi = 1$, $|N - \phi| = |p + q + 1| < 3\sqrt{N}$ so $|\frac{e}{N} - \frac{k}{d}| \leq \frac{3k}{d\sqrt{N}} < \frac{1}{2d^2}$ and $\frac{k}{d}$ arises as a convergent, $\alpha = \frac{e}{N}$.

**Coppersmith:** Let $f(x) \in \mathbb{Z}[x]$ be a monic polynomial of $deg(f) = d$, $N \in \mathbb{Z}$. If $\exists x_0 : f(x_0) = 0 \pmod{N}$ with $|x_0| \leq X = N^{\frac{1}{d}-\epsilon}$, one can find $x_0$ in time polynomial in $lg(N)$ and $\frac{1}{\epsilon}$ for fixed $d$. This can be used to extend the *Franklin Reiter* attack.

If $f(x) = f_0 + f_1x + \ldots + f_dx^d$ and $\exists x_0 : f(x_0) = 0 \pmod{n}$ with $|x_0| < N^{\frac{1}{d}}$, find $x_0$ efficiently. The idea is to find $h(x) \in \mathbb{Z}[x]$ which shares a root with $f \pmod{n}$ with $||h||^2 = \sum_{i=0}^{deg(h)} |h_i|^2$ with $||h||$ small.

**Lemma:** Let $h(x) \in \mathbb{Z}[x]$, $deg(h) \leq n, X, N \in \mathbb{Z}^{>0}$; suppose $||h(XN)|| < \frac{N}{\sqrt{n}}$; if $|x_0| < X$ satisfies $h(x_0) = 0 \pmod{N}$ then $h(x_0) = 0$.

Suppose $f(x_0) = 0 \pmod{n}$ then $f(x_0)^k = 0 \pmod{N^k}$. For some $m$, set $g_{u,v}(x) = N^{m-v}x^u f(x)^v$, $0 \leq u < d, 0 \leq v \leq m$ then $g_{u,v}(x_0) = 0 \pmod{N^m}$. Fix $m$, try to find $a_{u,v} \in \mathbb{Z} : h(x) = \sum_{u \geq 0} \sum_{v=0}^{m} a_{u,v} g_{u,v}(x)$ that satisfies the lemma; that is $||h(xX)|| \leq \frac{N^m}{\sqrt{d(m+1}}$ with $h(xX) = \sum_{u \geq 0} \sum_{v=0}^{m} a_{u,v} g_{u,v}(xX)$ that. Use LLL for this minimization problem. LLL conditions on $< b_1, b_2, \ldots, b_n >$ are $\mu_{ij} = \frac{<b_i, b_i^*>}{<b_i^*, b_i^*>}$, $b_i^* = b_i - \sum_{j<i} \mu_{ij} b_j^*$, $||b_i^*||^2 \geq (\frac{3}{4} - \mu_{i,i-1}^2)||b_{i-1}||^2$, if $x \in L, ||b_1|| \leq 2^{\frac{m-1}{2}}||x||$, $||b_1|| \leq 2^{\frac{m}{4}} \Delta^{\frac{1}{m}}$.

Example: $f(x) = x^2 + ax + b$. Want to find $x_0 : f(x_0) = 0 \pmod{N}$. Set $m = 2$. $g_{00}(xX) = N^2$, $g_{10}(xX) = XN^2x$, $g_{01}(xX) = bN + aXxN + XN^2x$, $g_{11}(xX) = bNXx + aX^2x^2N + N^2X^3x^3$, $g_{02}(xX) = b^2 + 2abXx + (a^2+2b)X^2x^2 + 2aX^3x^3 + X^4x^4$, $g_{12}(xX) = b^2Xx + 2abX^2x^2 + (a^2+2b)X^3x^3 + 2aX^4x^4 + X^5x^45$.

$$A = \begin{pmatrix} N^2 & 0 & bN & 0 & b^2 & 0 \\ 0 & XN^2 & aXN & bNX & 2abX & Xb^2 \\ 0 & 0 & NX^2 & aNX^2 & (a^2+2b)X^2 & 2abX^2 \\ 0 & 0 & 0 & NX^3 & 2aNX^3 & (a^2+2b)X^3 \\ 0 & 0 & 0 & 0 & X^4 & 2aX^4 \\ 0 & 0 & 0 & 0 & 0 & X^5 \end{pmatrix}. \quad det(A) = N^6X^{15},\ ||b_1|| < 2^{\frac{3}{2}}NX^{\frac{5}{2}}.$$

$b_1 = Au$, $Bu = (u_1, u_2, \ldots, u_6)$, $||h(xX)|| \leq \frac{N^2}{\sqrt{6}}$, $|x_0| \leq X = \frac{N^{\frac{2}{5}}}{48^{\frac{1}{8}}}$ and $|x_0| < N^{.39}$.

**Common Modulus attack:** Suppose $(e_1, e_2) = 1$ and $m$ is encrypted both with an $< n, e_1 >$ scheme and a $< n, e_2 >$ scheme; let $c_1 = m^{e_1} \pmod{n}$ and $c_2 = m^{e_2} \pmod{n}$ with $d_1e_1 + d_2e_2 = 1$ then $m = {c_1}^{d_1}{c_2}^{d_2}$.

**Small exponent attacks:** Suppose $e = 3$ and $c_1 = {m_1}^e$, $c_2 = {m_2}^e$ with $m_2 = m_1 + \delta$, where $\delta$ is known. Put $F(x) = x^e - c_1 \pmod{n}$ and $G(x) = (x + \delta)^e - c_2 \pmod{n}$ then $(x - m) \mid (F(x), G(x))$ and we can recover $m$. Now if $\delta$ is unknown but $|\delta| < n^{\frac{1}{9}}$ and there is an algorithm, $A$ (e.g.- Coppersmith's algorithm), that can find the roots, $\alpha$ of $f(x) = 0 \pmod{n}$ when $|\alpha| < n^{\frac{1}{9}}$, the foregoing attack can be extended. To do this, consider $F(x) = x^e - c_1 \pmod{n}$ and $G(x, y) = (x + y)^e - c_2 \pmod{n}$ and compute the resultant $h(y) = Res(F, G)$ in the ring $\mathbb{Z}_n[y]$; note $h$ has a root, $\delta$.

## 3.5 Lattice Methods

**LLL:** $\mathcal{F}(A) = \{\alpha_1 a_1 + \ldots + \alpha_n a_n : 0 \leq a_i \leq 1\}$ then $vol(\mathcal{F}(A) = det(A)$. Reduced basis: $\mu_{ij} < \frac{1}{2}$, $||b_i^*||^2 \leq \frac{4}{3}||b_{i+1}(i)||^2$. Let $(b_1, \ldots, b_n)$ be a reduced basis of $L$ then $||b_1|| \leq 2^{\frac{n-1}{2}}\lambda(L)$.

LLL motivation: $AU = B$ has a solution iff $M = \begin{pmatrix} I & 0 \\ A & -B \end{pmatrix}$ and $M[U, 1]^T = [U, 0]^T$ has a solution with $U$ a $0, 1$ vector. Since $||[U, 0]^T|| \leq n$, a short vector in the lattice generated by the column space of $M$ is likely to be close to a solution of $AU = B$. Let $L$ be a lattice generated by $M$, $vol(L) = |det(M)|$ Not all lattices are generated by linearly independent vectors; for example $< (1, 2), (1, 1), (2, 1) >$.

**Lattices in 2 dimensions** (vectors are columns) $[a, b]$ is reduced iff $||a|| \leq ||b||$ and $||a||, ||b|| \leq ||a + b||, ||a - b||$. Lemma: If $||x|| \leq ||x + y||$ then $||x + y|| \leq ||x + \alpha y||$, $\alpha > 1$. Let $\lambda_k = min_x|\{v \in \mathcal{L}(B) - \{0\} : ||v|| \leq x\}| \geq k$ (so $\lambda_1$ is the shortest vector in the lattice.) Theorem: If $a, b$ is a basis, $||a|| = \lambda_1$, $||b|| = \lambda_2$ iff $[a, b]$ is a reduced basis. Gauss algorithm: (1) Find $\mu$: $||b - \mu a||$ is minimal. (2) if $||a - b|| > ||a + b||$ replace $b$ with $-b$. (3) if $[a, b]$ is not reduced, swap $a$ and $b$ and go to 1. Note: LLL Gives an approximation to reduced basis $n > 2$.

Definitions: $\pi_i(x) = \sum_{j \leq i} \frac{<b_j^*, x>}{<b_j^*, b_j^*>} b_j^*$, $b_i^* = \pi(b_i)$. $\lceil x \rfloor$ is integer closest to $x$. $B = [b_1, b_2, \ldots, b_n] \in \mathbb{R}^{m \times n}$ is **LLL reduced** with respect to $\delta$ if

1. $|\mu_{i,j}| \leq \frac{1}{2}$, $i > j$.

2. $\delta||\pi(b_i)||^2 \leq ||\pi_i(b_{i+1})||^2$ (same as $\delta||b_i^*||^2 \leq ||b_{i+1}^* + \mu_{i+1,i}b_i^*||^2$).

Note: For $\delta = 1$ this just means $\pi_i[b_i, b_{i+1}]$ is reduced. Also observe that $(\delta - \frac{1}{4})||b_i^*||^2 \leq ||b_{i+1}^*||^2$. So $\lambda_1 \geq min||b_i^*|| \geq \frac{1}{\alpha}^{\frac{n-1}{2}}||b_1||$, $\alpha^{-1} = \delta - \frac{1}{4}$.

LLL single step for $\delta = \frac{1}{4}$:

1. $b_i^* = b_i - \sum_{j<i} b_j^*$ where $\mu_{i,j} = \lceil \frac{<b_i, b_j^*>}{<b_j^*, b_j^*>} \rfloor$.

2. if $\delta||\pi_i(b_i)||^2 > ||\pi_i(b_{i+1}||^2$, swap $b_i$ and $b_{i+1}$ and repeat reduction step.

Note: Terminates because $det(B)^2 \in \mathbb{Z}$ decreases by at least $\delta$ at each step.

**LLL Theorem:** Let $L \subseteq \mathbb{R}^n$ be a lattice with reduced basis $(b_1, b_2, \ldots, b_n)$ then (a) $\forall x \in L$, $||b_j|| \leq 2^{\frac{n-1}{2}}||x||$ and (b) replace $x$ by max of $t$ linearly independent vectors.

$K$ is **convex** and **symmetric** iff $x, y \in K$ implies $ax + by \in K$ provided $|a| + |b| \leq 1$. **Minkowski:** Let $L$ be a lattice of rank $r$. Let $v_1$ be the shortest vector, $v_i$ the shortest vector independent of $< v_1, \ldots, v_{i-1} >$, then $|v_1||v_2|\ldots|v_r| \leq \frac{2^r}{vol(B_r)}d(L)$ where $vol(B_r) = \frac{\pi^{\frac{r}{2}}}{\Gamma(1+\frac{r}{2}))}$.

**Minkowski's theorem on linear forms:** Let $\Lambda \in \mathbb{R}^N$ and $L_1, \ldots, L_N$ be linear forms with associated matrix $C$; if $det(C)d(\lambda) \leq \epsilon_1\epsilon_2\ldots\epsilon_N$, there is a lattice point $\lambda \neq 0$ such that $|L_m(\lambda)| \leq \epsilon_m$. Corollary: $\exists l : L_m(l) \leq (det(C))^{\frac{1}{N}}$.

Low density subset sum. $\sum a_i s_i = s$ look at matrix formed by $I_n$ with bottom row $(\frac{1}{2}, \ldots, \frac{1}{2}, ms)$ and first n entries in rightmost columns $(ma_1, ma_2, \ldots, ma_n)$. Round.

**Weakness due to partial knowledge:** If $n = pq$ has $m$ bits and we know the first or last $\frac{m}{4}$ bits of $p$, then $n$ is easy to factor. If plaintext is short, match $cx^{-e} = y^e$ to get $c = (xy)^e \pmod{n}$. If $q < p < 2q$ and $1 \leq d, e < \psi(n)$ with $de = 1 \pmod{\psi(n)}$ and $d < \frac{1}{3}n^{\frac{1}{4}}$ then d can be found easily.

**Attack on RSA using LLL:** Suppose message is of the form "M xxx" where only 'xxx' varies (e.g.- "The key is xxx"). Thus the message is of the form $B + x$ where $B$ is fixed and $|x| < Y$. $c = (B + x)^3 \pmod{n}$ and $f(T) = (B + T)^3 - c = T^3 + a_2T^2 + a_1T + a_0 \pmod{n}$. We want to find $x : f(x) = 0 \pmod{n}$. Let $v_1 = (n, 0, 0, 0)$, $v_2 = (0, Yn, 0, 0)$, $v_3 = (0, 0, Y^2n, 0)$, $v_4 = (a_0, a_1Y, a_2Y^2, Y^3)$. Then $||b_1|| \leq 2^{\frac{3}{4}}|det(v_1, v_2, v_3, v_4)| = 2^{\frac{3}{4}}n^{\frac{3}{4}}Y^{\frac{3}{2}}$. $b_1 = c_1v_1 + \ldots + c_4v_4 = (e_0, Ye_1, Y^2e_2, Y^3e_3)$; $e_0 = c_1n + c_4a_0$, $e_1 = c_2n + c_4a_1$, $e_2 = c_3n + c_4a_2$, $e_3 = c_4$, and $g(T) = e_3T^3 + e_2T^2 + e_1T + e_0$. Since $f(x) = 0 \pmod{n}$ and $c_4f(T) = g(T) \pmod{n}$, $0 = c_4f(x) = g(x) \pmod{n}$. If $Y < 2^{\frac{7}{6}}n^{\frac{1}{6}}$, $|g(x)| \leq 2||b_1||$ (use C-S) but $||b_1|| \leq 2^{-1}n$ so $|g(x)| < n$ and $g(x) = 0$ yielding 3 candidates for $x$. **Coppersmith** extended this to small solutions of polynomials of degree $d$ using a $d + 1$ dimensional lattice by examining the monic polynomial $f(T) = 0 \pmod{n}$ of degree $d$ when $|x| \leq n^{\frac{1}{d}}$.

## 3.6 Symmetric Key Analysis

**DES S Box Criteria:** (1) $S$ is not linear or affine in the inputs, (2) changing 1 bit of input changes at least 2 bits of output, (3) minimize differences between 1s and 0s if one input bit is held constant, (4)$Ham(S(x) \oplus S(x \oplus 001100)) > 1$, and (5) $S(x) \neq S(x \oplus 11ab00)$.

**Differential cryptanalysis:** Notation: $x \rightarrow y, p$ means input difference $x$ produces output $y$ with probability $p$. If $x' \rightarrow y'$ and $D_j(x', y') = \{u : S_j(u) \oplus S_j(u \oplus x') = y'\}$ then $x \oplus k \in D_j(x', y')$, and $k \in D_j(x', y') \oplus x$. Set $\tau_j(x, x', y') = \{k : k \in D_j(x', y') \oplus x\}$ and $test_j(E_j, {E_j}^*, C_j') = \tau_j(E_j, E_j \oplus E_j^*, C_j')$. Note: some candidate keys will scritch. To convert from chosen to known attack, select $2^{32}\sqrt{2m}$ pairs, about $m$ of these will have the right difference $x$ produces output $y$ with probability $p$. If $x' \rightarrow y'$ and $D_j(x', y') = \{u : S_j(u) \oplus S_j(u \oplus x') = y'\}$ then $x \oplus k \in D_j(x', y')$, and $k \in D_j(x', y') \oplus x$. Set $\tau_j(x, x', y') = \{k : k \in D_j(x', y') \oplus x\}$ and $test_j(E_j, {E_j}^*, C_j') = \tau_j(E_j, E_j \oplus E_j^*, C_j')$.

3-round Attack: $(L_0, R_0)$, $R_3 = L_2 + f(R_2, k_3) = L_0 + f(R_0, k_1) + f(R_2, k_3)$. Choose $R_0{}' = 000000$, so that $f(R_0, k_1) + f(R_0^*, k_1) = 0$, get $R_3' = L_0' + f(R_2, k_3) + f(R_2^*, k_3)$. Set $C' = P^{-1}(R_3' + L_0')$ which is the output xor for round 3. Compute $E = E(L_3)$, $E^* = E(L_3^*)$. Calculate $test_j(E_j, E_j{}^*, C_j')$, for j= 1,2,...,8 after choosing plaintexts. Can do this since $R_2 = L_3$ is known. Not that key bits overlap on initial and final rounds and must satisfy both conditions. Best differential cryptanalysis attack on DES uses two 13 round differentials in a 2R attack.

6-round Attack: Use $L_0', R_0'$: 0x40080000, 0x40000000, $L_1', R_1'$: 0x40000000, 0x00000000, $p = .25$; $L_2', R_2'$: 0x00000000, 0x40000000, $p = 1$; $L_3', R_3'$: 0x40000000, 0x40080000, $p = .25$.
$R_6 = L_5 + f(R_5, K_6) = R_4 + f(R_5, K_6) = L_3 + f(R_3, K_4) + f(R_5, K_6)$. Estimate $L_3' = 0x04000000$ and $R_3' = 0x40080000$ with $p = \frac{1}{16}$. Use this to estimate input xor for S-boxes of round 4. Get $C_1'C_2'...C_8' = P^{-1}(R_6' + 0x04000000)$ and $E_1'E_2'...E_8' = E(R_5) = E(L_6)$. Now compute $test_j(E_j, E_j{}^*, C_j')$, for j= 2,5,6,7,8. Right pair follows characteristic. Right pairs bump count for correct key bits, wrong pairs are random. Filter: If $|test_j(E_j, E_j{}^*, C_j')| = 0$, for any j= 2,5,6,7,8, this is a wrong pair. $\frac{2}{3}$ of the wrong pairs are detected this way, so ratio of right pairs remaining is $\frac{\frac{1}{16}}{\frac{1}{16}+\frac{15}{16}\times\frac{1}{3}} = \frac{1}{6}$. Number of suggested pairs is $\Pi|test_j(E_j, E_j{}^*, C_j')|$, for j= 2,5,6,7,8, correct values will be suggested $\frac{3n}{16}$ times; incorrect strings at random among approx $2^{30}$ values. Let $T_j$ be the counter vector of length 64. For each pair compute $T_j^i$, j= 2,5,6,7,8, $1 \leq i \leq n$. For $I \subseteq \{1, 2, \ldots n\}$, $\sum_{i \in I} T_j^i$. There should be some $I$ of size about $\frac{3n}{16}$ where all of the indexes have 1 in the vector. This is the suggested key.

Another 3-Round Characteristic: $L_0', R_0'$: 0x00200008, 0x00000400, $L_1', R_1'$: 0x00000400, 0x00000000, $p = .25$; $L_2', R_2'$: 0x00000000, 0x00000400, $p = 1$; $L_3', R_3'$: 0x00000400, 0x00200008, $p = .25$.

**Linear cryptanalysis:** $\alpha \cdot P + \beta \cdot C = \gamma \cdot C$ with $p = \frac{1}{2} + \delta$ requires about $c\delta^{-2}$ plaintexts. Last round estimation: $L(P) + M(C) + N(P_{n-1}, K_n) = P(K)$ then use MLE: $T=$ # plain cipher pairs = 0 if $|T_{max} - \frac{N}{2}| > |\frac{N}{2} - T_{min}|$ and $p > .5$, guess $P(K) = 0$. Best demonstrated effect on DES is $2^{42.6}$.

3-round: $P_L[7, 18, 24, 29] + C_L[7, 18, 24, 29] + P_R[15] + C_L[15] = K_1[22] + K_3[22]$.
8-round: $P_L[7, 18, 24] + P_R[12, 16] + C_R[7, 18, 24, 29] + C_L[15] + F_8(C_R, K_8) = K_1[19, 23] + K_3[22] + K_4[44] + K_5[22] + K_7[22]$.

**Gradual exercises:** Analyze 8 round RC5 with no rotation, 8 round RC5 with rotation equal to round number, 12-round DES with no S-box, 4 round DES, 6 Round DES. Best Linear attack on DES uses a 14 round differential with bias $2^{-21.75}$ forward and reverse and uses $2^{43}$ corresponding pairs with .85 probability of success.

An encryption scheme, E, is **semantically secure** if $\forall A, \exists B$ such that $\forall f, h, f, h : \{0,1\}^* \rightarrow \{0,1\}$ and all ensembles $\{X_n\}$ where $X_n$ $\{0,1\}^{n^2}$, $Pr[A(E(X_n), h(X_n)) = f(X_n)] < Pr[B(h(X_n)) = f(X_n))] + \mu(n)$ where $\mu$ is negligible. A deterministic PT algorithm G is pseudo-random if $\exists l : N \rightarrow N$, so that for any probabilistic PT algorithm D, and any positive polynomial P and all sufficiently large k, $|Pr[D[G(U_k)) = 1] - Pr[D(U_{l(n)}) = 1]| < \frac{1}{p(k)}$. ($l$ is a stretching function.)

A **linear trail** is $U = (u^{(0)}, u^{(1)}, \ldots, u^{(r)})$ associated with a composite function $\beta = \rho^{(r)}\rho^{(r-1)} \ldots \rho^{(1)}$ with correlation contribution at each step of $C((u^{(i)})^T \rho^{(i)}(a), u^{(i-1)}a)$ and overall correlation of $C_p(U) = \prod_i C^{(\rho^{(i)})}_{u^{(i)}, u^{(i-1)}}$.

**Theorem:** $C(u^T\beta(a), w^Ta) = \sum_{U, u^{(0)}=u, u^{(r)}=w} C_p(U)$.

**Some rules for Walsh transforms:** $V_f = \{w : f(w) \neq 0\}$ is called the support for $f$. Computing Walsh transforms of composite functions is easier if the components have non-intersecting support (as they do it they depend on different variables); in that case, $V_f \cap V_g = \emptyset$ trivially. If $w \in V_f$ and $h(x) = g(x) + w^Tx$, $H(u) = G(w \oplus u)$. If $V_f \cap V_g = \emptyset$ then $u \in V_g$, $H(u + w) = F(w)G(u)$.

**"Bricklayer" functions:** If

$$h(a(1), \ldots, a(n)) = (h(1)(a(1), a(2), \ldots, a(n)), h(2)(a(1), a(2), \ldots, a(n)), \ldots, h(n)(a(1), a(2), \ldots, a(n))),$$

then $C_{uv}^{(h)} = \prod_{i=1}^{n} C_{u(i),v(i)}^{h(i)}$. **Truncating Function:** Let $a' = h^{(r)}(a)$ taking $GF(2)^{n-1} \rightarrow GF(2)^n$ be defined by $a'_i = a_i$ for $i \neq s$ and $a'_s = \epsilon \oplus v^t a \oplus a_s$ where $v^T a = \epsilon$ defined the restriction. Then $C_{w,w}^{(h^r)} = 1$, $C_{v \oplus w,w}^{(h^r)} = (-1)^{\epsilon}$, $\forall w : w_s = 0$; note there are two non-zero entries both of amplitude 1. If $C' = CC^{(h^{(r)}}$, $C'_{u,w} = C_{u,w} \oplus (-1)^{\epsilon} C_{u,v\oplus w}$ if $w_s = 1$ and 0 if $w_s = 0$.

For **key alternating ciphers**, $C_p(U) = \prod_i (-1)^{(u^{(i)})^T k^{(i)}} C_{u^{(i)},u^{(i-1)}} = (-1)^{d_U \oplus \bigoplus_i (u^{(i)})^T k^{(i)}} |C_p(U)|$. Put $s_i = U^T K \oplus d_U$, $C(v^T \cdot \beta(a), w^T a) = \sum_{U,u^{(0)}=u,u^{(r)}=w} (-1)^{d_U \oplus U^T K} |C_p(U)|$. $C_p(U) = (-1)^{s_i} C_i$, averaging over the round keys we get $E(C_t^2) = 2^{-n_K} \sum_k (\sum_i (-1)^{s_i} C_i)^2$. After reduction, average correlation potential is $E(C_t^2) = \sum_i C_i^2$, note that $C_i C_j = 2^{n_K} \delta(i \oplus j)$.

For key schedule $K = M_\kappa k$, $E(C_t^2) = 2^{-n_K} \sum_i \sum_j (\sum_k (-1)^{(d_{U_i} \oplus d_{U_j})^T M_\kappa k \oplus d_{U_i} \oplus d_{U_j}}) C_i C_j$. The inner sum simplifies to $(-1)^{d_{U_i} \oplus d_{U_j}} 2^{n_K} \delta(M_\kappa^T (U_i \oplus U_j))$. If key schedule is not linear $K = f_\kappa(k)$, the coefficient of the mixed term is $(-1)^{(U_i \oplus U_j)^T f_\kappa(k) \oplus d_{U_i} \oplus d_{U_j}}$.

Multiround linear expressions correspond to linear trails. Generally, $|C_p(U)|$ is independent of round key but this is not the case in DES because of the shared bits between S-boxes. 32 bit input parities before $E$ give rise to $\alpha$ $2^{2l}$-48 bit patterns. If $l$ is the number of pairwise neighboring S-boxes, we can do this in$16l$ multiplications and additions. The probability that a multiround expression holds is $\frac{1}{2}(1 + C_p(U))$ for the associated trail.

**Question:** Is there an easy to compute function, $T_K$, obviously non-linear, so that $T_K E_K T_K^{-1}$ has good linear approximations? How do you find such $T_K$? Finding the best approximation reduces to finding an orthogonal transformation that maximizes the largest entry. Suppose $T$ is such a matrix; if $T$ has all bad affine approximations is it possible that there is another orthogonal transformation, $R$ with $T^R = R^{-1} T R$ such that $max_{ij}(|(T^R)_{ij}|) > max_{ij}(|(T)_{ij}|)$? If $\rho_1, \rho_2, \ldots, \rho_n$ is a series of such transformations (like the iterated components of a block cipher), note that $R^{-1} E_K(x) R = R^{-1} \rho_1 R R^{-1} \rho_2 R \ldots R^{-1} \rho_n R$ thus raising the possibility of better "per round" approximations on a related cipher.

Here is a motivating example in $\mathbb{R}^3$: $R = \begin{pmatrix} cos(\varphi) & sin(\varphi) & 0 \\ -sin(\varphi) & cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & sin(\theta) \\ 0 & -sin(\theta) & cos(\theta) \end{pmatrix}$ and

$$R^{-1}TR = \begin{pmatrix} cos^2(\varphi) + cos(\theta) sin^2(\varphi) & cos(\varphi) sin(\varphi) - cos(\theta) cos(\varphi) sin(\varphi) & -sin(\varphi) sin(\theta) \\ -cos(\varphi) sin(\varphi) + cos(\theta) cos(\varphi) sin(\varphi) & sin^2(\varphi) + cos(\theta) cos^2(\varphi) & sin(\varphi) sin(\theta) \\ sin(\varphi) sin(\theta) & -cos(\varphi) sin(\theta) & cos(\theta) \end{pmatrix}$$

$NL(f) \leq 2^{n-1} - 2^{\frac{n}{2}-1}$, $NL(f) \leq 2^{n-1} + \sqrt{2^n + max_{e \neq 0}(F(D_e(f)))}$, where $D_e f = f(x) \oplus f(x \oplus e)$.

**Prolog to computing DES correlation matrix:** Let $f(x_1, x_2, x_3, x_4) = (x_1 + f_1(x_3, x_4), x_2 + f_2(x_3, x_4), x_3, x_4)$ (first position most significant) then, with least significant positions indexing rows and columns, and $F_i(w)$ as the Walsh transform for $f_i(x_3, x_4)$ and $H(w)$ the Walsh transform of $h(x) = f_1(x) + f_2(x)$. Bit positions in this example are $(x_1, x_2, x_3, x_4)$.

$$C^{(f)} = \left(\begin{array}{cccc|cccc|cccc|cccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & F_2(0) & F_2(1) & F_2(2) & F_2(3) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & F_2(1) & F_2(0) & F_2(3) & F_2(2) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & F_2(2) & F_2(3) & F_2(0) & F_2(1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & F_2(3) & F_2(2) & F_2(1) & F_2(0) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F_1(0) & F_1(1) & F_1(2) & F_1(3) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F_1(1) & F_1(0) & F_1(3) & F_1(2) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F_1(2) & F_1(3) & F_1(0) & F_1(1) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & F_1(3) & F_1(2) & F_1(1) & F_1(0) & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H(0) & H(1) & H(2) & H(3) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H(1) & H(0) & H(3) & H(2) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H(2) & H(3) & H(0) & H(1) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & H(3) & H(2) & H(1) & H(0)
\end{array}\right)$$

**Feistel:** A typical round of DES consists of two involutions: $\tau$ and $\sigma_k$. $\sigma_k(L, R) = (L \oplus f(R, k), R)$, $f(x, k) = PS_1S_2 \dots S_8(E(x) + k))$. $\tau(L, R) = (R, L)$. First "line" of $\sigma_k$ is $y_9 = x_9 \oplus S_1^1(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$, $y_{17} = x_{17} \oplus S_1^2(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$, $y_{23} = x_{23} \oplus S_1^3(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$, $y_{31} = x_{31} \oplus S_1^4(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$.

Suppose $\tau(x_1, x_2, x_3, x_4) = (x_3, x_4, x_1, x_2)$, with position (0001) representing $x_4$, then

$$C^{(\tau)} = \left(\begin{array}{cccccccc|cccccccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{array}\right).$$

The column order from left to right in the forgoing is: 1, $(x_4)$, $(x_3)$, $(x_4, x_3)$, $(x_2)$, $(x_4, x_2)$, $(x_3, x_2)$, $(x_4, x_3, x_2)$, $(x_1)$, $(x_4, x_1)$, $(x_3, x_1)$, $(x_4, x_3, x_1)$, $(x_2, x_1)$, $(x_4, x_2, x_1)$, $(x_3, x_2, x_1)$, $(x_4, x_3, x_2, x_1)$ corresponding to the ordered sequence $0000, 0001, 0010, \dots$. The row order from top to bottom is 1, $(x_2)$, $(x_1)$, $(x_2, x_1)$, $(x_4)$, $(x_4, x_2)$, $(x_4, x_1)$, $(x_4, x_2, x_1)$, $(x_3)$, $(x_3, x_1)$, $(x_3, x_2)$, $(x_3, x_2, x_1)$, $(x_3, x_4)$, $(x_3, x_4, x_2)$, $(x_3, x_4, x_2)$, $(x_3, x_4, x_1)$, $(x_3, x_4, x_2, x_1)$.

**Todo:** Best Approximation of degree two. Correlation of decomposed function $(g(x_1, x_2, \dots, x_k, h(x_{k+1}, \dots, x_n)))$. Minimum distance.

**Standard Functions:** For $h(x) = x \oplus k$, $C^{(h)}_{u,u} = (-1)^{u^T \cdot k}$. For $h(x) = Mx \oplus w$, $C^{(h)}_{u,w} = \delta(M^T u \oplus w)$. $\hat{c}_{fg} = 2^{-n} \sum_a (-1)^{\hat{f}(a)\hat{g}(a+b)}$, $\hat{r}_f = \hat{c}_{ff}$.

**Theorem:** All correlation matrices are doubly stochastic and orthogonal. Correlation matrices for involutions are symmetric.

To calculate the round correlation for DES, decompose it into three involutions. The first, adds output from odd numbered S-boxes but is otherwise the identity. The second, adds output from even numbered S-boxes but is otherwise the identity. The third transposes $L$ and $R$. The first and second involutions don't overlap on input variables to the SBoxes so the Walsh transforms of components of the S-Boxes are all that is needed. In both the first and second transformations, each position affected by an S-box is multiplied by $(-1)^{w^T \cdot k}$ (i.e. -$\pm 1$) for the relevant round keys. Thus, if $\sigma_k(L, R) = (L \oplus f(R, k), R)$, $f(x, k) = PS_1 S_2 \ldots S_8(E(x) + k))$, the first "line" is $y_9 = x_9 \oplus S_1^1(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$, $y_{17} = x_{17} \oplus S_1^2(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$, $y_{23} = x_{23} \oplus S_1^3(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$, $y_{31} = x_{31} \oplus S_1^4(x_{64} + k_1, x_{33} + k_2, x_{34} + k_2, x_{35} + k_2, x_{36} + k_2, x_{37} + k_2)$. $Tr(C^{(AES)})$ is the number of fixed points of AES. Since $Tr(AB) = Tr(BA)$, $Tr(C^{(AES)}) = Tr(C^{(k_{14})} C^{(k_{13})} \ldots C^{(k_1)} C^{(RS)} (C^{(MRS)})^{13})$.

The difference propagation probability denoted by $R_p(a' \rightarrow_h b')$ is $Prob^h(a', b') = 2^{-n} \sum_a \delta(b' + h(a + a') + h(a))$; we have $0 \leq R_p(a' \rightarrow_h b') \leq 1$. The restriction weight is defined as $w_r(a' \rightarrow_h b') = -lg(R_p(a' \rightarrow_h b'))$ (restriction weight reflect loss of entropy). $w_c(U) = -lg(|C_p(U)|)$ (correlation weight). For bricklayer function, $Prob^h(a', b') = \prod_i Prob^{h_{(i)}}(a'_{(i)}, b'_{(i)})$ and $w_r(a', b') = \sum_i w_r(a'_{(i)}, b'_{(i)})$.

**Theorem:** $Prob^f(a', 0) = \frac{1}{2}(1 + \sum_w (-1)^{w^T a'} F(w)^2)$. The differential probability and correlation potential table of a boolean function satisfy $Prob(a', b') = 2^{-m} \sum_{u,w} (-1)^{w^T a' \oplus u^T b'} C_{u,w}^2$.

A **differential trail**, $Q = (q^{(0)}, q^{(1)}, \ldots, q^{(r)})$ with steps $(q^{(i-1)}, q^{(i)})$ having weight $w_r^{\rho^{(i)}}(q^{(i-1)}, q^{(i)})$ have trail weight $w_r(Q) = \sum_i w_r^{\rho^{(i)}}(q^{(i-1)}, q^{(i)})$. $Prob(a', b') = \sum_{q^{(0)} = a', q^{(r)} = b'} Prob(Q)$. For a differential trail, $Q$, with weight $< (n-1)$, $Prob(Q) \approx 2^{-w_r(Q)}$. For a differential trail, $Q$, with weight $w_r(Q) > (n-1)$, for expected proportion $2^{n-1-w_r(Q)}$ of keys, there will be a right pair. $\sum_{b'} R_p(a' \rightarrow_h b') = 1$. $R_p(a' \rightarrow_h b') = 2^{-n} \sum_{u,w} (-1)^{wa' + ub'} (C_{u,w})^2$ and dually $C_{u,w}^2 = 2^{-n} \sum_{a',b'} (-1)^{wa' + ub'} R_p(a' \rightarrow_h b')$.

**Block cipher design:** To eliminate low weight trails, there are two strategies: (1) Choose S-boxes with difference propagations that have high restriction weight and input-output correlations with high correlation weights; or, (2) Design round transformations so that only trails with many S-boxes occur. Linear cryptanalysis requires correlation $> 2^{-\frac{n_b}{2}}$ over most rounds. This can't happen if we choose the number of rounds so that there are no such linear trails with correlation contribution $> n_k^{-1} 2^{-\frac{n_b}{2}}$ Each output parity is correlated to an input parity since $\sum_w F(w)^2 = 1$ but if it occurs by constructive interference over many trails that share input/output selection then any such must be the result of at least $n_k$ linear trails which are unlikely to be key dependent. Differential cryptanalysis requires input to output difference propagation with probability $> 2^{1-n_b}$. If there are no differential trails with low weight, difference propagation results from multiple trails which again will not likely be key dependent.

**Design strategy for Rijndael:** Choose number of rounds so that there is no correlation over all but a few rounds with amplitude significantly larger than $2^{-\frac{n_b}{2}}$ by insuring there are no linear trails with correlation contribution above ${n_k}^{-1} 2^{-\frac{n_b}{2}}$ and no differential trails with weight below $n_b$.

Examine round transformations $\rho = \lambda \circ \gamma$, where $\lambda$ is the mixing function and $\gamma$ is a bricklayer function that acts on bundles of $n_t$ bits. Block size is $n_b = mn_t$. The correlation over $\gamma$ is the product of correlations over different S-box positions for given input and output patterns. Define weight of correlation as $-lg(Amplitude)$. If output selection pattern is $\neq 0$, the S-box is active. Looking for maximum amplitude of correlations and maximum difference propagation probability. The weight of a trail is the sum of the weights of the selection patterns or the sum of the active S-box positions it is greater than the number of active S-boxes times the minimum correlation weight per S-box. Wide trail: design round transformations so there are no trails with low bundle weight.

Define $w_b(a)$ as the bundle weight of $a$. $\mathcal{B}_d(\phi) = min_{a, b \neq a}(w_b(a \oplus b) + w_b(\phi(a) \oplus \phi(b)))$. $\mathcal{B}_l(\phi, \alpha) = min_{\alpha, \beta, C(\alpha^T x, \beta^T \phi(x)) \neq 0}(w_b(\alpha) + w_b(\beta))$. **Theorem:** In an alternating key block cipher with $\gamma\lambda$ round functions, the number of active bundles in a two round trail is $\geq$ the bundle branch number of $\lambda$. If $\psi = \gamma \Theta \gamma \lambda$ is a four round function, $\mathcal{B}(\psi) \geq \mathcal{B}(\lambda) \times \mathcal{B}^c(\Theta)$ where $\mathcal{B}$ can be either the linear or differential branch number. The linear and differential branch numbers for an AES round is 5.

**Linearized polynomial:** $L(x) = \sum_{i=0}^{t} \beta_i x^{2^i}, \beta \in GF(2^n)$. **Discrete Fourier Transform:** $A_k = \sum [f(x) + f(0)] x^{-k}$, $f(x) = \sum A_k x^k$. $A_{2^i k} = A_k^{2^i}$. Coset leaders: $C_s = \{s, 2s, 2^2 s, \ldots, 2^{n_s - 1} s\}$, coset leader $s$ is smallest: $s = s2^{n_s - 1} \pmod{2^n - 1}$. For any non-zero function $f : GF(2^n) \rightarrow GF(2)$ can be represented as $f(x) = \sum_{k \in \Gamma(n)} Tr_1^{n_k}(A_k x^k) + A_{2^n - 1} x^{2^n - 1}$ where $\Gamma(n)$ are the coset leaders $\pmod{2^n - 1}$, $n_k \mid n$ and $Tr_1^{n_k}(x)$ is the trace function from $GF(2^{n_k}) \rightarrow GF(2)$. Let $\alpha$ be a primitive element of $GF(2^n)$ and $f(0) = 0$ with $a_t = f(\alpha^t), t = 0, 1, 2, \ldots, 2^n - 1$, $x = x_0 + x_1\alpha + x_2\alpha^2 + \ldots + x_{n-1}\alpha^{n-1}$.

Any function $f : GF(2^{n_k}) \rightarrow GF(2)$ corresponds to a binary sequence with period $N \mid 2^n - 1$; TBD—what is $k$. **Hadamard-Walsh:** $\hat{f}(\lambda) = \sum_{x \in GF(2^n)} (-1)^{Tr(\lambda \cdot x) + f(x)}$. Polynomials $\rightarrow_{eval}$ Periodic sequences $\leftrightarrow_{trace}$ Boolean Functions.

Low degree approximations $\exists g \neq 0 : fg = 0$ and $fg$ has low degree $deg(fg) \geq deg(f)$. $|S_d| = \sum_{i=0}^{d} \binom{n}{i}$.

Let $f$ be a boolean function of $n$ variables. The annihilator ideal of $f$, $AN(f) = \{g : g(x)f(x) = 0\}, \forall x \in GF(2^n)$, $AN_d(f) = \{g \in AN(f) : deg(g(x)) \leq d\}$. The algebraic immunity, $AI(f)$ is the smallest degree non-zero polynomial in $AN(f) \cup AN(1 + f)$. $AI(f) \leq \lceil \frac{n}{2} \rceil$.

Suppose $\mathcal{L}$ is an $n$-bit NLFSR based filter generator with filter function $f$ and that $L$ takes the current $n$-bit state to the next $n$-bit state. Suppose the initial state is $\vec{x_0}$. Then the generated keystream is $s_t = f \circ L^t(\vec{x_0})$. $s_t = 1$ if $\exists g \in AN_d(f) : g \circ L^t(\vec{x_0}) = 0$, $s_t = 0$ if $\exists h \in AN_d(1 + f) : h \circ L^t(\vec{x_0}) = 0$. Collect all functions of degree $\leq d$ for $N$ known keystream bits; then, (1) $g \circ L^t(x_1, x_2, \ldots, x_n) : \forall g \in AN_d(f), \forall 0 \leq t < N : s_t = 1$; and, (2) $h \circ L^t(x_1, x_2, \ldots, x_n) : \forall g \in AN_d(1 + f), \forall 0 \leq t < N : s_t = 0$. Using linearization to solve these equations, requires identifying the subset of monomials forming a linear system of up to $\sum_{i=1}^{d} \binom{n}{i}$ variables. Gaussian reduction on this system takes time $O((\sum_{i=1}^{d} \binom{n}{i})^{\omega}) \approx n^{\omega d}$ where $\omega \approx 2.37$ and the the number of monomials is $\approx \frac{2n^d}{d!(dim(AN_d(f)) + dim(AN_d(1+f)))}$.

**Akelarre:** Akelare; Rounds $0 \leq R < R$. $(B_0, B_1, B_2, B_3) = (A_0, A_1, A_2, A_3) <<< K_{13r+4}[25, 26, \ldots, 31]$. Initial Prep: $I_j = X_j = K_j$. Round $r$: $(I_0', I_1', I_2', I_3') = (I_0, I_1, I_2, I_3) <<< K_{13r+4}[25, 26, \ldots, 31]$. $A_R(I_0' \oplus I_2', I_1' \oplus I_3') = a_L || a_R$. $O_0 = I_0' \oplus a_R$, $O_1 = I_1' \oplus a_L$, $O_2 = I_2' \oplus a_R$, $O_3 = I_3' \oplus a_L$. Final Out: $Y_j = I_j' + K_{13R+5+j}$. Todo: describe $A_R$.

**FEAL-4:** 32 bit blocks, 64 bit keys. Four round Feistel with input/output whitening. Key, $K$, is used to generate 12 16-bit keys $K_0, K_1, \ldots, K_{11}$. To define the key schedule and the round function $F$ put $G_0(a, b) = (a + b \pmod{256}) <<< 2$, $G_1(a, b) = (a + b + 1 \pmod{256}) <<< 2$. Key Schedule: Define $f_K : \mathbb{Z}_2^{32} \times \mathbb{Z}_2^{32} \rightarrow \mathbb{Z}_2^{32}$ as follows: $f_K(a, b) = c$, $a = a_0||a_1||a_2||a_3$, $b = b_0||b_1||b_2||b_3$, $c = c_0||c_1||c_2||c_3$, then $d_1 = a_0 \oplus a_1$, $d_2 = a_2 \oplus a_3$, $c_1 = G_1(d_1, a_2 \oplus b_0)$, $c_2 = G_0(d_2, c_1 \oplus b_1)$, $c_0 = G_0(a_0, c_1 \oplus b_2)$, $c_3 = G_1(a_3, c_2 \oplus b_3)$. Then put $B_{-2} = 0$, $B_{-1} = K_L$, $B_0 = K_R$, and $B_{i+1} = f_K(B_{i-2}, B_{i-1} \oplus B_{i-3}$, $K_{2(i-1)} = (B_i)_L$, $K_{2i-1} = (B_i)_R$. Encryption: If $P_L, P_R$ is the cipher input and $C_L, C_R$ is the cipher output, $L_0 = P_L \oplus (K_4||K_5)$ and $R_0 = L_0 \oplus P_R \oplus (K_6||K_7)$. Each round is defined as: $R_{i+1} = L_i \oplus F((K_{2(i-1)}||K_{2i-1}) \oplus R_i)$ and $L_{i+1} = R_i$. $F$ is defined by: $F(x_0, x_1, x_2, x_3) = (y_0, y_1, y_2, y_3)$ where $y_1 = G_1(x_0 \oplus x_1, x_2 \oplus x_3)$, $y_0 = G_0(x_0, y_1)$, $y_2 = G_0(y_1, x_2 \oplus x_3)$, and $y_3 = G_1(y_2, x_3)$. Finally, $C_L = L_4 \oplus (K_8||K_9), C_R = R_4 \oplus L_4 \oplus (K_{10}||K_{11})$. Note that $A_0 \oplus A_1 = 0x80800000 \rightarrow F(A_0) \oplus F(A_1) = 0x02000000$. For differential attack, pick $P_L$ at random and $P_1 = 0x8080000080800000$. Suppose $X'$ is the output differential of $F$ in round 3, $Y'$ is the input differential to $F$ in round 4 and $Z'$ is the output differential in Round 4, then $C_L' = 0x02000000 \oplus Z'$ and $C_R' = C_L' \oplus Y'$ and $Y = C_L \oplus C_R$. Now we can solve for $K_3$ with standard differential techniques. For linear analysis, denote $S_{i,j}(X) = x_i \oplus x_j$, $S_i(X) = x_i$. Then, $S_5(G_0(a, b)) = S_7(a \oplus b)$ and $S_5(G_0(a, b)) = S_7(a \oplus b) \oplus 1$. The following hold: $S_{13}(Y) = S_{7,15,23,31}(X) \oplus 1$, $S_5(Y) = S_{15}(Y) \oplus S_7(X)$, $S_{15}(Y) = S_{21}(Y) \oplus S_{23,31}(X)$, $S_{23}(Y) = S_{29}(Y) \oplus S_{31}(X) \oplus 1$ and $a = S_{23,29}(P_L \oplus P_R \oplus C_L) \oplus S_{31}(P_L \oplus C_L \oplus C_R) \oplus S_{30}(P_L \oplus C_L \oplus K_0)$.

**WEP Attack:** WEP is data level encryption using a long term secret $K$ and per message initial vector, $IV$ which is 3 bytes which we call $K_0, K_1, K_3$. The IV and the key bytes $K_3, ...$ form a single RC4 key $K_0, K_1, K_2, K_3, \ldots$. Attack involve selecting $IV = 3|255|V$. The RC4 initialization at $i = 0$ step is $j = j + S_0 + 255 = 3 \pmod{256}$ then swap $S[0], S[3]$; this leaves S:

| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | ... |
|---|---|---|---|---|---|---|---|---|---|
| S[i] | 3 | 1 | 2 | 0 | 4 | 5 | 6 | 7 | ... |

The $i = 1$ step is $j = j + S_1 + K_1 = 3 + 1 + 255 \pmod{256} = 3$; this leaves S:

| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | ... |
|---|---|---|---|---|---|---|---|---|---|
| S[i] | 3 | 0 | 2 | 1 | 4 | 5 | 6 | 7 | ... |

The $i = 2$ step is $j = j + S_2 + K_2 = 5 + V \pmod{256} = 3$; this leaves S:

| i | 0 | 1 | 2 | 3 | 4 | ... | 5+V | ... | ... |
|---|---|---|---|---|---|---|---|---|---|
| S[i] | 3 | 0 | 5+V | 1 | 4 | ... | 2 | ... | ... |

Finally, at $i = 3$ step is $j = j + S_3 + K_3 = 5 + V + S_3 + k + 3 \pmod{256} = 6 + V + K_3$; this leaves S:

| i | 0 | 1 | 2 | 3 | 4 | ... | 5+V | ... | 6+V+K[3] | ... |
|---|---|---|---|---|---|---|---|---|---|---|
| S[i] | 3 | 0 | 5+V | 6+V+K[3] | 4 | ... | 2 | ... | 1 | ... |

$Stream[0] = S[3] = 6 + V + K_3$ if initialization stops here. Attack works if $S[0], S[1], S[2]$ don't change. The probability of this is $\frac{253}{256}^{255} \approx .0513$.

$\Delta^{\otimes} X = X \otimes X^{-1}$. $r-$round characteristic: sequence of differences $< \alpha_0, \alpha_1, \ldots, \alpha_r >$. Definition (Lai): An iterated cipher is called a Markov cipher if $Pr(\Delta C_1 = \beta | \Delta C_0 = \alpha, C_0 = \gamma)$ is independent of $\gamma, \forall \alpha, \beta \neq e$. **Homogeneous Markov Chain:** $Pr(v_{i+1} | v_i = \alpha)$ is independent of $i, \forall \alpha, \beta$.

If an $r-$round iterated cipher is a Markov and the $r$ round keys are independent and uniformly distributed then $\Delta P = \Delta C_0, \Delta C_1, \ldots, \Delta C_r$ is a homogeneous Markov chain and $Pr(\Delta C_s = \alpha_s | \ldots | \Delta C_1 = \alpha_1 | \Delta P = \alpha_0) = \prod Pr(\Delta C_i | \Delta C_{i-1})$.

**Differentials:** Right pair follows differential. Assume $m$ pairs of chosen text, $p$ is probability of characteristic, $k$ is the number of keys, $\gamma$ is number of suggested keys. There are about $mp$ right pairs. If $\lambda$ is the ratio of non-discarded pairs to the number of discarded pairs, wrong key is suggested $\frac{m\gamma\lambda}{k}$. $S/N = \frac{mp}{\frac{m\gamma\lambda}{k}} = \frac{kp}{\gamma\lambda}$. For DC to succeed, $S/N > 1$. $\Delta_\alpha f(x) = f(x + \alpha) - f(x)$. $\Delta_{\alpha_1, \ldots, \alpha_i} = \Delta_{\alpha_i}(\Delta_{\alpha_1, \ldots, \alpha_{i-1}})$. If $a_i$ is linear independent of $a_1, \ldots, a_{i-1}$, $\delta_{a_1, a_2, \ldots, a_i} f(x) = 0$. $ord(\Delta_a(f(x)) \leq ord(f(x)) - 1$. If $\delta_{a_1, a_2, \ldots, a_i} f(x) \neq c$ then the non-linear order of $f(x) \geq i$.

Let $P = (p_{ij})$ be the transition probabilities of a homogeneous Markov chain and ${p_{ij}}^s$ is the probability that state $j$ can be reached from state $i$ in $s$ steps. **Ergodic:** aperiodic and irreducible. If a random cipher is selected from $\Sigma_{2^n}$, $Pr(P$ is ergodic $) \rightarrow 1$.

**Theorem (OConner):** Most Feistel ciphers are resistant to differential attack. Let $p_g$ be the probability of the best linear approximation of $g$. $|p_g - \frac{1}{2}| = max_k max_{\alpha \neq 0 \neq \beta} |Pr_x(g(x,h) \cdot \beta = x \cdot \alpha) - \frac{1}{2}|$ and the best $s$ round linear approximation satisfies $|p_L - \frac{1}{2}|^2 \leq |p_g - \frac{1}{2}|^2$. For DES, $s \geq 4$, $|p_L - \frac{1}{2}|^2 \leq 8|p_f - \frac{1}{2}|^4$.

An $r-$round iterated $2m$ bit block cipher with $r$-round keys each has $n$ bits. A **strong key schedule** is one in which (1) For any $s$ bits of the $r$ round keys derived from $k$ where $s < rn$, it is "hard" to find any of the remaining $rn - s$ bits from the $s$ bits, (2) given a relation between two different master keys, is it "hard" to predict the relationship between any of the round keys. $< RK_l >= nMSB(E_{k_i}(IV \oplus l))$.

## 3.7 New Ciphers

### 3.7.1 AES-Rijndael

Arithmetic in $GF(2^8)$ with minimum polynomial $m(x) = x^8 + x^4 + x^3 + x + 1$. If $m(\theta) = 0$, matrix for multiplication by $\theta$ over $GF(2)$ is denoted by $T$ and squaring by $S$, then

$$T = \begin{pmatrix} 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 1&0&0&0&1&0&0&0 \\ 1&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&1&0 \\ 1&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0 \end{pmatrix}, S = \begin{pmatrix} 1&1&0&0&0&0&0&0 \\ 0&0&1&0&1&0&0&0 \\ 0&1&1&0&0&0&0&0 \\ 1&0&0&1&0&1&0&0 \\ 1&1&1&1&0&0&0&0 \\ 0&0&1&0&0&0&1&0 \\ 1&1&0&1&0&0&0&0 \\ 0&1&0&1&0&0&0&1 \end{pmatrix}$$

$Tr(a) = a + a^p + a^{p^2} + \ldots + a^{p^{d-1}}$ and $N(a) = aa^p a^{p^2} \ldots a^{p^{d-1}}$. Linearized polynomial: $L(x) = a_0 x + a_1 x^p + a_2 x^{p^2} + \ldots + a_{d-1} x^{p^{d-1}}$; linear functions can be expressed as linearized polynomials.

**Rijndael** input: $p$ consisting of $Nb$ words, $k$ with $Nk$ words. State: 4 rows, $Nb$ columns. Key: 4 rows, $Nk$ columns. Both key rows are filled in the following order: Fill leftmost column $s_{i,0}, i = 0, 1, 2, 3$, then next column, etc.

| Nb/Nk | 4 | 6 | 8 |
|---|---|---|---|
| 4 | 10 | 12 | 14 |
| 6 | 12 | 12 | 14 |
| 8 | 14 | 14 | 14 |

```
Rijndael(p, k, Nb, Nk)  {
    ComputeRoundKeys(K, W[i])
    state= p
    AddRoundKey(state)
    for (i=0, i<Nr, i++) {
         for each byte, b in state, ByteSub(b)
         ShiftRow(state)
         if(i<Nr-1)
             MixCol(state)
         AddRoundKey(state)
         }
}

ByteSub(b) {
    t= 0
    if b!=0 {
        t= 1/b;
    // M= circ(1,0,0,0,1,1,1,1)
    // [Shift right going down].
    // a= (1,1,0,0,0,1,1,0)^T.
    return(Mt + a);
    }

ShiftRow(state) {
    shift right row 1 by 0.
    shift right row 2 by 1.
    shift right row 3 by 2 if Nb<8,
                     3 otherwise.
    shift right row 4 by 3 if Nb<8,
                     4 otherwise.
    }

MixCol(state) {
    multiply each col of state by
             c(x) (mod  x**4+1);
    // c(x)= 0x03x**3+0x01x**2+0x01x+0x02
    // d(x)= 0x0bx**3+0x0dx**2+0x09x+0x0e
    }

AddRoundKey(state) {
    state= state + W[i];
    }

ComputeRoundKeys(K[4*Nk], W[Nb*(Nr+1)]) {
    for(i=0; i<Nk; i++)
           W[i]= (K[4i], K[4i+1],
                  K[4i+2], K[4i+3])
    for(i=Nk; i<Nb*(Nr+1)); i++) {
        t= W[i-1];
        if((i mod Nk)==0)
           t= SubByte(RotByte(t))^RCon(i/Nk);
        if((i mod Nk)==4 and Nk>6)
           t=SubByte(t);
        W[i]= W[i-Nk] ^ t;
        }
    }

SubByte(w) {
    w= ByteSub(w);
    }

RotByte(w= (a,b,c,d)) {
    w= (b,c,d,a);
    }
```

```
RC[1]= 0x01;
RC[i+1]=  RC[i]*x (x in poly over GF(2));
```

```
RCon[i]= (RC[i], 0x00, 0x00, 0x00);
```

Note $[ShiftRow, MixCol] = 1$. Rounds Key: $K_{r,0}, K_{r,1}, \ldots, K_{r,15}$. First Round is input key. For $s = r+1$, $T_0 = S[K_{r,13}] + \theta^r$, $T_1 = S[K_{r,14}]$, $T_2 = S[K_{r,15}]$, $T_3 = S[K_{r,12}]$ and $K_{s,i} = K_{r,i} + T_i, 0 \leq i \leq 3$, $K_{s,i} = K_{r,i} + K_{s,i-4}, 4 \leq i \leq 15$. Note that key expansion is equivalent to: $W[i] = W[i-1] \oplus W[i-4]$, if $i \neq 0 \pmod 4$ $W[i] = T(W[i-1]) \oplus W[i-4]$, if $i = 0 \pmod 4$ where $T(a,b,c,d) = (SB(b) \oplus r(i), SB(c), SB(d), SB(a)), r(i) = 0x02^{\frac{i-4}{4}}$ in $GF(2^8)$. Inverse provides linear/differential immunity, linear diffusion provides algebraic complexity.

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} y_7 \\ y_6 \\ y_5 \\ y_4 \\ y_3 \\ y_2 \\ y_1 \\ y_0 \end{pmatrix} = L \begin{pmatrix} x_7 \\ x_6 \\ x_5 \\ x_4 \\ x_3 \\ x_2 \\ x_1 \\ x_0 \end{pmatrix}$$

$S[w] = L[w^{(-1)}] + 0x63$. Combined RowShift, ColumnMix and Diffusion and AddRound is $x \mapsto Mx + 0x63 + k_i$ where $M$ is a $16 \times 16$ matrix and $min_M(x) = (x+1)^{15})|(x^{16}+1)$ which can be transformed into $P^{-1}MP = V_1 \oplus \ldots \oplus V_{15}$ with $dim(V_i) = (16, 14^3, 10^3, 8^2, 6, 4, 4^4, 2)$.

Motivation. **Linear cryptanalysis resistance** is provided if no linear trail has a correlation coefficient $> 2^{\frac{n}{2}}$. **Differential cryptanalysis resistance** is provided if there is no differential trail with prop ratio $> 2^{1-n}$. The **prop ratio** of differential trail is approximately the product of the prop ratios of its active S-boxes. The **correlation** of a linear trail is approximately the product of the I/O correlations of its active S-boxes. The **wide trail** strategy is: (1) choose an S-box with maximum prop ratio and correlation $\approx 2^{-6}, 2^{-3}$, respectively; (b) construct diffusion layer in such a way that there are no multiple round trails with few active S-boxes. **Theorem:** The weight of a two round trail with $Q$ active columns at the input and output is $\geq 5Q$; The minimum number of active S-boxes in a four round differential or linear trail is 25.

### 3.7.2 Tea, TwoFish

```
Tea(unsigned K[4], ref unsigned L, ref unsigned R) {
    unsigned d= 0x9e3779b9;
    unsigned s= 0;
    for(int i=0; i<32;i++) {
        s+= d;
        L+= ((R<<4)+K[0])^(R+s)^((R>>5)+K[1]);
        R+= ((L<<4)+K[2])^(L+s)^((L>>5)+K[3]);
        }
    }
```

(1) 4 different $8 \times 8$ bijective, key dependent S boxes. (2) MDS code. (3) PHT: $a' = a + b \pmod{2^{32}}$, $b' = a + 2b \pmod{2^{32}}$. Basic algorithm: whiten, 16 rounds, whiten.

$$MDS = \begin{pmatrix} 0x01 & 0xef & 0x5b & 0x5b \\ 0x5b & 0xef & 0x5b & 0x01 \\ 0xef & 0x5b & 0x01 & 0xef \\ 0xef & 0x01 & 0xef & 0x5b \end{pmatrix}$$

$Round(w_1, w_2, w_3, w_4, k_1, k_2) = (w_1', w_2', w_1, w_2)$: $w_1' = w_3 + F_1(w_1, w_2, r) >>> 1$; $w_2' = (w_4 <<< 1) + F_1(w_1, w_2, r)$; $F_r(w, v) = PHT(g(w), g(v <<< 8)) + k_r \pmod{2^{32}}$; $g(x,y,z,w) = MDS \begin{pmatrix} S_1(x) \\ S_2(y) \\ S_3(z) \\ S_4(w) \end{pmatrix}$. All calculations over $GF(2^8)$.

### 3.7.3 Miscellaneous

**Cramer-Shoup:** $G = \mathbb{Z}_p$, $G =< g >=< g' >$, $H$, a collision resistant hash whose image is $\mathbb{Z}_p{}^*$. $PK = (G, g, g', h, k, k')$, $s, t, t', u, u'$ randomly selected. $h = sg$, $k = tg + t'g'$, $k' = ug + u'g'$. Encrypt (m): Choose $r$, random, set $n = H(rg|rg'|m + rnk')$. $E(m) = (x, y, z, w) = (rg, rg', m + rh, rk + rnk')$. Decryption: $D(x, y, z, w)$, check that $(nu + t)x + (nu' + t'))y = w$. If so, compute $z - sx$.

**Bit Commitment and coin flips:** $b, b' \in \{0, 1\}$. Alice sends Bob $c = commit(b)$, Bob sends Alice $b'$, Alice sends Bob $reveal(c)$. Result is $b \oplus b'$.

**Zero Knowledge** using 3 color: For each round, Prover randomly permutes colors and *commits* color at each vertex. For each round, Verifier asks to *reveal* color at the vertices of an edge. blob: commit with equality.

**Shalevi-Micali Commit:** $h$ is a one way function like $SHA1$. $commit(m) = h(r||m)$, $r$, random. $p$ a 161 bit prime. Pick $a, b$: $ax + b = z \pmod p), y = h(x), c = (y, a, b)$. $reveal(c) = x, m$.

**Time memory tradeoff:** Fix a plaintext block, $P$ and pick $SP_i, i = 1, 2, \ldots, m$. For each $i$, set $K^i_0 = SP_i$ and $K^i_{j+1} = F(E(K^i_j, P)), j = 0, 1, \ldots, t-1$ where $F$ is a randomizing function to avoid short cycles and put $EP_i = K^i_t$. For each $i$, store $(SP_i, EP_i)$. Phase 2: Get $C = E(P, K)$ from oracle where $K$ is unknown. Compute $X_0 = C$, $X_{i+1} = E(P, X_i)$ until $X_i = EP_j$ for some $i, j$. Then compute $Y_0 = EP_j$ and $Y_{j+1} = E(P, Y_j)$ until $Y_k = C$ then $K = Y_{k-1}$. If $m$ is the number of starting points for each $F$, $t$ is the number of encryptions per chain and $r$ is the number of tables. Attack requires $mr$ memory and $tr$ time with the probability of success $1 - e^{-\frac{trm}{k}}$.

**Nostradamus ("herding") attack:** Let $h$ be a Merkle-Damgard hash with compression function $f$ and initial value $IV$. Goal is to hash a prefix value (P) quickly by appending random suffixes (S). Procedure Phase 1: Pick $k$ and generate $2^k$ random values $d_{0i}$ from each pair of the values $f(IV||d_{i,i+1})$ find two messages $M_{0,j}, M_{1,j}$ which collide under $f$ and call this value $d_{1,j}$ this takes effort $2^{n/2}$ for each pair. Keep doing this (colliding $d_{i,j}, d_{i+1,j}$ under $M_{i,j}, M_{i+1,j}$ to produce $d_{i,j+1}$ until you reach $d_{2^k,0}$. This is the diamond. Publish $y = w(d_{2^k,0})$ where $w$ is the final transformation in the hash as the hash (i.e. - claim $y = h(P||S)$. The cost of phase 1 is $(2^k - 1)2^{n/2}$. In phase 2, guess $S'$ and compute $T = f(IV||P||S')$; keep guessing until $T$ is one of the $d_{ij}$. Once you get a collision, follow a path through the $M_{ij}$ to $d_{2^k,0}$, append these $M_{ij}$ to $P||S'$ and apply $w$ to get right hash.

## 3.8 Cryptographic Hashes

**Weak collision resistance:** Given $x$, it is computationally infeasible to find $x' \neq x$ with $h(x) = h(x')$. **Strong collision resistance:** It is computationally infeasible to find $x' \neq x$ with $h(x) = h(x')$ for any $x$, $x'$. **One-way:** Given a digest $z$, it is computationally infeasible to find $x$ with $h(x) = z$. Strongly collision resistant implies one-way.

**Merkle Damgard construction:** $z_0 = IV, z_{i+1} = f(z_i, m_i), h(m) = g(z_r)$, where $f$ is a compression function, $r$ is the number of rounds and $m = m_1||m_2||\ldots||m_r$. If $f$ is collision resistant then so is $h$. **Hash from Block Cipher:** $g_i = e_{g_{i-1}}(x_i) + x_i$, $g_i = e_{g_{i-1}}(x_i) + x_i + g_{i-1}$, $g_i = e_{g_{i-1}}(g_{i-1} + x_i) + x_i$, $g_i = e_{g_{i-1}}(g_{i-1} + x_i) + g_{i-1} + x_i$.

**Chaum Hash:** $\alpha$, $\beta$ two primitive elements of $\mathbb{Z}_p$, $h(x, y) = \alpha^x \beta^y \pmod p$. If there's a collision, $log_\alpha(\beta)$ can be computed efficiently. $h(0^{t+1}||y_1)$, $g_{i+1} = h(g_i||1||y_{i+1})$. Todo: do reduction proof.

**Iterative construction is vulnerable to multi-collision (Joux):** Suppose $M1, M1'; M2, M2'; \ldots; Mt, Mt'$ all collide. From these we get $2^t$ collisions. If $r$ people each have one of $N$ possible birthdays, there is a greater than .5 chance of $k$ collisions if $r > N^{\frac{k-1}{k}}$. Todo: prove this fact.

**Random Oracle Model:** Let $f$ be a OWF with trapdoor, $(y_1, y_2) = (f(r), h(r) + m)$ is used as encryption. An oracle with $l$ requests $L$, $Pr(guess\ right) = P(r \in L) + \frac{1}{2}P(\neg r \in L)$. Set $p = \frac{1}{2} + e$, $e \leq Pr(r \in L)$. Canetti, Goldreich, Halevi constructed a cryptosystem that is secure in Random Oracle Model but insecure for any concrete hash.

**MD-4:** In description below, K[0]= 0, K[1]= 0x5a827999, K[2]= 0x6ed9eba1. $F(A,B,C) = (A \wedge B) \vee (\neg A \wedge C)$, $G(A,B,C) = (A \wedge B) \vee (A \wedge B) \vee (B \wedge C)$, $H(A,B,C) = (A \oplus B) \oplus C$. $W_i = X_{\sigma(i)}, i = 0, 1, \ldots, 47$. $Q_{-4} = A$, $Q_{-3} = D$, $Q_{-2} = C$, $Q_{-1} = B$. $Q_i(A,B,C) = (Q_{i-4} + F(Q_{i-1}, Q_{i-2}, Q_{i-3}) + W_i + K_0) <<< s_i, 0 \leq i \leq 15$, $Q_i(A,B,C) = (Q_{i-4} + G(Q_{i-1}, Q_{i-2}, Q_{i-3}) + W_i + K_1) <<< s_i, 16 \leq i \leq 31$, $Q_i(A,B,C) = (Q_{i-4} + H(Q_{i-1}, Q_{i-2}, Q_{i-3}) + W_i + K_2) <<< s_i, 32 \leq i \leq 47$.

```
MD-4(Y[0] , ..., Y[N-1])
    K[0]= 0; K[1]= 0x5a827999; K[2]= 0x6ed9eba1;
    (A, B, C, D)= (0x67452301, 0xefcdab89, 0x98badcfe, 0x10325476);
    for(i=0; i<(N/16); i++) {
        X[j]= Y[16i+j], j= 0, 1, ..., 15;
        W[j]= X[SIGMA(j)], j= 0, 1, ..., 47;
Q[-4]= A;
Q[-3]= D;
Q[-2]= C;
Q[-1]= B;
// Calculate Q[i] recursively according to formula above
        (A, B, C, D)+= (Q[44], Q[45], Q[46], Q[47]);
(A, B, C, D)= (A, D, C, B);
        }
    return (A, B, C, D);
```

**Dobbertin attack on MD4:** Let $M$ and $M'$ be 512 bit messages consisting of 16, 32-bit works $X_0, X_1, \ldots, X_{15}$ with $X_i = X'_i$ for all $i$ except $i - 12$ and let $X'_{12} = X_{12} + 1 \pmod{2^{32}}$. We want to find a collision. $X_{12}$ is first used in step 12 and last used in step 35. $\Delta_i = (Q'_j - Q_j, Q'_{j-1} - Q_{j-1}, Q'_{j-2} - Q_{j-2}, Q'_{j-3} - Q_{j-3})$ after step $i$. Dobbertin attack consists of three steps: (1) Show that if $\Delta_{19} = (0, 2^{25}, -2^5, 0)$ then $\Delta_{35} = (0,0,0,0)$ with probability $p > 2^{-30}$ (actually, $p > 2^{-22}$); (2) get conditions on $M$ (i.e. on the $X_i$) based on round 12, that guarantee $\Delta_{19} = (0, 2^{25}, -2^5, 0)$; (3) find $X_0, X_1, \ldots, X_{11}$ that produce candidates that present the desired conditions at step 12, after about $2^{22}$ of these, you'll get a collision. The work factor is about $2^{20}$.
**1.** Steps 19-35. Suppose $\Delta_{19} = (0, 2^{25}, -2^5, 0)$ and $G(Q_{19}, Q_{18}, Q_{17}) = G(Q'_{19}, Q'_{18}, Q'_{17})$, then the following table holds:

| $j$ | $\Delta(Q_j)$ | $\Delta(Q_{j-1})$ | $\Delta(Q_{j-2})$ | $\Delta(Q_{j-3})$ | $i$ | $s_j$ | $p$ | In |
|---|---|---|---|---|---|---|---|---|
| 19 | $2^{25}$ | $-2^5$ | 0 | 0 | $*$ | $*$ | $*$ | $*$ |
| 20 | 0 | $2^{25}$ | $-2^5$ | 0 | 1 | 3 | 1 | $X_1$ |
| 21 | 0 | 0 | $2^{25}$ | $-2^5$ | 1 | 5 | $\frac{1}{9}$ | $X_5$ |
| 22 | $-2^{14}$ | 0 | 0 | $2^{25}$ | 1 | 9 | $\frac{1}{3}$ | $X_9$ |
| 23 | $2^6$ | $-2^{14}$ | 0 | 0 | 1 | 13 | $\frac{1}{3}$ | $X_{13}$ |
| 24 | 0 | $2^6$ | $-2^{14}$ | 0 | 1 | 3 | $\frac{1}{9}$ | $X_2$ |
| 25 | 0 | 0 | $2^6$ | $-2^{14}$ | 1 | 5 | $\frac{1}{9}$ | $X_6$ |
| 26 | $-2^{23}$ | 0 | 0 | $2^6$ | 1 | 9 | $\frac{1}{3}$ | $X_{10}$ |
| 27 | $2^{19}$ | $-2^{23}$ | 0 | 0 | 1 | 13 | $\frac{1}{3}$ | $X_{14}$ |
| 28 | 0 | $2^{19}$ | $-2^{23}$ | 0 | 1 | 3 | $\frac{1}{9}$ | $X_3$ |
| 29 | 0 | 0 | $2^{19}$ | $-2^{23}$ | 1 | 5 | $\frac{1}{9}$ | $X_7$ |
| 30 | $-1$ | 0 | 0 | $2^{19}$ | 1 | 9 | $\frac{1}{3}$ | $X_{11}$ |
| 31 | 1 | $-1$ | 0 | 0 | 1 | 13 | $\frac{1}{3}$ | $X_{15}$ |
| 32 | 0 | 1 | $-1$ | 0 | 2 | 3 | $\frac{1}{3}$ | $X_0$ |
| 33 | 0 | 0 | 1 | $-1$ | 2 | 9 | $\frac{1}{3}$ | $X_8$ |
| 34 | 0 | 0 | 0 | 1 | 2 | 11 | $\frac{1}{3}$ | $X_4$ |
| 35 | 0 | 0 | 0 | 0 | 2 | 15 | 1 | $X_{12}, X_{12}+1$ |

**Steps 12 to 19**. To get $\Delta_{19} = (0, 2^{25}, -2^5, 0)$, $Q_{16} = Q'_{16}$, $Q_{19} = Q'_{19} + 2^{25}$, $Q_{18} + 2^5 = Q'_{18}$, $Q_{17} = Q'_{17}$ and $Q_i = Q'_i, 8 \leq i \leq 11$.

| $j$ | $i$ | M In | M' In |
|---|---|---|---|
| 12 | 0 | $X_{12}$ | $X_{12}+1$ |
| 13 | 0 | $X_{13}$ | $X_{13}$ |
| 14 | 0 | $X_{14}$ | $X_{14}$ |
| 15 | 0 | $X_{15}$ | $X_{15}$ |
| 16 | 1 | $X_0$ | $X_0$ |
| 17 | 1 | $X_4$ | $X_4$ |
| 18 | 1 | $X_1$ | $X_8$ |
| 19 | 1 | $X_{12}$ | $X_{12}+1$ |

These yield the following conditions: $(Q'_{12} <<< 29) - (Q_{12} <<< 29) = 1$, $F(Q'_{12}, Q_{11}, Q_{10}) - F(Q_{12}, Q_{11}, Q_{10}) = (Q'_{13} <<< 25) - (Q_{13} <<< 25)$, $F(Q'_{13}, Q_{12}, Q_{11}) - F(Q_{13}, Q_{12}, Q_{11}) = (Q'_{14} <<< 21) - (Q_{14} <<< 21)$, $F(Q'_{14}, Q_{13}, Q_{12}) - F(Q_{14}, Q_{13}, Q_{12}) = (Q'_{15} <<< 13) - (Q_{15} <<< 13)$, $G(Q'_{15}, Q'_{14}, Q_{13}) - G(Q_{15}, Q_{14}, Q_{13}) = Q_{12} - (Q'_{12}$, $G(Q'_{16}, Q'_{15}, Q_{14}) - G(Q_{16}, Q_{15}, Q_{13}) = Q_{13} - (Q'_{13}$, $G(Q'_{17}, Q'_{16}, Q_{15}) - G(Q_{17}, Q_{16}, Q_{14}) = Q_{12} - Q'_{12} + (Q_{18} <<< 23) - (Q_{18} <<< 23)'$, $G(Q'_{18}, Q'_{17}, Q_{16}) - G(Q_{18}, Q_{17}, Q_{15}) = Q_{15} - Q'_{15} + (Q_{19} <<< 19) - (Q_{19} <<< 19)')$. For the solutions, $(Q_{10}, Q_{11}, Q_{12}, Q_{13}, Q_{14}, Q_{15}, Q_{16}, Q_{17}, Q_{18}, Q_{19}, Q'_{12}, Q'_{13}, Q'_{14}, Q'_{15})$, $\Delta_{19}$ will hold if $X_{13} = \ anything$, $X_{14} = (Q_{14} <<< 21) - Q_{10} - F(Q_{13}, Q_{12}, Q_{11})$, $X_{15} = (Q_{15} <<< 13) - Q_{11} - F(Q_{14}, Q_{13}, Q_{12})$, $X_0 = (Q_{16} <<< 29) - Q_{12} - G(Q_{15}, Q_{14}, Q_{13}) - K_1$, $X_4 = (Q_{17} <<< 27) - Q_{13} - G(Q_{16}, Q_{15}, Q_{14}) - K_1$, $X_8 = (Q_{18} <<< 23) - Q_{14} - G(Q_{17}, Q_{16}, Q_{15}) - K_1$, $X_{12} = (Q_{19} <<< 19) - Q_{15} - G(Q_{18}, Q_{17}, Q_{16}) - K_1$, $Q_9 = (Q_{13} <<< 25) - F(Q_{12}, Q_{11}, Q_{10}) - X_{13}$, $Q_8 = (Q_{12} <<< 19) - F(Q_{11}, Q_{10}, Q_{09}) - X_{12}$. Can choose $Q_{12} = -1$, $Q'_{12} = 0$, $Q_{11} = 0$ to simplify. This means we can pick $Q_{14}, Q_{15}, Q_{16}, Q_{17}, Q_{18}, Q_{19}$ arbitrarily and determine $Q_{10}, Q_{13}, Q'_{13}, Q'_{14}, Q'_{15}$ subject to the checks $G(Q_{15}, Q_{14}, Q_{13}) - G(Q'_{15}, Q'_{14}, Q'_{13}) = 1$ and $F(Q'_{14}, Q'_{13}, 0) - F(Q_{14}, Q_{13}, -1) - (Q'_{15} <<< 13) + (Q_{15} <<< 13) = 0$. Finally, we must insure the solutions is admissible by checking that $G(Q'_{19}, Q'_{18}, Q_{17}) = G(Q_{19}, Q_{18}, Q_{17})$. Under these circumstances the solution is a candidate for the differential. Once one candidate is found use the "continuity" of $F$ and $G$ by modifying one bit of the candidate at a time, the continuity makes it likely this will work.

**Steps 0 to 11**. Having found $Q_8, Q_9, Q_{10}, Q_{11}$ such that

$$MD4_{12,\ldots,47}(Q_8, Q_9, Q_{10}, Q_{11}, X) = MD4_{12,\ldots,47}(Q_8, Q_9, Q_{10}, Q_{11}, X')$$

we need to find $MD4_{0,\ldots,11}(IV, X) = (Q_{11}, Q_{10}, Q_9, Q_8)$. We are free to choose $X_j, j = 1, 2, 5, 6, 7, 9, 10, 11$. We pick $X_1, X_2, X_3, X_5$ at random and compute $X_6, X_7, X_9, X_{10}, X_{11}$ such that $MD4_{6,\ldots,11}(Q_2, Q_3, Q_4, Q_5, X) = (Q_{11}, Q_{10}, Q_9, Q_8)$. Since $Q_{11} = (Q_7 + F(Q_{10}, Q_9, Q_8) + X_{11}) <<< 19$, if can do this by making $X_{11} = (Q_{11} <<< 13) - Q_7 - F(Q_{10}, Q_9, Q_8)$ and similarly for $X_{10}, X_9$. We can't do this for $X_9$ but since $Q_8 = (Q_4 + F(Q_7, Q_6, Q_5) + X_8) <<< 3$, if $Q_7 = -1, Q_6 = (Q_8 <<< 29) - Q_4 - X_8$ the desired equation holds for all such $X_8$; in particular, by picking $X_6 = (Q_6 <<< 21) - Q_2 - F(Q_5, Q_4, Q_3)$ and $X_7 = (Q_7 <<< 13) - Q_3 - F(Q_6, Q_5, Q_4)$. These guarantee $\Delta_{35} = 0$.

```
SHA1($M$,$n$)
// $M$ is message, $n$ is number of 512 bit blocks
    M= SHA1Pad(M)
    $f_i(B,C,D) = (B \wedge C) \vee (\overline{B} \wedge D), 0 \leq i \leq 19$
    $f_i(B,C,D) = (B \oplus C \oplus D), 20 \leq i \leq 39$
    $f_i(B,C,D) = (B \wedge C) \vee (B \wedge D) \vee (C \wedge D), 40 \leq i \leq 59$
    $f_i(B,C,D) = (B \oplus C \oplus D), 60 \leq i \leq 79$

    $K_i = 0x5a827999, 0 \leq i \leq 19; K_i = 0x6ed9eba1, 20 \leq i \leq 39$
    $K_i = 0x8f1bbcdc, 40 \leq i \leq 59; K_i = 0x6a62c1d6, 60 \leq i \leq 79$

    $H_0 = 0x67452301, H_1 = 0xefcdab89, H_2 = 0x98badcfe, H_3 = 0x10324576, H_4 = 0xc3d2e1f0$

    for (i=0, $i < n$, i++) {
        $M_i = W_0||W_1||\ldots||W_{15}$
        for($j = 16, j < 80, j++$) {
            // $ROTL^1$ below is difference between SHA-0 and SHA-1
            $W_j = ROTL^1(W_{j-3} \oplus W_{j-8} \oplus W_{j-14} \oplus W_{j-16})$
        }
        $A = H_0, B = H_1, C = H_2, D = H_3, E = H_4$
        for($j = 0, j < 80$; j++) {
           $ROTL^5$ below is correlated to lowest wt differential
           $t = ROTL^5(A) + f_j(B, C, D) + E + W_j + K_j$
           $E = D, D = C, C = ROTL^{30}(B), B = A, A = t$
           }
        $H_0+ = A, H_1+ = B, H_2+ = C, H_3+ = D, H_4+ = E$
        }
```

```
SHA-1Pad(x)   // with MD strengthening
    Append 1 and enough 0's until there are 64 bits remaining
    Append size hashed in 64 bit format
    return(x)
```

**Shamir's non-linear functions with maximal period:** $x \rightarrow x^2 \wedge c$, $x \rightarrow x + 4h(x) + 1$, Example: $x \rightarrow (x+1)(2x+1)$.

**Changes from MD4 to MD5:** (1) 64 steps, function for final 16 rounds is $I(A, B, C) = B \oplus (A \vee \neg C)$, (2) $G(A, B, C) = (A \wedge C) \vee (B \wedge \neg C)$, (3) each round uses different constant, (4) each step adds result of previous step, (5) the order of input words to the steps is different,(6) shift values are different. Chinese attack uses "precise" differential (signed difference) where 0 indicates no difference, + indicates 1 → 0 difference and − indicates 0 → 1 difference. This is different from both xor and modular difference; for example, if $z' = 10100101, z = 10010101$, $\nabla(z', z) = 00 + -0000$.

**Chinese attack on MD5.** Attack proceeds in four phases: (1) specify input differential patters via modular difference (hard and "done by hand" according to Wang), (2) specify output differential pattern (only 1 known) that is easily satisfied in earlier rounds, (3) derive sufficient conditions propagation; (4) generate pairs of 1024 bit numbers that satisfy 3 (deterministically when possible). To do step 4: (a) generate $M_0$ at random; (b) use single step modification to $M_0$ to satisfy sufficient conditions; (c) use multi-step modifications to insure conditions hold in middle rounds; (d) check conditions for all remaining steps; (e-f) do the same for $M_1$; compute $M_0' = M_0 + \Delta M_0$ and $M_1' = M_1 + \Delta M_1$ according to the input differential. **Conditions:** $T_j = F(Q_{j-1}, Q_{j-2}, Q_{j-3}) + Q_{j-4} + K_j + W_j$, $R_j = T_j <<< s_j$, $Q_j = Q_{j-1} + R_j$, now apply modular difference and derive conditions on $\Delta T_j$ and $\Delta Q_j$ for differential (below) to hold.

$\Delta X = X' - X$. $\Delta H_0 \rightarrow_{(M_0, M_0')} \Delta H_1 \rightarrow_{(M_1, M_1')} \Delta H_2 \ldots \rightarrow_{(M_{i-1}, M_{i-1}')} \Delta H_i = H$ with each composed of $\Delta H_i \rightarrow_{P_2} \Delta R_{i+1,1} \rightarrow_{P_2} \Delta R_{i+1,2} \rightarrow_{P_3} \Delta R_{i+1,3} \rightarrow_{P_4} \Delta R_{i+1,4} = \Delta H_{i+1}$. Let $\Delta i, j = x_{i,j}' - x_{i,j} = \pm 1$ and $\Delta x_i[j_1, j_2, \ldots, j_l] = x_i[j_1, j_2, \ldots, j_l] - x_i$. Collision is caused by 1024 bit input: $(M_0, M_1)$ with $\Delta M_0 = (0, 0, 0, 0, 2^{31}, 0, 0, 0, 0, 0, 0, 2^{15}, 0, 0, 2^{31}, 0)$ and $\Delta M_1 = (0, 0, 0, 0, 2^{31}, 0, 0, 0, 0, 0, 0, -2^{15}, 0, 0, 2^{31}, 0)$. Sufficient conditions insure that differential holds with high probability. At 8th iteration, $b_2 = c_2 + (b_1 + F(c_2, d_2, a_2) + m_7 + t_7) <<< 22$, we try to control $(\Delta c_2, \Delta d_2, \Delta a_2, \Delta b_1) \rightarrow \Delta b_2$ with the following (A) non-zero bits of $\Delta b_2$: $d_{2,11} = 1, b_{2,1} = 0$, $d_{2,26} = \overline{a_{2,26}} = 1, b_{2,16} = 0$, $d_{2,28} = \overline{a_{2,28}} = 0, b_{2,i} = 0$, $d_{2,11} = 1, b_{2,24} = 0$; (B) zero bits of $\Delta b_2$: $c_{2,i} = 0$, $d_{2,i} = a_{2,i}$, $c_{2,1} = 1$, $d_{2,6} = \overline{a_{2,6}} = 0$, $d_{2,i} = 0$, $d_{2,12} = 1$, $a_{2,24} = 0$, 7th bit of $c_2, d_2, a_2$ result in no change in $b_2$. Algorithm 1: Repeat until first block is found (a) Select random $M_0$, (b) Modify $M_0$, (c) $M_0, M_0' = M_0 + \Delta M_0$ produce $\Delta M_0 \rightarrow (\Delta H_1, \Delta M_1)$ with probability $2^{-37}$, (d) Test characteristics. 2: Repeat until first block is found (a) Select random $M_1$, (b) Modify $M_1$, (c) $M_1, M_1' = M_1 + \Delta M_1$ produce $\Delta M_1 \rightarrow 0$ with probability $2^{-30}$, (d) Test characteristics.

Comments from NIST: Randomization (prevent offline computation for herding): $RMX(r, M_1 | \ldots | M_L) = (r | m_1 \oplus r | \ldots | m_L \oplus r)$. $H_r(M_1 | \ldots | M_L) = H(r | m_1 \oplus r | \ldots | m_L \oplus r)$. Transmit $r$. Herding attack: first committing to an output $h$, then mapping messages with arbitrary starting values to $h$. Joux: If $H_1, H_2$ are n bit hashes; $H_1(M) || H_2(M)$ can be broken in $O(n 2^{\frac{n}{2}})$. Haifa: $h_{i+1} = CF(h_i, M_i, bitlength, salt)$.

**Joux attack on SHA-0:** For SHA-0, change bit 1 which shifts to bit 31 and because of no carry: it is linear in $\oplus$. Disturbance bit vector: $(m_0^{(0)}, m_0^{(1)}, \ldots, m_0^{(79)})$. Perturbation mask: $-5 \leq i \leq -1, M_0^{(i)} = 0$, $0 \leq i \leq 79, M_{0,k}^{(i)} = 0$, if $k \neq 1$ $0 \leq i \leq 79, M_{0,1}^{(i)} = M_0^{(i)}$. Corrective masks: $-4 \leq i \leq 79, M_1^{(i)} = ROL_5(M_1^{(i-1)})$, $-3 \leq i \leq 79, M_2^{(i)} = M_1^{(i-2)}$, $-2 \leq i \leq 79, M_3^{(i)} = ROL_{30}(M_1^{(i-3)})$, $-1 \leq i \leq 79, M_3^{(i)} = ROL_{30}(M_1^{(i-4)})$, $0 \leq i \leq 79, M_3^{(i)} = ROL_{30}(M_1^{(i-5)})$. Early round differentials are prescribed and later round differentials hold with non-negligible probability ($2^{-61}$, $2^{-56}$ using neutral bits — A bit is neutral if flipping it doesn't change differential pattern). Multi-block: patch final round errors in next block. Early rounds are non-linear and prescribed. Late rounds linear and probabilistic. Final rounds can be "patched". Procedure: Fix linear characteristic, fix non-linear characteristic, modify message (keeping differential) if conflict in mid round.

**SHA-256 definitions:** $Ch(x,y,z) = (x \wedge y) \oplus (\neg x \wedge z)$, $Maj(x,y,z) = (x \wedge y) \vee (x \wedge z) \vee (y \wedge z)$.
$\psi_{256}^{i,j,k}(x) = ROTR^i(x) \oplus ROTR^j(x) \oplus ROTR^k(x)$, $\phi_{256}^{i,j,k}(x) = ROTR^i(x) \oplus ROTR^j(x) \oplus SHR^k(x)$.
$\Sigma_0^{256}(x) = \psi_{256}^{2,13,22}(x)$, $\Sigma_1^{256}(x) = \psi_{256}^{6,11,25}(x)$.
$\sigma_0^{256}(x) = \phi_{256}^{7,18,3}(x)$, $\sigma_1^{256}(x) = \phi_{256}^{17,19,10}(x)$.

**SHA-512 definitions:** $Ch(x,y,z) = (x \wedge y) \oplus (\neg x \wedge z)$, $Maj(x,y,z) = (x \wedge y) \vee (x \wedge z) \vee (y \wedge z)$.
$\psi_{512}^{i,j,k}(x) = ROTR^i(x) \oplus ROTR^j(x) \oplus ROTR^k(x)$, $\phi_{512}^{i,j,k}(x) = ROTR^i(x) \oplus ROTR^j(x) \oplus SHR^k(x)$.
$\Sigma_0^{512}(x) = \psi_{512}^{28,34,39}(x)$, $\Sigma_1^{512}(x) = \psi_{512}^{14,18,41}(x)$.
$\sigma_0^{512}(x) = \phi_{512}^{1,8,7}(x)$, $\sigma_1^{512}(x) = \phi_{512}^{19,61,6}(x)$.

**SHA-256($M_1||M_2||\ldots||M_N$):**

```
for($i = 1; i \leq N; i++$) {
    $W_t = M_t^{(i)}, 0 \leq t \leq 15$,
    $W_t = \sigma_1^{256}(W_{t-2}) \oplus W_{t-7} \oplus \sigma_0^{256}(W_{t-15}) \oplus W_{t-16}, 16 \leq t \leq 63$;
    $a = H_0^{(i-1)}$; $b = H_1^{(i-1)}$; $c = H_2^{(i-1)}$; $d = H_3^{(i-1)}$;
    $e = H_4^{(i-1)}$; $f = H_5^{(i-1)}$; $g = H_6^{(i-1)}$; $e = H_7^{(i-1)}$;
    for($t = 0; t < 64; t++$) {
        $T_1 = h + \Sigma_1^{256}(e) + Ch(e,f,g) + K_t^{256} + W_t$; $T_2 = \Sigma_0^{256}(a) + Maj(e,f,g)$;
        $h = g$; $g = f$; $f = e$; $e = d + T_1$; $d = c$;
        $c = b$; $b = a$; $a = T_1 + T_2$;
        }
    $H_0^{(i)} = a + H_0^{(i-1)}$; $H_1^{(i)} = b + H_1^{(i-1)}$; $H_2^{(i)} = c + H_2^{(i-1)}$; $H_3^{(i)} = d + H_3^{(i-1)}$;
    $H_4^{(i)} = e + H_4^{(i-1)}$; $H_5^{(i)} = f + H_5^{(i-1)}$; $H_6^{(i)} = g + H_6^{(i-1)}$; $H_7^{(i)} = h + H_7^{(i-1)}$;
    }
```

SHA-512 is the same except there are 79 rounds and the words are 64 bits long.

## 3.9 Elliptic Curve Crypto

$E_F(a,b) : y^2 = x^3 + ax + b$ where $a, b \in F$ and $char(F) \neq 2, 3$; we sometimes write $E_q(a,b)$ if $F = GF(q)$. For ECC, also require smooth; namely, $4a^3 + 27b^2 \neq 0 \pmod{p}$, $p = char(F)$. For $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ define $P+Q = (x_3, y_3)$ with $x_3 = \lambda^2 - x_1 - x_2, y_3 = \lambda(x_1 - x_3) - y_1$ where $\lambda = \frac{(y_1 - y_2)}{(x_1 - x_2)}$ if $P \neq Q$ and $\lambda = \frac{(3x_1^2 + a)}{(2y_1)}$ if $P = Q$. For $char(F) = 2$, $E_F(a,b) : y^2 + xy = x^3 + ax + b$ and $x_3 = \lambda^2 + \lambda + a + x_1 + x_2, y_3 = \lambda(x_1 + x_3) + x_3 + y_1$ where $\lambda = \frac{(y_1 - y_2)}{(x_1 - x_2)}, P \neq Q$ and $\lambda = x_1 + \frac{y_1}{x_1}, P = Q$. For an ECC system, the public key parameters are $q, a, b, P$ ($P$ is called the base point); pick $1 < x < p$, $x$ is the private key. Public key is $Q = xP$. **ECDLP:** Find $x$ knowing $Q$. **ECC Encrypt:** To encrypt $m$ (already an integer in the right range), map it to a point on the curve $P_M$, pick $1 < k < p$, send $(kP, kQ + P_M)$. **ECC Decrypt:** Receive $(L, M)$ calculate $M - xL = P_M$ and map it back to the integer message. Here is a way to embed integers in curves: For $q = p^r$, odd, select parameter $\kappa$ so that the probability of failure is $2^{-\kappa}$; $m$ is message and $0 \leq m < M, q > \kappa M$ and $x = m\kappa + j \in F_q$ now for the first $j$ for which $x^3 + ax + b$ is a square, use the corresponding point $P = (x, \sqrt{x})$. **ECDSA sign:** Select $k$ at random, compute $kP, r = f_E(kP), s = k^{-1}(H(M) + xr)$. Signature is $(r, s)$. Verify: $u_1 = s^{-1}H(M), u_2 = s^{-1}r$, accept if $f_E(u_1P + u_2Q) = r$. Note: $(k, \#E) = 1$.

**Curve selection:** Avoid **anomalous curves** (Definition: $char(F) \mid \#E_F(a,b)$), and **supersingular curves** (Definition: $\#E_q(a,b) = q + 1 - t, q \mid t$ — $t$ is Frobenius trace satisfying $(\phi_q)^2 - t\phi_q + q = 0$; also $t$ is $Tr(\phi_q)$), CM 3 ($a = 0, p = 3 \pmod{4}$), MOV-vulnerable (Frey-Ruck) For comparison, attacks on DLP: $L(v,c,n) = exp(c(ln(p)^v(ln(ln(p))^{1-v})$, NFS discrete log is $L_n[\frac{1}{3}, (\frac{64}{9})^{\frac{1}{3}}]$. Best known ECDLP is $EC(n) = \sqrt{n}$. In comparisons, usually put $n = lg(\lceil q \rceil), N = lg(\lceil p \rceil)$ and put $\frac{E_{EC}}{E_{CONV}} = \frac{2^{\frac{n}{2}}}{exp(cN^{\frac{1}{3}}(log(N(log(2))^{\frac{2}{3}}}$.

**NIST Curves:** Use prime fields $\mathbb{F}_p$ with $p = 2^{192}-2^{64}-1, 2^{224}-2^{96}+1, 2^{256}-2^{224}+2^{192}+2^{96}-1, 2^{384}-2^{128}-2^{96}+2^{32}-1, 2^{521}-1$ or binary fields $\mathbb{F}_q$ with $q = 2^{163}, 2^{233}, 2^{283}, 2^{409}, 2^{571}$. $\#E_p(a,b) = q+1-t, |t| \leq 2\sqrt{q}$ and $t$ is called the trace of $E$. $E_q(a,b)$ has rank 1 or 2, that is: $E_q(a,b) \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2}$ and $n_2 \mid n_1, n_2 \mid (q-1)$. If $n_2 = 1, E_q(a,b) \cong \mathbb{Z}_{n_1} = \{kP : 0 < k < n_1\}$ and $P$ is a generator. $E_q(a_1,b_1) \cong E_q(a_2,b_2)$ if $a_1 = u^4 a_2$ and $b_1 = u^4 b_2$. $E_q, q = p^n$ is supersingular if $p \mid t$. Field represented as polynomial or normal basis. Hyperelliptic: higher genus. **Weil-Deligne:** Set $\zeta(t, E/F_q) = exp(\sum_r \frac{N_r t^r}{r})$, where $N_r$ is the number of solutions of $E/F_{q^r}$. $\zeta(t,E) = \frac{a-at+qt^2}{(1-t)(1-qt)}, N_1 = q+1-a, N_r = q^r+1-\alpha^r-\beta^r$ where $\alpha, \beta$ are reciprocal roots of the numerator. Random selection of $(E,B)$: Generate $x,y,a$ at random and compute $b = y^2-(x^3+ax)$, check there are not multiple roots. To compute $|E|$, use Schoof.

**MOV Attack:** $E_q(a,b) \mapsto F^*_{q^k}$ if $n$, the curve order, satisfies $n \mid (q^k-1)$ then use index calculus, small probability of supersingular or $k \leq log^2(q)$. Attack fails if $k > log^2(q)$ (Frey and Ruck extended the attack).

**IBE:** Suppose $p = 6q-1$, $E_p : y^2 = x^3+1 \pmod{p}$ and suppose $\#E = 6q$. $\exists P_0 \neq \infty$ and $qP_0 = \infty$. Finally, suppose there is a bilinear map, $\tilde{e}(P,Q)$, from points into $q$-th roots of unity that is easy to compute with $\tilde{e}(aP_0, bP_0) = \tilde{e}(P_0,P_0)^{ab}$. $\tilde{e}(P_0,P_0) \neq 0$ and two hash functions: $H_1 :< 2^\infty > \rightarrow kP_0$ and $H_2 : \{\omega^i\} \rightarrow < 2^n >$. Pick a secret $s : P_1 = sP_0$. To encrypt to $ID$: set $D_U = sH_1((ID)$, $g = \tilde{e}(H_1(ID), P_1)$, choose $r \neq 0 \pmod{q}$ and compute $t = m \oplus H_2(g^r)$, $A \rightarrow B : c = (rP_0, t)$. To decrypt: Get $(u,v)$, compute $h = \tilde{e}(H_1(D_u, u)$, $m = v \oplus H_2(h)$. Note $h = g^r$.

**ECC Point Operation Costs:** $I$ = inverse cost $/GF(p)$. $S$ = square cost $/GF(p)$. $M$ = multiply cost $/GF(p)$.

| **Operation** | **Cost** | **Modular Op** | **Cost** |
|---|---|---|---|
| $2P$ | $I+2S+2M$ | Add, Sub | $O(lg(n))$ |
| $P+Q$ | $I+S+2M$ | Multiply | $O(lg(n)^2)$ |
| $2P+Q$ | $2I+2S+2M$ | Invert | $O(lg(n)^2)$ |
| $P+Q$, $P-Q$ | $I+2S+4M$ | Exp | $O(lg(n)^3)$ |

If $X =< X_1, X_2, \ldots, X_n >$ and $Y =< Y_1, Y_2, \ldots, Y_n >$ then $Pr(\Delta X, \Delta Y) = \frac{1}{2^n}$ for perfect differential resistance. $(\Delta X, \Delta Y)$ is a differential characteristic. $N_D = \frac{c}{p_D}$ and $p_D = \prod_i^\gamma \beta_i$ where $\gamma$ is the number of active boxes.

$Tr(x) = x + x^p + \ldots + x^{p^{n-1}}$. $e, d$ is a dual basis if $Tr(d^{(i)}e^{(j)}) = \delta(i \oplus j)$.

## 3.10 Algebraic and other attacks

**Hadamard-Walsh:** $W_f(w)$, measures distance to affine and completely determines $f$. **Autocorrelation:** $r_f(w)$ measures differential and does not determine $f$.

**Balanced:** weight is $2^{n-1}$. $CI_f(t)$: output is statistically independent on any $t$ input bits. **Resilient:** $R_f(t)$ is $CI_f(t)$ and balanced. **Non-linearity:** $N_f$ is distance to affine. $N_f = min_{g \in RM(1,n)} d(f,g) = 2^{n-1} - \frac{1}{2} max_w |W_f(w)|$. $\epsilon = \frac{N_f}{2^n} - \frac{1}{2}$ **Linearity:** $L_f = max_w |W_f(w)|$. $D_w(f(x)) = f(x) \oplus f(w+x)$
**Theorem:** $r_f(w) = 2^{-n} \sum_u W_f(u)^2 (-1)^{u \cdot w}$. For iterated ciphers, once the number of rounds is high enough to generate $G$ (usually $A_n$), more rounds don't help.

**AES:** $8j+m$ component is $v_{(j,m)}$. $0 = w_{0,(j,m)} + p_{(j,m)} + k_{0,(j,m)}$, $0 = x_{i,(j,m)} w_{i,(j,m)} + 1, i = 1,2,\ldots,9$. $0 = w_{i,(j,m)} + (Mx_{i-1})_{(j,m)} + k_{i,(j,m)}, i = 1,2,\ldots,9$, $0 = c_{(j,m)} + (M^* x_9)_{(j,m)} + k_{10,(j,m)}$. $M$ is the combined effect of ShiftRow, MixColumn and the Linear diffusion. 5248 equations, 3840 sparse quadratic, 1408 linear diffusion, 7808 terms, 2560 state variables, 1408 key variables. $1280 + 1408 = 2588$ state/key variables eliminated, $4288 - 2688 = 1600$ unknown. 2688 equations, 1280 sparse quadratic, 5248 terms, 2560 state, 1408 linear diffusion, 1408 key variables.

For AES: $M : x \mapsto CRLx + 63$ (Everything but subByte). Minimal polynomials: $C : (x^4+1)$, $R : (x^4+1)$, $L : (x+1)^3$, $C : (x+1)^{15}$. **BES:** $b \rightarrow M_B b^{-1} + k_B$. $w_0 = p + k_0$, $x_i = w_i^{-1}$, $w_i = M_B x_{i-1} + k_i$, $c = M_B^* x_9 + k_{10}$. $AES_k(P) = C \leftrightarrow BES_{\phi(k)}(\phi(P)) = \phi(C)$, $\phi(a) = (a^{2^0}, a^{2^1}, a^{2^2}, a^{2^3}, a^{2^4}, a^{2^5}, a^{2^6}, a^{2^7})$.

**Circulant as linearized polynomial:** $x \mapsto 0x05x^{2^0} + 0x09x^{2^1} + 0xf9x^{2^2} + 0x25x^{2^3} + 0xf4x^{2^4} + 0x01x^{2^5} + 0xb5x^{2^6} + 0x8fx^{2^7}$, $S : w \mapsto \sum_{i=0}^{7} \lambda_i w^{255-2^i} + 0x63$, modified: $S : w \mapsto \sum_{i=0}^{7} \lambda_i w^{-2^i}$. **Rank of system** is $\frac{equations}{monomials}$.

**Equation Solving:** If $n$ =number of equations, $M$ = number of variables. Solution takes $2^n$, if $n = m$, $n$, if $n = m + 1$ and $\sqrt{n}$ if $m >> n$.

**Buchberger:**

```
Input: $F = \{f_1, f_2, \ldots, f_m\}$. Output: Grobner $G = \{g_1, g_2, \ldots, g_s\}$.
$G \leftarrow F$;
Do {
    $G' \leftarrow G$;
    for($p, q \in G', p \neq q$) {
        Compute $S(p, q)$;
        $r \leftarrow REM(S(p, q), G')$;
        if($r \neq 0$) {
            $G' \leftarrow G' \cup \{r\}$;
            }
        }
    } while($G! = G'$)
```

Theorem: Foregoing algorithm yields Grobner Basis.

**F4/F5:** Grobner by matrix reduction. Example: $f_1 = 3x^3yz - 5xy$, $f_2 = 5x^2z^2 + 3xy + 1$, $g_1 = xy - 2z$, $g_2 = x^2z - 3yz$.

| | $x^3yz$ | $x^2z^2$ | $yz^2$ | $xy$ | $z$ | $1$ |
|---|---|---|---|---|---|---|
| $f_1$ | 3 | 0 | 0 | $-5$ | 0 | 0 |
| $f_2$ | 0 | 5 | 0 | 3 | 0 | 1 |
| $x^2zg_1$ | 1 | $-2$ | 0 | 0 | 0 | 0 |
| $1g_1$ | 0 | 0 | 0 | 1 | $-2$ | 0 |
| $zg_2$ | 0 | 1 | $-3$ | 0 | 0 | 0 |

Complexity of F5 is ${N_D}^{\omega}$ where $N_D$ is the size of the largest matrix containing polynomials of degree $D$. If $m = n, D \approx .09n$.

| Condition | Complexity |
|---|---|
| $m = an$ | exponential in $n$ |
| $n << m << n^2$ | subexponential in $n$ |
| $m = an^2$ | polynomial in $n$ |

**AES Design Criteria:** Invertibility, minimize largest non-trivial correlation between input and output, minimize largest non-trivial xor, complexity of algebraic expressions, Simplicity of expression. Estimation of linearly independent equations for XSL on AES-128.

**XL: The Extended Linearization.**

```
Input: $F = \{f_1, f_2, \ldots, f_m\}$.
Output: univariates.
$S \leftarrow \emptyset$;
Pick $D = d + 1$;
$G \leftarrow F$;
for($i = 1; i \leq n + 1; i + +$) {
    Generate $p_{\beta j} = x^{\beta} f_j, f_j \in F$;
    Do Gaussian reduction.
    If there is a univariate $f(x)$ {
        Solve;
        $S \leftarrow S \cup \{(x - a_i)\}$;
        Substitute.
        }
else
        $D \leftarrow D + 1$;
}
```

For each round ($0 \leq i \leq 9$) and each S-box ($0 \leq j \leq 15$), we get $r = 8 \times 3 = 24$ quadratics. $S$: Total S-boxes, $P-1$: passive S-Boxes, Highest degree: $2P$. $R$: Equations. $B$: S-boxes/round. $|R| = \binom{S}{P}(t^P - (t-r)^P)$, $|R'| = \binom{S}{P-1} SB(N_r+1)(t-r)^{P-1}$, $|R''| = \binom{S}{P-1}(S_k - L_k)(N_r+1)(t-r)^{P-1}$, $L_k$: independent key variables, $S_k$: key variables. Total terms: $T = \binom{S}{P} t^P$. For $P = 2$, $(R + R' + R'') = 33,665,888, T = 33,788,100$. For $P = 3$, $(R + R' + R'') = 95.18 \times 10^9, T = 91.9 \times 10^9$.

**Saturation Attack:** $\Lambda$-set has 256 states which are either all the same in a byte position or all different. In either case $\bigoplus_{x \in \Lambda} x_{i,j} = 0$. Mixcolumn is the only operation that changes this condition and only if there is more than one active byte in the column. To capitalize on this at final round (where mixing disrupts condition), guess key byte. If condition holds, it's right; otherwise it isn't.

**Boomerang:** $E = E_1 E_0$. $E_0 : \alpha \rightarrow \beta, p$, $E_1 : \gamma \rightarrow \delta, q$. (1) Pick $P_1 \oplus P_2 = \alpha$; (2) Ask for $C_1 = E(P_1), C_2 = E(P_2)$; (3) Compute $C_3 = C_1 \oplus \gamma$, $C_4 = C_2 \oplus \gamma$; (4) Request $P_4 = E^{-1}(C_4)$, $P_3 = E^{-1}(C_3)$. $E_0(P_1) = I_1$, $E_0(P_2) = I_2$, $E_0(P_3) = I_3$, $E_0(P_4) = I_4$. $E_1(I_1) = C_1$, $E_1(I_2) = C_2$, $E_1(I_3) = C_3$, $E_1(I_4) = C_4$. What is probability that $P_3 + P_4 = \alpha$. $e_1 : Pr[I_1 + I_3 = \gamma] = q$, $e_2 : Pr[I_2 + I_4 = \delta] = q$. $Pr[e_1 \wedge e_2] = q^2$. $Pr[I_3 + I_4 = \beta] = q^2$, $Pr[P_3 + P_4 = \alpha] = p^2 q^2$. If $(pq)^2 > 2^{-n}$, $pq > 2^{\frac{-n}{2}}$.

**Amplified Boomerang:** Use two short differentials instead of one differential. Start with quartet $P_1 \oplus P_2 = P_3 \oplus P_4 = \alpha$, each has $\alpha \rightarrow \beta$ with probability $p$. $E_0(P_1) \oplus E_0(P_2) = E_0(P_3) \oplus E_0(P_4) = \beta$. $E_0(P_1) \oplus E_0(P_3) = E_0(P_2) \oplus E_0(P_4) = \gamma$. $C_2 \oplus C_4 = C_1 \oplus C_3 = \delta$ and we want to use $\gamma \rightarrow \delta$. Probability that quartet becomes right is $\binom{Np}{2} 2^{-n} q^2$. Distinguishers count quartets $((P_1, P_2), (P_3, P_4))$ satisfying $C_1 \oplus C_3 = C_2 \oplus C_4 = \delta$.

**Bilinear Attack:** Notation: $L_r[0, 1, 2, \ldots, n-1]$, $R_r[0, 1, 2, \ldots, n-1]$ are the input to round $r$ and $I_r[0, 1, 2, \ldots, n-1]$, $O_r[0, 1, 2, \ldots, n-1]$ are the input (without key) and output to the round functions. If $\alpha \subseteq \{0, 1, 2, \ldots, n-1\}$, define $L_r[\alpha] = \bigoplus_{s \in \alpha} L_r[s]$. Consider the bilinear $L_{r+1}[\beta] \cdot R_{r+1}[\alpha] \oplus R_r[\beta] \cdot L_r[\alpha] = I_r[\beta] \cdot O_r[\alpha]$.

**Square/Integral:** Gives one linear combination of 4 key bits in round 4. Properties of sets of texts preserved by encryption. Example: 256 plaintexts that agree on 15 input bytes. $\theta$ - linear map, $\gamma$ - non-linear transform, $\pi$ - byte transposition, $\sigma$ - key addition, $\Lambda$ - 256 active states, $\lambda$ - set of indices of active bytes. Then $\bigoplus_{b=\theta(a), a \in \Lambda} b_{i,j} = 0$. $\forall x, y \in \Lambda$, $x_{i,j} \neq b_{i,j}$ if $(i,j) \in \lambda$, $x_{i,j} = b_{i,j}$ if $(i,j) \notin \lambda$. $a_{i,j} = b_{i,j} \oplus S_\lambda[b_{i,j}] \oplus^4_{i,j}$; if the result is not balanced, key is wrong. (See saturation attack earlier.)

**Truncated differentials:** Suppose $g : GF(2)^n \times GF(2)^n \times GF(2)^m \rightarrow GF(2)^n \times GF(2)^n \times$ implements a Feistel cipher round that is $g(X, Y, Z) = (Y, f(Y, Z) \oplus X)$. The S/N ratio is $\frac{|K|p}{\gamma\lambda}$ where $p$ is the differential probability, $\gamma$ is the number of suggested keys and $\lambda$ is the ratio of non-discarded keys to all keys. A full differential $a' \rightarrow b'$ specifies all $n$ bits, a truncated differential specifies a subset of bits. Here is an example of its usefulness. Let $f(x) = x^{-1}$. It has non-linear order $n-1$. If $n$ is odd the map is differentially 2-uniform $p = 2^{1-n}$; if $n$ is even the map is differentially 4-uniform $p = 2^{2-n}$. For 3 rounds, the differential probability is $2^{3-2n}$ and the S/N is $2^{3-n}$. For $r > 3$ the attack can't succeed. For 2 rounds, $p = 2^{1-n}$ and the S/N is $2^{n+1}$ so the attack requires $2^n$ texts and is $O(2^{3n})$ but for $a' \neq 0$, there are only $2^{n-1}$ possible $b'$ and we get one bit of information — the S/N is $\frac{2^{2n}}{2^{2n-1}} = 2$. Let $f(x, k)$ be the non-linear function in a 5 round Feistel cipher with block size $2n$. Let $\alpha \neq 0$ be an input differential for which only a fraction, $W$, of all output differences are possible. Then a truncated differential attack requires $2L$ chosen plain-cipher pairs and is $O(L 2^{2n})$ where $L$ is the smallest integer: $W^L < 2^{-2n}$.

**Higher order differentials:** Define

$$\Delta_a^{(1)}(f(x)) = f(x+a) - f(x), \Delta^{(i)}_{a_1, a_2, \ldots, a_i}(f(x)) = \Delta^{(1)}_{a_i}(\Delta^{(i-1)}_{a_1, a_2, \ldots, a_{i-1}}(f(x))).$$

Let $L[a_1, a_2, \ldots, a_i]$ is the set of all linear combinations of $< a_1, a_2, \ldots, a_i >$. Then $\Delta^{(i)}_{a_1, a_2, \ldots, a_i}(f(x)) = \sum_{\gamma \in L[a_1, a_2, \ldots, a_i]} f(P + \gamma)$ and $ord(\Delta_a^{(1)}(f(x))) \leq ord(f(x)) - 1$. Here is an example application. Let

$f(x,k) = (x+k)^2 \pmod p$ be the Feistel round function with size is $lg(p)$. $f$ is differentially 1-uniform and the round differential has probability $\frac{1}{p}$, $f''(x)$ is constant. The first order differential attack on a 5 round cipher requires $2p$ texts and is $O(p^3)$; a second order differential attack requires 8 texts and is $O(p^2)$. [Use $\Delta_{\alpha,\beta}(f(x)), \alpha = a||0, b = b||0, S/N = r^2$]. For a 5 round Feistel with $f$ non-linear of degree $r$ using an $r$th order differential requires $2^{r+1}$ texts and is $O(2^{2n+r})$.

**SFLASH attack:** The idea of SFLASH is to hide an easy-to-invert quadratic map, $F(x)$ with two "secret" invertible linear transformations $U, T$. If $e = q^i + q^j$, $F(x) = x^e$ is quadratic; in particular, if $e = q^\theta + 1$ (and from now on, it is) and $P = T \circ F \circ U$, $F$ is (easily) invertible if $(q^\theta + 1, q^n - 1) = 1$ (so $q = 2^k$) but without knowledge of $U, T$, $P$ isn't. This is the $C^*$ scheme Patarin broke. If we remove $r$ of $n$ quadratic equations in the base field that represent $P$, Patarin's attack doesn't work and the new scheme $C^{*-}$ can be used for signatures. Let $\Pi : (x_1, x_2, \ldots, x_n) \mapsto (x_1, x_2, \ldots, x_{n-r})$. $P$ is public key; to sign $m$, choose $r$ coordinates at random. Signer recovers $s$: $P_\Pi(s) = \vec{r}$. Signature is $(m, s)$. The idea of Shamir's attack is to use a multiplicative property of the linear transformation induced by a field element, $\xi$, on the differential to obtain a different set of linear combinations of the $F$ quadratics and then apply Patarin's attack. Define the differential $DF(a,x) = F(x+a) - F(x) - F(a) - F(0)$. For $F(x) = x^e$, $e = q^\theta + 1$ in field of characteristic $q$, $DF(\xi \cdot a, x) + DF(a, \xi \cdot x) = (\xi + \xi^{q^\theta}) DF(a,x)$. Denote $M_\xi$ as the matrix for the linear transformation induced by multiplying by $\xi$, $L(\xi)$ as the matrix induced by $\xi + \xi^{q^\theta}$ and $\Lambda(L(\xi)) = T_\Pi M_{L(\xi)} T^{-1}$. Let $Q$ be the space of quadratic forms, $V$ the subspace generated by $TFU$ and $V_\Pi$ the space generated by $T_\Pi FU$. $V_\Pi \subseteq V \subseteq Q$. There is a corresponding set of bilinear forms $B$, and sets $W$ and $W_\Pi$ and setting $N_\xi = U^{-1} M_\xi U$, the relation $DP(N_\xi(a), x)) + DP(N_{a,\xi}(x)) = \Lambda(L(\xi)) DP(a,x)$ holds. This equation relates unknown coefficients of $N_\xi$ on the left with unknown coefficients of $\Lambda(L(\xi))$ on the right. Setting $S_M(a,x) = DP_\Pi(N_\xi(a), x)) + DP_\Pi(N_{a,\xi}(x))$ we note the LHS is in $W_\Pi$ with probability $q^{-r}$ if $M$ represents a matrix for some $\xi$ induced value and probability $q^{n^2/2}$ if not. These identify transforms that can produce other $P$ equations to fill out the $r$ unknown quadratics to apply Patarin. SFLASH-1 parameters: $q = 2^7, n = 37, \theta = 11, r = 11$; SFLASH-2 parameters: $q = 2^7, n = 67, \theta = 33, r = 11$.

**Impossible differentials:** Suppose $\alpha \rightarrow \beta$ for $E_1$ is impossible and $E = E_2 \circ E_1 \circ E_0$. Encrypt many plaintexts with possible output $\alpha$ after $E_0$ and decrypt pairs with all possible subkeys through $E_2$. If these suggest $\alpha \rightarrow \beta$ the keys are impossible.

**Related Key Attacks:** If $K \rightarrow (K_1, K_2, \ldots, K_r)$ and $K^* \rightarrow (K_2, \ldots, K_r, K_1)$ and $F(X, K_i)$ is the round function then $n-1$ of the rounds are identical. If $P^* = F(P, K_1)$ and we know $2^{n/2}$ P/C pairs $(P,C)_K$ and $2^{n/2}$ P/C pairs $(P^*, C^*)_{K^*}$ try to solve $F(P, K') = P^*$ and $F(C, K') = C^*$; this gives $K_1$. Related key differential: $\alpha \rightarrow \beta$ for $E^0$ with $p > 2^{-n}$ then $Pr_{X,K}[E^0_K(X) \oplus E^0_{K \oplus \Delta K}(X \oplus \alpha) = \beta] = p > 2^{-n}$.

Structural: Prior to MixCol $(x^i_0, x^i_1, x^i_2, x^i_3)^T$ and after $(y^i_0, y^i_1, y^i_2, y^i_3)^T$ then $y^0_0 \oplus y^0_1 \oplus \ldots \oplus y^0_{255} = 00$.

**Slide Attack:** Let $F$ be a per-round function. If $C = E(P) = F^n(P)$ and $P' = F(P)$ then $C' = E(P') = F(C)$. Effective against rounds which implement weak permutations.

**Wiedemann:** Solve $A\vec{x} = \vec{b}$ in $O(n\omega)$ time over $F = GF(q)$ where $\omega$ is the number of non-zero elements of $A$. Let $S =< A^i b >, det(A) \neq 0$ and suppose $f(z) = \sum_{j=0}^{d}$ is the minimal polynomial normalized so the trailing coefficient $(f_0)$ is 1. Let $x = -\sum_{i=1}^{d} f_i A^{i-1} b$. Then $Ax = (1 - f(A))b = b$ so $x$ is a solution, this requires $2n(\omega + 1)$ field operations. To find $f$, look at the **linear recurrent sequence** $s_i = (u, A^i b)$, the associated polynomial $f_u | f$ and can be computed from the first $2n$ terms is $O(n^2)$.

Let $F = GF(q)$. Every $k$th order linear recurrent sequence is ultimately periodic with period $r$ satisfying $r \leq q^k$ ($r \leq q^k - 1$ if homogeneous). If $s_{n+k} = a_{k-1} s_{n+k-1} + \ldots + a_0 s_n$ the associated matrix is $A = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & a_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 0 & 1 & a_{n-1} \end{pmatrix}$ and the least period divides $A^k - 1$. If $D^{(r)}_n =$

| $j$ | $g_j(x)$ | $h_j(x)$ | $m_j$ | $b_j$ |
|---|---|---|---|---|
| 0 | $1$ | $x$ | 0 | 0 |
| 1 | $1$ | $x^2$ | 1 | 2 |
| 2 | $1+x^2$ | $2x$ | $-1$ | 1 |
| 3 | $1+x+x^2$ | $2x^2$ | 0 | 0 |
| 4 | $1+x+x^2$ | $2x^3$ | 1 | 2 |
| 5 | $1+x+x^2+2x^3$ | $2x+2x^2+2x^3$ | $-1$ | 2 |
| 6 | $1+x^3$ | $2x^2+2x^3+2x^4$ | 0 | 1 |
| 7 | $1+x^2+2x^3+x^4$ | $x+x^4$ | 0 | 1 |
| 8 | $1+2x+x^2+2x^3$ | - | 0 | - |

Figure 3.1: Berlekamp-Massey for $G(x) = 1 + x + x^4 + x^6 + x^7 \in F_2[x]$

$\begin{pmatrix} s_n & s_{n+1} & s_{n+2} & \cdots & s_{n+r-1} \\ s_{n+1} & s_{n+2} & s_{n+3} & \cdots & s_{n+r-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ s_{n+r-1} & s_{n+r} & s_{n+r+1} & \cdots & s_{n+2r-1} \end{pmatrix}$ then $s_0, s_1, \ldots$ is a linear recurrent sequence iff $D_n^{(r)} = 0$ for all but finitely many $n \geq 0$. If a linear recurrent sequence has minimal polynomial $m(x)$ of degree $\leq k$ and $r = \lfloor k + \frac{1}{2} - \frac{1}{2} m_{2k} \rfloor$ then $m(x) = x^r g_{2k}(\frac{1}{x})$ and $m(x)$ depends only on the first $2k$ terms.

```
Wiedemann's Algorithm
1. Set b[0]= b, k=0, y[0]= 0, d[0]= 0
2. If b[k]=0, x= -y[k].  Terminate.
3. Select u[k+1] at random
4. Compute first 2(n-d[k]) terms of (u[k+1], A**i b[k])= s[0,..]
5. Set f[k+1](z)= minimum poly in 4
6. Set y[k+1]= y[k]+f[k+1](z) b[k], b[k+1]= b[0]+A[y[k+1]), d[k+1]= d[k]+deg(f[k])
7. k= k+1, go to 2
```

```
Berlekamp's Algorithm
Given s[0], s[1], ... with generating function G(x)= s[0] + s[1]x + ... + s[i] x**i + ... in
F=GF(q)

1. g[0](x) = 1, h[0](x)=x, m[0]= 0
2. b[j]= coefficient of x**j in G(x) g[j](x)
   g[j+1]= g[j](x)- b[j] g[j](x),
   h[j+1] = 1/b[j] x g[j(x), if b[j] !=0 and m[j] >=0; x h[j](x), otherwise
   m[j+1]= -m[j], if b[j] !=0 and m[j] >=0; m[j+1]+1, otherwise
```

```
Version 2
Input: F=GF(q), 2n coefficients of a Linear recurrence <a[0], a[1], ..., a[2n-1]>
Output:  Minimal polynomial P

   R0=x**(2n); R1= a[0]+a[1]x+ ... + a[2n-1] x**(2n-1); V0=0; V1=1;
   while(n<=deg(R1) {
        R0= QR1+R;  // Division Algorithm
        V= V0-Q V1;
        V0= V1; V1= V; R0= R1; R1= R;
        }
   d= max(deg(V1), 1+deg(R1));
   P= x**d V1(1/x);
   return(P/leading-coeff(P));
```

## 3.11 Quantum Crypto

**Key Distribution:** Choose two basis: $B_1 = (|0>, |1>)$ and $B_2 = (|\frac{1}{2}>, |-\frac{1}{2}>)$. Alice chooses a random sequence of basis $\beta_i$ from $\{B_1, B_2\}$ and a random sequence of bits $b_i$ and encodes $b_i with \beta_i$. Bob chooses a random sequence of basis $\beta_i$ from $\{B_1, B_2\}$ and obtains sequence $c_i = \beta_i b_i$. Bob reveals his sequence of basis choices and then Alice reveals hers. Each confirms subset of bits for which the sequences agree using some classical system.

Consider three polarizers $A, B, C$ which have phases $0, 45, 90$. If $A$ and $C$ are placed in series, no light comes through but if $A$, $B$ and $C$ are placed in series, some light gets through. Let $|0>, |1>$ be two orthogonal vectors in a complex $2-$dimensional space. A **qubit** is a unit vector in this space. It can have many basis. **Shor:** Choose $m : n^2 \leq 2^m < 2n^2$ and let $v = \frac{1}{\sqrt{2^m}}(|0> + |1> + \ldots + |2^m - 1>$. Let $f$ be a function and $x = \frac{1}{C} \sum |x>$. System computes $t = \frac{1}{C} \sum |x, f(x)>$. If $f(x) = a^x \pmod{n}$, measurement of last $\frac{m}{2}$ bits fixes sequence $t = \frac{1}{C} \sum |x, u = f(x)>$ for fixed $u$ measuring the Fourier transform identifies period, that is $m : a^i = a^{i+r}$ so that $a^r = 1 \pmod{n}$ but that means $r$ is a universal exponent and we can (probably) factor $n$. **Universal exponent method:** Suppose $a^r = 1 \pmod{n}, \forall a : (a, n) = 1$. Put $r = 2^k m$, $m$ odd. Choose $a$ at random if $(a, n) \neq 1$, we have a factor; otherwise, put $b_0 = a^m \pmod{n}$ and $b_{n+1} = b_n^2 \pmod{n}$. If $b_0 = 1$ or $b_j = -1 \pmod{n}, 0 \leq j < k$ or $b_{j+1} = 1 \pmod{n}$ and $b_j = -1 \pmod{n}$, stop and pick a new $a$. If $b_{j+1} = 1 \pmod{n}$ but $b_j \neq \pm 1 \pmod{n}$ then $(b_j - 1, n)$ is a factor.

## 3.12 Protocols, Models

**Bell-Lapadula (BLP):** Subjects and Objects labeled. Simple Security property: S can read O iff $L(O) \leq L(S)$. *-Property: S can write O iff $L(S) \geq L(O)$. Tranquility: Labels never change. Biba: S can write O iff $I(O) \leq I(S)$. S can read O iff $I(S) \leq I(O)$.

**Perfect Forward Security** and ephemeral Diffie-Hellman with authentication. Both Alice and Bob agree on modulus $p$ and base $g$. Alice picks secret $a$ and Bob $b$ for signing; signing public keys have been previously exchanged. To for session key, Alice picks random $x$ and Bob picks random $y$. In the protocol below, $r_A = g^x \pmod{p}$, $r_B = g^y \pmod{p}$ and $K = g^{xy} \pmod{p}$. (1) $A \rightarrow B$ : "Alice", $r_A$. (2) $B \rightarrow A$ : "Bob", $r_B$, $E_K(sig_B(r_A, r_B))$ (3) $A \rightarrow B : E_K(sig_A(r_A, r_B))$. Throwing away $r_A, r_B, x, y$ yields perfect forward secrecy.

**Kerberos:** $L$ is lifetime. $T_X$ is the timestamp from $X$.

1. $A \rightarrow S$: $A, B$
2. $S \rightarrow A$: $\{T_S, L, K_{AB}, B, \{T_A, L, K_{AB}, A\}_{K_{BS}}\}_{K_{AS}}$.
3. $A \rightarrow B$: $\{T_S, L, K_{AB}, A\}_{K_{BS}}, \{A, T_A\}_{K_{AB}}$.
4. $B \rightarrow A$: $\{T_A + 1\}$.

Protocol layers: Application (DNS, TLS, HTTP, SSH), Transport (TCP), Network (IPv4), Link (ethernet, Wi-Fi).

**TLS:** Three phases: (1) Peer negotiation, (2) PK based key exchange (including certificate exchange), (3) encrypted traffic. TLS exchanges records each record has a content-type and MAC; all records are numbered. Content type 22 is handshake. Results in 2 encryption keys, 2 integrity keys and 2 IV's.

M1: $(C \rightarrow S)$ ClientHello(Client-random[28], cipher-suites, compression methods, highest protocol version),

M2: $(S \rightarrow C)$ ServerHello(ServerRandom[28], cipher-suite, certificates),

M3: $(C \rightarrow S)$ ClientKeyExchange(E(PkS, Pre-Master Secret), MD5-SHA1(M1 —— M2—— M3A)), [Master Secret is PRF(Pre-master secret, "master secret", ClientRandom —— ServerRandom)],

M4: $(S \rightarrow C)$ Finish MD5-SHA1(M1 —— M2 —— M3A —— M3C).

**IPSEC:** Two protocols: securing packets and key negotiation. Two modes: transport and tunnel. In transport mode only payload is encrypted. Packets can be secured for authentication and integrity only (AH) or authentication, confidentiality and integrity. IKE Phase1: CP (crypto proposed), CS (crypto selected), IC (initiation cookie), RL (response cookie), $K = h(IC, RC, g^{ab} \pmod p, R_A, R_B)$. $SKEYID = h(R_A, R_B, g^{ab} \pmod p)$. $Proof_A : [h(SKEYID, g^a \pmod p, g^b \pmod p, IC, RC, CP, \text{``}Alice''\text{]}_{Alice}$. Public Key: (1) $A \rightarrow B$ : IC, CP. (2) $B \rightarrow A$ : IC, RC, CS. (3) $A \rightarrow B$ : IC, CP $g^a \pmod p, \{R_A\}_{Bob}, \{\text{``}Alice''\}_{Bob}$. (4) $B \rightarrow A$ : IC, CP $g^b \pmod p, \{R_B\}_{Alice}, \{\text{``}Bob''\}_{Alice}$. (5) $A \rightarrow B$ : IC, CP $E(Proof_A; K)$. (6) $B \rightarrow A$ : IC, CP $E(Proof_B; K)$.

**Fiat-Shamir:** Prove knowledge of a secret, $s$, where $v = s^2 \pmod n$, $n = pq$; $v, n$, public. $A$ proves she knows $s$: (1) $A$ picks $r$ at random and computes $x = r^2 \pmod n$ — commitment, (2) $B$ chooses $e \in \{0, 1\}$ at random and sends $e$ to $A$ — challenge, (3) $A$ computes $y = rs^e \pmod n$ and sends it to Bob — response, (4) finally, $B$ verifies $y^2 = r^2 s^{2e} = xv^e \pmod n$ — verify this.

S/Mime: Todo.

```
DSig:
<Signature>
  <SignedInfo>
    <CanonicalizationMethod/>
    <Reference URI=?>
      <Transforms/>
      <DigestMethod/>
      <DigestValue/>
  </SignedInfo>
  <SignatureValue/>
  <KeyInfo/>
  <Object>
</Signature>

XML Encryption:
<EncryptedData>
  <EncryptionMethod/>
  <KeyInfo>
    <AgreementMethod/>
    <KeyName/>
    <RetrievalMethod/>
  </KeyInfo>
  <CipherData/>
</EncryptedData>

SAML: Authn/AuthZ Request/Response over SOAP.
      Assertion, conditions, advice.
XACML: Authorization Rules:
      Subjects, Resources, Actions.
REL: Grant, Principal, Right, Resource, Condition.
WS-Policy: security policy
WS Trust: Trust
WS-Privacy including WS-Secure
   Conversation, Federation.
WS-Authorization: Principal, Claim, Token.
```

More Timings: P4, 2.1 GHz. AES: 44 operations/round.

| **Algorithm** | **Key Size** | **Speed(MB/sec)** | **Algorithm** | **Key Size** | **Speed(MB/sec)** |
|---:|---:|---:|---:|---:|---:|
| DES | 56 | 21 | 3DES | 168 | 9.8 |
| SHA-1 | NA | 68 | SHA-256 | NA | 44 |
| TEA | 64 | 23 | AES | 128 | 61 |

**Reestimation:** Rotor modeled by $S(r_j, R) = C^r R C^{-r}$ and represented by a $q \times q$ permutation matrix. Key space is $D_1 \times D_2 \times \ldots \times D_k$, $D_i$ is all $q!$ permutation matrices. $\chi^{cs} = {\chi_1}^{cs} \times {\chi_2}^{cs} \times \ldots \times {\chi_k}^{cs}$, ${\chi_i}^{cs}$ is all possible $q \times q$ stochastic matrices. Suppose $\vec{p}$ is plaintext distribution. $d(r, x) = S(r, x)\vec{p}$. Likelihood $L(X|\{c, r\}) = Pr(ciphertext = \{c_1\}^N | \{r_1\}^N; X) = \prod_{n=1}^{N} {e_{c(n)}}' d(r(n); X)$. Want to maximize $L$ by adjusting $X$. The MLE of $X$ exists and is strongly consistent. Use the following result: Let $P(z)$ be a polynomial with non-negative coefficients homogeneous of degree $d$ in $z_{ij}$, $\mathcal{Z} = \{z_{ij} : z_{ij} \geq 0, \sum_i^{q_j} z_{ij} = 1\}$. $\mathcal{T}(z)_{ij} = z_{ij} \frac{\frac{\partial P}{\partial z_{ij}}}{\sum_i^{q_j} z_{ij} (\frac{\partial P}{\partial z_{ij}})_z}$. Computations requires is $\approx kq^2 N$ and a 2 rotor machine with $N = 1024$ ciphertext letters requires about 60 iterations.

## 3.13 Random Number Quality

Traditional approach for getting $n$ bit value: (1) Get large sample. (2) Calculate the relative frequency, $r_w$, of each word $w$ in $b$-bit block. (3) Estimate $H = -\sum_{w=0}^{2^r - 1} r_w lg(r_w)$. Repeat $\frac{n}{H}$ times. Total bits checked: $\lceil \frac{nb}{H} \rceil$ Concern: small set of possible values and deterministic mixing reduces entropy. Entropy is not the best measure of security. Consider the following **Theorem.** The entropy of a source $P =< p_1, p_2, \ldots, p_N >$ which is mixed by $F : [1..N] \rightarrow [1..m]$ is greater than the entropy of the mixed sequence. Let $Prob(O = j) = q_j$ and $Q =< q_1, q_2, \ldots, q_m >$. $H_{out} = H_Q = -\sum_{j=1}^{m} q_j lg(q_j) = -\sum_{j=1}^{m} [\sum_{f(i)=j} p_i] lg([\sum_{f(i)=j} p_i]) = -\sum_{i=1}^{N} p_i lg(p_i + S_i) < H_p = H_{in}$, where $S_i = \sum_{j \neq i, F(i)=F(j)} p_j$ for the standard Shannon entropy $H_Q = -\sum_{i=1}^{N} p_i lg(p_i)$. Suppose $T$ values are required in cryptoperiod; if $Q =< q_1, q_2, \ldots, q_m >$ is the distribution and $q_{i_1} \geq q_{i_2} \geq \ldots \geq q_{i_m}$, adversary's best strategy is to guess $< i_1, i_2, \ldots >$ until success. This motivates a different entropy measure. Define $H_\alpha(Q) = \frac{1}{1-\alpha} \sum_{j=1}^{M} {q_j}^2$. $H_2(Q)$ is a good measure for collision resistance (not secrecy) since $\sum_{j=1}^{m} {q_j}^2 = 2^{-H_2(Q)}$; the waiting time for repeats is $\sqrt{\pi 2^{H_2(Q)-1}}$. $H_\infty(Q)$ is a good measure for the quality of resulting key generation, since the expected cost of the guessing attack is $\frac{1}{2q_{max}} = 2^{H_\infty(Q)-1}$. As an example, consider the distribution, $Q$ over 128 bit quantities consisting of one value that occurs with probability $2^{-80}$ and is otherwise flat. $H_2(Q) \approx 128$, $H_\infty(Q) \approx 80$.

If $X$ is a event with $n$ possible outcomes having respective probabilities $p_1, p_2, \ldots, p_n$ the **min-entropy** of $X$ is $H_\infty(X) = min_{1 \leq i \leq n} - lg(p_i) = -lg(max_i(p_i))$. To get an estimate of the min-entropy or $W(Q)$, we need $S(Q)$. Suppose randomizer produces $m = 2^n$ outputs with probability distribution $Q =< q_1, q_2, \ldots, q_m >$. Quality of $Q$ is $S(Q) = \sum_{j=1}^{m} q_j^2 \geq \frac{1}{m}$ which is the probability of repeated output. $W(Q) = \sum_{j=1}^{m} j q_j \leq \frac{m+1}{2}$ which is the adversary's work factor. Estimating either $H_2(Q)$ or $H_\infty(Q)$ consists of four steps. (1) Form Markov model of input source data (over $L$ consecutive samples), (2) Compute source data repeat probability, (3) Estimate $S(Q)$, (4) Use $S(Q)$ to estimate lower bound on $W(Q)$ and/or $H_\infty(Q)$.

**Entropy Order Paradox:** Consider $Q_1 =< 0.258, 0.116, 0.146, 0.032, 0.140, 0.266, 0.038, 0.004 >$ and $Q_2 =< 0.256, 0.232, 0.076, 0.130, 0.006, 0.157, 0.005, 0.129 >$. $H(Q_1) = 2.54542$ and $H(Q_2) = 2.54495$ but $S(Q_1) = 0.194176$ and $S(Q_2) = 0.188076$ while $W(Q_1) = 2.844$ and $W(Q_2) = 2.903$.

**Step 1 - Markov Model:** The model consists of $\Theta =< \theta_1, \theta_2, \ldots, \theta_L >$ states where where $\rho$ is the initial probability distribution, and $T = \begin{pmatrix} \tau_{1,1} & \tau_{1,2} & \cdots & \tau_{1,s} \\ \tau_{2,1} & \tau_{2,2} & \cdots & \tau_{2,s} \\ \cdots & \cdots & \cdots & \cdots \\ \tau_{s,1} & \tau_{s,2} & \cdots & \tau_{s,s} \end{pmatrix}$ is the transition matrix. The procedure is to (1) Model source as sequence of states $< s_1, s_2, \ldots, s_s >$, (2) Get $\rho$ (use steady state estimate), (3) Determine state defining bits. For multiple sources, $\Theta^{(k)} =< \theta_1, \theta_2, \ldots, \theta_{i_k} >$ and $\sum_{i=1}^{N} {p_i}^2 = \sum_{i_1=1}^{N_1} (p_{i_1}^{(1)} p_{i_2}^{(2)} \ldots p_{i_k}^{(k)})^2$. **Step 2 - Compute source data repeat probability:** $\sum_{j=1}^{N} {p_j}^2 = [\rho_1, \rho_2, \ldots \rho_s] T [1, 1, \ldots, 1]^T$. **Step 3 - Estimate $S(Q)$:** $S(Q) = \sum_{j=1}^{m} q_j^2 = \frac{1}{m}(1 + \epsilon_s)$ where $(1 + \epsilon_s) = (m-1) \sum_{i=1}^{N} p_i^2$. **Step 4 (for $H_2$) - Estimate $W(Q)$ using $S(Q)$ for $L$ source inputs:** To get the best possible bound on $W(Q)$ given $S(Q)$ ($q_j$ unknown): Let $m' = min(m, \frac{3S(Q)+4+\sqrt{9S(Q)^2+16}}{6S(Q)}) \approx min(m, \frac{4}{3S(Q)})$ then $W(Q) \geq B$ where $B = \frac{1}{6}(3m' + 3 - \sqrt{3(m'^2 - 1)(m'S(Q) - 1)})$. To obtain this result use Lagrange multipliers to minimize $W(Q)$ subject to $\sum_{j=1}^{m} {q_j}^2 = S(Q)$ and $\sum_{j=1}^{m} q_j = 1$. **Step 4 (for $H_\infty$):** Use Dynamic Programming compute $p_{max}$ or proceed as follows: Set $y_1 = F(x_1)$, $q_1 = Pr[y_1] = p_{max} + \sum_{i=2}^{N} p_i I_{i,1}$ and $q_j = \sum_{i=2}^{N} p_i I_{i,j}$. $\mu_1 = E[q_1] = \frac{1}{M}[1 + (M-1)p_{max}]$, $\mu_2 = E[q_j] = \frac{1}{M}[1 - p_{max}]$, $\sigma_1^2 = \sum_{i=2}^{N} {p_i}^2 Var(I_{i,1}) = (\frac{1}{M} - \frac{1}{M^2}) \sum_{i=2}^{N} {p_i}^2$ and for $j \geq 2$, $\sigma_j^2 = \frac{M-1}{M} \sum_{i=2}^{N} {p_i}^2$. $-lg(\mu_1)$ is a good estimate for $H_\infty(Q)$. Want $|-lg(\mu_1) - H_\infty(Q)| \leq \frac{1}{2} 10^{s-d+1}$, whereas $s$ is largest integer: $10^s \leq H_\infty(Q)$. If $Y$ is the number of $q_j$ exceeding $B$, $Pr[q_{max} \leq B] = 1 - Pr[Y > 0] \geq 1 - E[Y] > 1 - \epsilon$. $Pr[E_j] = Pr[z > \frac{\mu_1^{1-\frac{1}{2}10^{-d}}}{\sigma}, j > 1$. Put $B = max(\mu_1 + T_1\sigma, \mu_2 + T_2\sigma)$, where $z$ is normally distributed and $Pr(z > T_1) = \frac{\epsilon}{3}$ while $Pr(z > T_2) = \frac{\epsilon}{3(M-1)}$ then $Pr(\mu_1^{1+\frac{1}{2}10^{-d}} \leq p_{max} \leq \mu_1^{1-\frac{1}{2}10^{-d}}) \geq (1 - \epsilon)$.

Example ($L = 3$): Let $b_t, b_{t+1}, b_{t+2}$ be three successive states and $Prob(b_{t+2} = b_{t+1} \oplus b_t) = .8$ with $s = 4$ states then $T = \begin{pmatrix} .8 & .2 & 0 & 0 \\ 0 & 0 & .2 & .8 \\ .2 & .8 & 0 & 0 \\ 0 & 0 & .8 & .2 \end{pmatrix}$ and the initial distribution $\rho = (.25, .25, .25, .25)$. The state distribution is $\Theta =< \theta_1, \theta_2, \theta_3 >$. In SHA-1 mixing example, $\sum_{i=1}^{N} p_i^2 = 4.87 \times 10^{-44}$, $L = 256$ and we compute $S(Q) = \sum_{i=1}^{m} q_j^2 \approx \frac{1}{m}[1 + (m-1) \sum_{i=1}^{N} p_i^2]$. $m = 2^{160}$. $m = 2^{160}$, $m' = 2.74 \times 10^{43}$, $W(Q) \geq 9.1 \times 10^{42}$.

**Parameter Estimate:** $N = \alpha\Gamma^{L-1}U$, $\alpha$ is initial $\rho$, $\Gamma$ is initial $T$. $U = [1, 1, \ldots, 1]^T$. $I_{i,j}$ is 1 if $F(i) = j$ and 0 if $F(i) \neq j$. $q_j = \sum_{i=1}^{N} I_{i,j}p_i$, $E(q_j) = \sum_{i=1}^{N} p_i E(I_{i,j}) = \frac{1}{m}$. $Var(q_j) = \sum_{i=1}^{N} p_i^2 Var(I_{i,j})$. $Var(I_{i,j}) = \frac{1}{m} - \frac{1}{m^2}$. $Var(q_j) = \sum_{i=1}^{N} p_i^2 Var(I_{i,j}) = \frac{m-1}{m^2} \sum p_i^2$. $E(\sum_{j=1}^{m})q_j^2) = \sum_{i=1}^{N} E(q_j^2) = \sum_{j=1}^{N} E(q_j^2) + Var(q_j) = \frac{1}{m}(1 + (m-1)\sum_{i=1}^{N} p_i^2$.

**Extension to HMM:** Transition matrix $T = \tau_{i,j}$, $s$ states, $\vec{\rho}$ initial distribution, $\theta_t \in \{1, 2, \ldots, r\}$ is the output at time $t$, $C^{(n)} = (c_{i,j}^{(n)})$, $c_{i,j}^{(n)} = \sum_{\theta_1,\ldots,\theta_n} Pr(\theta_1, \ldots, \theta_n, \sigma_n = i)Pr(\theta_1, \ldots, \theta_n, \sigma_n = j)$. $\sum_{i,j} c_{i,j}^{(n)} = \sum_{i=1}^{N} p_i^2 = \sum_{\theta_1,\ldots,\theta_n} Pr(\theta_1, \ldots, \theta_n)^2$ and $C^{(n)} = (BB^T) \cdot (T^T C^{(n-1)} T)$ where $\cdot$ means elementwise multiplication. Recursion step requires $2s^3$ multiplications. $\approx 7$ minuses for 400 outputs without eigenvalue.

$f : G \to \mathbb{C}$, $g : G \to \mathbb{C}$, $E(f) = E(g) = 0$, $E(|f|^2) = E(|g|^2) == 1$. $S_{fg} = f(x) \cdot \overline{g(y)}$, $L_{ab}(X, Y) = \chi^a(x)\chi^{-b}(y)$. Imbalance of $S$: $I(S) = |E(S)|^2$, $\overline{I}(S) = \frac{1}{K} \sum_{k \in K} I(S|K = k)$. $C = (c_{ab}), c_{ab} = \overline{I}(L_{ab}(X, Y))$, $a, b \in G \setminus \{0\}$. Let $y = e_k(x) = x + k$, $c_{ab} = \delta(a \oplus b)$. $c_{ab} = \frac{1}{|K|} \sum_k |\mathcal{F}(\chi^b \cdot e_k)(a)|^2$, $\overline{I}(S) = \hat{f}^T C \hat{g}$. $\hat{f}_a = |\mathcal{F}(f)(a)|^2$, $\hat{g}_b = |\mathcal{F}(g)(b)|^2$. **Likelihood estimate of correlation:** $\tilde{I}(S) = |\frac{1}{N} \sum_{x,y} f(x)\overline{g(\tilde{y})}|^2$. $\xi(x, J, N)$ is the imbalance distribution with imbalance parameter $J$. $\xi(x, J, N)\frac{2N}{1-J}h(\frac{2N}{1-J}x, \frac{2N}{1-J}J)$ where $h(\cdot, s)$ is the probability density of $\chi^2$ with 2 degrees of freedom and skewness parameter $s$. $\xi(x, J, N) = \frac{N}{(1-J)\sqrt{\pi}} e^{-\frac{N(x+J)}{1-J}} \sum_{r=0}^{\infty} \sigma_r$ where $\sigma_r = \frac{1}{(2r)!}((\frac{2N}{1-J})^2 Jx)^r \frac{\Gamma(r+\frac{1}{2})}{\Gamma(r+1)}$. If $J << (\frac{1}{2N})^2$, $\xi(x, J, N) \approx h(x, J, N)$, $h(x, J, N) = \frac{N}{1-J} e^{-\frac{N(x+J)}{-1J}}$ with accumulated error $\epsilon = 1 - e^{-\frac{JN}{1-J}}$.

Let $S$ be an I/O product and $S_1, \ldots, S_N$ samples, $\tilde{I}(S) = |\frac{1}{N} \sum_{j=1}^{N} S_j|^2$. Let $E$ be an $n - 1 \times n - 1$ matrix with $E_{ij} = \frac{1}{n-1}$ and $C$ the truncated correlation matrix. $C^r - E = (C - E)^r$ and $\sigma_2(C) = \sigma_1(C - E)$ where $\sigma_k(M)$ is the $k$-th largest singular value of $M$. Let $D = C^T C = V^{-1}\Lambda V$, $\Lambda = diag(1, \sigma_2(C), \ldots)$. **Theorem:** Let each of the $r$ rounds of an interactive cipher have correlation matrix $C$ then $\overline{I}(S) \leq \frac{1}{n-1} + ||C - E||^r$; also, $\overline{I}(S) \leq \frac{1}{n-1} + \sigma_2(C)^r$, $\sigma_2(C) \leq min < (1 - \sum_b min_a(C^T C)_{ab})^{\frac{1}{2}}, (1 - \sum_a min_b(C^T C)_{ab})^{\frac{1}{2}} >$.

Let $\otimes$ be the Kroneker product. $\Phi(M, N) = \sum_{a,b} g_{ab} M^a \otimes N^b$, $\phi(x, y) = \sum_{a,b} g_{ab} x^a y^b$. The eigenvalues of $\Phi(M, N)$ are $\phi(\lambda_r(M), \lambda_s(N))$. if $C = A \otimes B$, the singular values of $C$ are products of the singular values of $A$ and $B$. The correlation of a non-keyed permutation $R = G \to G$ is $C = (F^* PF)(\overline{F^* PF})$ where $F = (f_{ab}))$, $f_{ab} = \frac{1}{\sqrt{n}}\chi^{-a}(b)$, $P = (p_{ab})$, $p_{ab} = \delta(a \oplus \phi(b))$. $C = U \cdot \overline{U}$ where $U$ is unitary. **Theorem:** The correlation matrix of a keyed permutation $e_k : G \to G$ is $C = \frac{1}{|K|} \sum_k C^{(k)}$, $C^{(k)} = U^{(k)}\overline{U}^{(k)}$, $U^(k) = FP^{(k)}F^*$, $P^{(k)} = \delta(a \oplus e_k(b))$.

# Chapter 4

# Physics

## 4.1 Basic Laws

**Classical Mechanics:** $\vec{F} = \frac{d\vec{p}}{dt}$, $\vec{p} = m\vec{v}$, $m = m_0\sqrt{1-(\frac{v}{c})^2}$, $F = -G\frac{m_1 m_2}{{r_{12}}^2}$. $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$. Special Relativity: If primed (') coordinate system moving at constant velocity $u$ in the $x$ direction with respect to the unprimed system, $x' = \frac{x-ut}{\sqrt{1-\frac{u^2}{c^2}}}$, $y' = y$, $z' = z$, $t' = \frac{t-\frac{ux}{c^2}}{\sqrt{1-\frac{u^2}{c^2}}}$.

**Maxwell's Equations:** $\nabla \cdot \vec{j} = -\frac{\partial\rho}{\partial t}$, $\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0}$, $\nabla \times \vec{E} = -\frac{\partial\vec{B}}{\partial t}$, $\nabla \cdot \vec{B} = 0$, $c^2\nabla \times \vec{B} = \frac{j}{\epsilon_0} + \frac{\partial\vec{E}}{\partial t}$, $c = \frac{1}{\sqrt{\mu_0\epsilon_0}}$.

**Solution to Maxwell Equations:** $E = -\nabla\phi - \frac{\partial A}{\partial t}$, $B = \nabla \times A$. **Gauge Transformation:** $A' = A + \nabla\psi$, $\phi' = \phi - \frac{\partial\psi}{\partial t}$. Choosing gauge $\nabla \cdot A = -\frac{1}{c^2}\frac{\partial\phi}{\partial t}$ in Maxwell's equations yields: $\nabla^2\phi - \frac{1}{c^2}\frac{\partial\phi}{\partial t} = \frac{-\rho}{\epsilon_0}$ and $\nabla^2 A - \frac{1}{c^2}\frac{\partial A}{\partial t} = \frac{-j}{\epsilon_0 c^2}$. Solving these produces $\phi(1,t) = \int \frac{\rho(2,t-(r/c))}{4\pi\epsilon_0 r_{12}}dV$, and $A(1,t) = \int \frac{j(2,t-(r/c))}{4\pi\epsilon_0 c^2 r_{12}}dV$. Lenart-Weichart: $\phi(1,t) = \frac{q}{4\pi\epsilon_0(r-\frac{v\cdot r}{c})_{retarded}}$.

**Fundamental constants:** $G = 6.671 \times 10^{-11}\frac{Nm^2}{kg^2}$, $c = 2.99725 \times 10^{10}\frac{cm}{s}$, $k = 1.38 \times 10^{-16}\frac{ergs}{mol-deg}$, $h = 6.6262 \times 10^{-27} erg - sec$, $q_e = 1.60219 \times 10^{-19}C$, $\epsilon_0 = \frac{10^7}{4\pi c^2} = 8.854 \times 10^{-12}\frac{C}{N-m^2}$, STP: $22.4 \times 10^3\frac{cm^3}{mol}$, $R = 8.3143\frac{J}{mol-deg}$ $N_0 = 6.022 \times 10^{23} mol^{-1}$.

**EMF:** total accumulated force through wire. Some consequences: $E = \frac{-q}{4\pi\epsilon_0}(\frac{e_{r'}}{r'^2} + \frac{r'}{c}\frac{d}{dt}\frac{e_{r'}}{r'^2} + \frac{1}{c^2}\frac{d^2}{dt^2}\frac{e_{r'}}{r'^2})$, $E = cB$. $E^2 - (pc)^2 = (mc^2)^2$, $ED = \frac{1}{8\pi}(E^2 + B^2)$, $S = \frac{1}{\mu_0}E \times B$, $E = cB$ for EM waves. For conservative electric field: $\Delta\phi = -\int_a^b qEds$, $\Delta V = \frac{\Delta\phi}{q}$. $E = -\nabla\phi$.

Gauss (always): $\Phi_E = \int_S E \cdot dA = \frac{q_{in}}{\epsilon_0}$, $S$, closed. $\Phi_B = \int_S B \cdot dA = 0$, $S$, closed. $B = \frac{\mu_0}{2\pi}\frac{qv \times e_r}{r^2}$. Ampere: $\int_C B \cdot dl = \mu_0(I_{enclosed} + \epsilon_0\frac{d\Phi_E}{dt})$, Faraday: $\mathcal{E} = \int_C E \cdot dl = -\frac{d\Phi_B}{dt}$. Biot-Savart (steady currents only): $dB = \frac{\mu_0}{2\pi}\frac{I \times dl}{r^2}$. $E = 0$ for conductor in electrostatics. $C = \kappa_0 C_0$, Steady current in conductor: $J = nqV_d = \sigma E$, $E = \rho J$. Wire (steady current): $B = \frac{\mu_0 I}{2\pi r}$. AC: $V = IZ$.

*Devices and circuits.* $\mathcal{E} = -L\frac{di}{dt}$. $Iz = V$. $Z_C = \frac{1}{i\omega C}$, $Z_L = i\omega L$, $Z_R = R$. **Low-pass** (Inductance in series, capacitance across EMF), **high-pass** (switch capacitance and inductance). **Reactive:** no real term. **Dissipative:** real term $> 0$. **Propagation factor:** $\alpha = \frac{V_{n+1}}{V_n}$. **Transmission line:** $\frac{\partial^2 I}{\partial x^2} = L_0 C_0 \frac{\partial^2 I}{\partial t^2}$ impedance is $z_0 = \sqrt{\frac{L_0}{C_0}}$. **Mutual Inductance:** $\mathcal{E}_2 = -M\frac{di_1}{dt}$, $\mathcal{E}_1 = -M\frac{di_2}{dt}$. $U_L = \frac{1}{2}LI^2$, $U_C = \frac{1}{2}CV^2$. **Kirchoff:** $\sum_k v_k = 0$, $k$ covers loop; $\sum_k i_k = 0$, $k$ covers node. **Thevinen equivalence:** Two terminal linear network is equivalent to voltage source $V_{Th}$ and impedance in series. **Norton equivalence:** Two terminal linear network is equivalent to current source $V_N$ and conductance $G_N$ in parallel. **Resistor:** $R = \frac{\rho L}{A}$. **Battery:** $\mathcal{E} - Ir_{internal} = V_{ab}$. **Op Amp:** $v_o = A_{OL}v_d$. Transfer and two terminal input and output.

**Reflection on string:** $\frac{\partial^2\psi}{\partial t^2} = \frac{T}{\rho}\frac{\partial^2\psi}{\partial x^2}$, $v_\phi = \frac{\omega}{k}$, $v_g = \frac{d\omega}{dk}$, $Z = \sqrt{T\rho}$. Power: $P(t) = F\frac{\partial\psi}{\partial t}$, For travelling wave: $P(t) = Z(\frac{\partial\psi}{\partial t})^2$. Consider a wave train on a string from the left ($L$) with a change at $x = 0$

of medium (i.e. a denser string) to a string on the right ($R$). For perfect termination: $F_{term}$(“R on L”) $= -Z_L \frac{\partial \psi_{inc}}{\partial t}(0,t)$. For excess force: $F_{term}$(“R on L”) $= Z_L \frac{\partial \psi_{ref}}{\partial t}(0,t)$. $-Z_L \frac{\partial \psi_{inc}}{\partial t}(0,t) + Z_L \frac{\partial \psi_{ref}}{\partial t}(0,t) = -Z_R(\frac{\partial \psi_{inc}}{\partial t}(0,t) + \frac{\partial \psi_{ref}}{\partial t}(0,t))$. So, $\frac{\partial \psi_{ref}}{\partial t}(0,t) = \frac{Z_L - Z_R}{Z_L + Z_R} \frac{\partial \psi_{inc}}{\partial t}(0,t)$.
$R = \frac{Z_L - Z_R}{Z_L + Z_R}$ is called the reflection coefficient. **Wave Transmission:** String: $v = \sqrt{\frac{F}{\mu}}$, Fluid: $v = \sqrt{\frac{B}{\rho}}$, Solid: $v = \sqrt{\frac{Y}{\rho}}$, Adiabatic Gas: $v = \sqrt{\frac{\gamma p}{\rho}}$. Standing wave transmits no energy. Oscillating Dipole (Antenna): $E = \frac{p_0 k^2}{4\pi\epsilon_0} \frac{sin(\theta)}{r} sin(\omega t - kr)$.

**Early Quantum Mechanics:** $\Delta p \Delta x \geq \frac{h}{4\pi}$, $\lambda = \frac{h}{p}$, $\nu = \frac{E}{h}$, $p = \frac{hk}{2\pi}$, $E = \frac{h\omega}{2\pi}$, $p_{av} = nkT$. **Black-body radiation:** $E(\lambda, T) = \frac{8\pi hc}{(\lambda^5)}(e^{(hc)/(\lambda kT)} - 1)^{-1}$. **Photoelectric Effect:** $hf = KE + \phi$. **Bohr hydrogen atom:** $E_n = -\frac{13.6ev}{n^2}$, $r_n = n^2 a_0$, $a_0 = \frac{h^2}{2\pi kmc^2} = .0529nm$. **Time Independent Schrodinger:** $\frac{d^2\psi}{dt^2} + \frac{4\pi m}{h}(E - U(x))\psi = 0$. **Schrodinger:** $\frac{ih}{2\pi}\frac{\partial \psi}{\partial t} = -\frac{h^2}{8\pi^2 m}\nabla^2\psi + V\psi$.

**Relativity:** Proper Interval: $I(x,y,z,t) = x^2 + y^2 + z^2 - c^2t^2$, $I(x,y,z,t) = I(x'y'z't')$. $ds^2 = g_{ij}dx^i dx^j$, $g_{ij} = g_{ji}$, $\delta \int ds = 0$. **Action:** $S = \int_{t_1}^{t_2} L(x,x',t)dt$, $\delta S = 0 \rightarrow \frac{\partial L}{\partial x} - \frac{d}{dx}\frac{\partial L}{\partial x'} = 0$. $L(x,x',t) = -m_0c^2\sqrt{1 - \frac{v^2}{c^2}} - q(\phi + v \cdot A)$. $R_{excess} = \sqrt{\frac{A}{4\pi}} - r_{meas} = \frac{G}{3c^2}M$, $\frac{G}{3c^2} = 2.5 \times 10^{-29}\frac{cm}{gm}$. From principle of equivalence, $\omega = \omega_0(1 + \frac{gH}{c^2})$ - doppler shift measured by Pound and Rebka.

## 4.2 Physical Constants

1 $in = 2.54\ cm$. 1 $meter = 39.370\ in$. 1 $AU = 1.496 \times 10^{11}\ m$. 1 $lb = 4.448\ N$. 1 $Pa = 1\ \frac{N}{m^2}$. 1 $Atm = 1.013 \times 10^5\ Pa$. 1 $hp = 745.7\ W$. 1 $J = 10^7\ erg$. 1 $ev = 1.602 \times 10^{-19}\ J$. 1 $BTU = 1055\ J$. 1 $cal = 4.186\ J$. 1 $L = 1000\ cm^3$. 1 $Gal = 3.785 \times 10^{-3}\ m^3$.

**Atomic constants:** $M_e = .510998Mev = 9.10939 \times 10^{-31}kg$, $M_p = 938.256Mev(= 1836M_e) = 1.67262 \times 10^{-27}kg$, $M_n = 939.55Mev = 1.67493 \times 10^{-27}kg$, $\sigma_{SB} = 5.67 \times 10^{-8}Wm^{-2}K^{-4}$, $1ev = 1.6 \times 10^{-12}erg = 1.6 \times 10^{-19}J$, $1curie = 3.7 \times 10^{12}decays$, $c_s = 3.32 \times 10^4 cm/s$, $1cal = 4.1855J$, $1BTU = 252cal$, $1kgTNT = 4.2MJ$, $1A = 10^{-8}cm$. **HDNA:** 2,900,000 kilobases.

**Astronomical constants:** $H_0 = 100km(s - Mpc)^{-1}$, 1 pc= 3.26 l-y, $10^{80}nucleons$, $10^{28}cm - diam$, $10^{11}$ galaxies.
**Milky Way:** $\epsilon_{ecliptic/MW} = 62.5$, $1.6 \times 10^{11}stars$, $10^{23}cm - diam$, $8 \times 10^{44}gm$.
**Sun:** $E_{sun} = 4 \times 10^{33}ergs/sec$, $R_{Sun} = 3.5 \times 10^{10}cm$, $1.99 \times 10^{33}gm$, $\lambda_{sun} = 30days$.
**Earth:** $\epsilon_{earth} = 23.5$, 50% clouds, $R_{moon} = 2160mi$, $\epsilon_{moon} = 5$, $\lambda_{sider} = 27d7h43m12s$, $\lambda_{synod} = 29d12h44m3s$, $RA_{Greenwich}(1986.0) : 6.6245$, $0\ Jan\ 1986 = 2,446,430.5JD$.

**Geological:** For seismic wave, $v_P = \sqrt{\frac{(k + \frac{4}{3}\mu)}{\rho}}$, $v_S = \sqrt{\frac{\mu}{\rho}}$.
$\mu_{granite} = 1.6 \times 10^{10}dynes/cm$, $k_{granite} = 27 \times 10^{10}dynes/cm$, $k_{water} = 2.0 \times 10^{10}dynes/cm$, $\mu_{water} = 0$.
$v_{P-granite} = 5.5km/sec$, $v_{S-granite} = 3.0km/sec$, $v_{P-water} = 1.5km/sec$, $v_{S-water} = 0$.

**Materials:** Dry (static, sliding) Friction: Steel (.78,.42), Teflon; (.04,-). Expansion: $\alpha_l = l^{-1}\frac{\partial l}{\partial t} \times 10^6$, C: (Al, 24), (Cu, 17), (Granite, 8.3), (Ice, 50), (Fe, 12), (Water, 207). Heat Capacity: $c_v = m^{-1}\frac{\partial Q}{\partial T}$ : (He, 12.5), ($O_2$, 21.1), ($N_2$, 20.6), ($C_2H_6$ , 39.3), MFP $N_2 = 10^{-5}$ cm, $C_{v,solid} = 3R$. Melting/Boiling: MP/BP (K): Au, 1336, 3081; $O_2$, 54, 90; Cu, 1356, 2839. Heat Conduction: $Q' = -\kappa A\frac{\partial T}{\partial l}W(cmK)^{-1}$ : (Cu, 4), (Fe, 0.80), (Si, 1.5), ($H_2$,.00024-.0018), (Rock, 2.8 kc/mhK). Dielectric: $\epsilon = K\epsilon_0$: (Glass, 6.7), (Water, 78), (Nylon, 3.6).
Resistivity: $R = \rho\frac{L}{A} \times 10^{-8}$: (Ag, 1.4), (Cu, 1.7), (Al, 2.8), (Fe, 9.8).
Density: $\rho/\rho_{water}$: Al, 2.7; Cu, 8; Rock, 5.5; Au, 19; Fe, 8; Gas, .68; air, .0012; wood, .75.
Moduli: $B = \frac{\Delta P}{\frac{\Delta V}{V}}$ : Al, 70; Cu, 140; Fe, 100; Water, 200.
$Y = \frac{\frac{\Delta F}{A}}{\frac{\Delta l}{l}} \times 10^{12}dy/cm^2$: Al, 70; Cu, 110; Fe, 190.
$M_s = \frac{\frac{\Delta F}{A}}{\frac{\Delta x}{l}}$ : Al, 30; Cu, 42; Fe, 100.

| **Name** | **RA** | **Dec** | **Vmag** | **Dist** | **Name** | **RA** | **Dec** | **Vmag** | **Dist** |
|---|---|---|---|---|---|---|---|---|---|
| Polaris | 01 23 | 88 46 | 2.06 | 200 | Mizar | 13 20 | 55 27 | 2.12 | 26 |
| Aldeberan | 04 30 | 16 19 | .8 | 21 | Capella | 05 09 | 45 54 | .09 | 14 |
| Rigel | 05 10 | -08 19 | .11 | 270 | Bellatrix | 05 20 | 06 16 | 1.63 | 140 |
| Betelgeuse | 05 50 | 07 23 | .4 | 180 | Sirius | 06 41 | -16 35 | -1.44 | 2.7 |
| Canopus | 06 22 | -52 38 | -.72 | ? | Castor | 07 28 | 32 06 | 1.56 | 14 |
| Procyon | 07 34 | 05 29 | .36 | 3.5 | Pollux | 07 39 | 28 16 | 1.15 | 10.7 |
| Regulus | 10 03 | 12 27 | 1.34 | 26 | Merak | 10 56 | 56 55 | 2.36 | 23 |
| Spica | 13 20 | -10 38 | .97 | 65 | Arcturus | 14 11 | 19 42 | -.05 | 11 |
| Antares | 16 23 | -26 13 | .94 | 130 | Vega | 18 34 | 38 41 | .03 | 8.1 |
| Altair | 19 46 | 08 36 | .77 | 4.9 | Deneb | 20 38 | 44 55 | 1.25 | 500 |

Figure 4.1: Stars

| **Planet** | **$D_{av}$(km × $10^6$)** | **$\lambda$(rev)** | **e** | **i** | **$L_{node}$** | **$L_{Per}$** | **$P_{epoch}$** | **M(gm)** | **R(km)** | **Rot** |
|---|---|---|---|---|---|---|---|---|---|---|
| Mercury | 57.9 | 87.97d | .2 | 7 | 47.9 | 76.8 | 222.6 | 3.3e26 | 2439 | 58.7d |
| Venus | 108.2 | 224.7d | .007 | 3.4 | 76.3 | 131.0 | 174.3 | 4.9e27 | 6050 | 243d |
| Earth | 149.6 | 365.26 | .017 | 0 | 0 | 102.3 | 100.2 | 6e27 | 6378 | 23h56m |
| Mars | 227.9 | 686.98 | .093 | 1.8 | 49.2 | 335.3 | 258.8 | 6.4e26 | 3394 | 24h37m |
| Jupiter | 778.3 | 11.8yr | .048 | 1.3 | 100.0 | 13,7 | 259.8 | 1.9e30 | 71880 | 9.8h |
| Saturn | 1427.0 | 29.46 | .056 | 2.5 | 113.3 | 92.3 | 280.7 | 5.7e29 | 60400 | 10.66h |
| Uranus | 2869 | 84 | .047 | .8 | 73.8 | 170.0 | 141.3 | 8.8e28 | 23540 | 17.24h |
| Neptune | 4496 | 164.79 | .009 | 1.8 | 131.3 | 44.3 | 216.9 | 1e29 | 24600 | 16h |
| Pluto | 5900 | 247.7 | .250 | 17.2 | 109.9 | 224.2 | 181.6 | - | - | - |

Figure 4.2: Planetary data - Epoch: 1960 Jan 1.5UT, Orbit: $a = b\sqrt{1-e^2}$.

**Air:** 28.96 m-w, $c_p$ = 1005 J/kg-K, $c_v = 718 J/kg - K$. $1\ atm = 1.013 \times 10^5 Pa$, $Pa = 10^6 dyne/cm^2 = 1N/m^2 = 760mm - Hg$. $\rho : 1.293mg/cm^3$, $\kappa : 2.4 \times 10^{-2} W/m - K$, $visc@20 : .00018g/cm - s$. Water: 273.15K, 18 m-w, 540 cal/gm (vaporization), 80 cal/gm (fusion), $\rho_{ice} = 917kg/m^3$, $\kappa : .19W/m{-}K, visc@20 : .01gm/cm - s, ST : @20 : 73d/cm$.
**Sound Strength:** $g = 10log(\frac{I}{I_0})$ in db. $I_0 = 10^{-12}W/m^2$. Normal Conversation: 60 db, Jet: 130 db.
**Speed of Sound:** $\approx 330m/s$ at normal conditions, $v_{av} = \sqrt{3kT/m}$.

Misc units: $1\ in = 2.54\ cm$, $1\ kg = 2.2046\ lbs$, $1\ fluid - oz = 0.0338\ ml$, $1\ gal = 3.3785\ liters$.

**Stellar Evolution** ($'$: means differentiate wrt $r$): $P' = -\rho\frac{GM(r)}{r^2}$, $M' = 4\pi r^2\rho$, $L' = 4\pi r^2\epsilon$, $L' = \frac{(-3\chi\rho)}{(4acT^3(4\pi r^2))}$ (rad), $L' = (1-\gamma^{-1})T\rho^{-1}P'$ (conv), $P = RT\frac{\rho}{\mu}$, $\chi = C\rho T^{-3.5}$, $\alpha = \frac{10^6}{T^{1/3}}$.

**Optics:** $n_{glass} = 1.52$, $n_{water} = 1.33$, $n_{diamond} = 2.42$. Lensmaker's law (air to glass, one surface): $\frac{1}{s} + \frac{n}{s'} = \frac{1}{f}$.
Lensmaker's law (double surface): $\frac{n_1}{s} + \frac{n_2}{s'} = \frac{1}{f}$, $\frac{1}{f} = (n_2 - n_1)\frac{1}{R_1} - \frac{1}{R_2}$. $\frac{h_i}{h_o} = \frac{d_i}{f} = \frac{f}{d_o}$. **Resolving Power:** $4.54/D_{inches}$ arc-seconds, $f_{ratio} = \frac{L_{focus}}{R_{diameter}}$, $3 \le f_{ratio} \le 6$, $Mag = \frac{L_{focus-objective}}{L_{focus-eyepiece}}$.

**Chemical bonds:** covalent: 80-200 kcal/mole (C=C is 200), ionic: 4-7 kcal/mole, hydrogen 5kcal/mole, vanderWaal $< 1kcal/mole$ (methane). Thermal: .6 kcal/mole. Acid added to $H_2O$ increases $H^+, pH = -log[H^+]$, acid $< 7$.

**Fluids:** $P + \phi + \frac{1}{2}\rho v^2 = const$, $\nabla\rho v = -\rho'$, $\nabla v = 0$, $\nabla \times v = 0$.

**Interference:** $R = A[cos(\omega t) + cos(\omega t + \phi) + \ldots + cos(\omega t + (n-1)\phi)]$. $A_R = A\frac{sin(\frac{n\phi}{2})}{sin(\frac{\phi}{2})}$. $I = I_0\frac{sin^2(\frac{n\phi}{2})}{sin^2(\frac{\phi}{2})}$.
For $f(t) = A_1e^{i\omega_1 t} + A_2e^{i\omega_2 t}$, $I = A_1^2 + A_2^2 + 2cos((\omega_1 - \omega_2)t)$. Group velocity and modulation.

| Place | Lat | Long | Place | Lat | Long | Place | Lat | Long |
|---|---|---|---|---|---|---|---|---|
| Beijing | 40.1 | 116.33 | SF | 37.45 | -122.33 | NY | 41.44 | -73.8 |
| Boston | 42.35 | -71.05 | Chicago | 41.87 | -87.63 | Dallas | 32.78 | -96.78 |
| Madison, Wi | 43.07 | -89.38 | Santa Fe | 35.68 | -105.93 | Seattle | 47.61 | -122.33 |
| Tucson | 32.22 | -110.97 | DC | 38.88 | -77.0 | Denver | 39.75 | -104.99 |
| Atlanta | 33.75 | -84.39 | London | 51.5 | 0.0 | Paris | 48.83 | 2.3 |
| Berlin | 52.5 | 13.42 | Rome | 41.88 | 12.5 | Moscow | 55.75 | 37.7 |
| Athens | 37.97 | 23.75 | Jerusalem | 31.75 | 35.22 | Tokyo | 35.75 | 139.75 |
| Sidney | -33.87 | 151.2 | MKea | 19.826 | -155.47 | CTlo | -70.82 | -30.17 |
| New Orleans | 29.93 | -90.07 | Redmond,OR | 44.27 | -121.15 | Portland | 45.52 | -122.68 |
| LA, CA | 34.05 | -118.24 | San Diego | 32.7 | -117.15 | Orlando | 28.52 | -81.38 |
| Milan | 45.45 | 9.28 | Amsterdam | 52.3 | 4.77 | Auckland | -36.92 | 138.58 |
| Bombay | 18.93 | 74.58 | Delhi | 28.67 | 77.23 | Perth | -31.93 | -115.83 |
| Toronto | 43.65 | -79.38 | Bagdad | 33.3 | 44.43 | Cairo | 30.03 | 31.35 |

Figure 4.3: Places on Earth

**Spectrum:** 30cps audio 30K 500K AM 1500K 3M HF 30M 88M FM VHF 210M 400M UHF 800M 1.5G $H_2$ S-band 3G 7600A IR 6300 Visible 3900A UV 100A X-ray .1A gamma 67 Mev.
Red: 650nm, Yellow: 580 nm, Green: 500nm, Blue: 475nm, Violet: 400Nm.

**Middle C:** 256Hz. Octave has 12 notes in uniformly divided log scale. Octave is factor of 2.

**Central Forces:** $\vec{F}(r) = f(r)\hat{r}$. $m(\ddot{r} - r\dot{\theta}^2) = f(r)$ and (conservation of angular momentum) $m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = 0$. $r^2\dot{\theta} = h$, $\ddot{r} - \frac{h^2}{r^3} = \frac{f(r)}{m}$. If $V(r) = -\int f(r)dr$, $\frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + V(r) = E$; ellipse if $E < 0$, parabola if $E = 0$, hyperbola if $E > 0$. Force from path: $f(r) = \frac{mh^2}{r^4}[\frac{d^2r}{d\theta^2} - \frac{2}{r}(\frac{dr}{d\theta})^2 - r]$.

**Rotating frames:** Suppose $XYZ(F)$ is inertial system and $xyz(M)$ is rotating frame with a common origin $O$. $(\frac{d\vec{A}}{dt})_{|F} = (\frac{dA}{dt})_{|M} + \omega \times \vec{A}$. ${D_F}^2\vec{r} = {D_F}^2\vec{r} + D_M(\vec{\omega}) \times \vec{r} + 2\vec{\omega} \times D_M\vec{r} + \vec{\omega} \times (\vec{\omega} \times \vec{r})$. Last two terms are Coriolis and Centripetal. If $O$ is moving too, $D_F(\vec{r}) = \dot{R} + D_M\vec{r} + \vec{\omega} \times \vec{r}$ and ${D_F}^2\vec{r} = \ddot{R} + {D_M}^2\vec{r} + D_M(\vec{\omega}) \times \vec{r} + 2\vec{\omega} \times D_M\vec{r} + \vec{\omega} \times (\vec{\omega} \times \vec{r})$. Object dropped from rotating sphere from a height $h$ is deflected by $\frac{1}{3}\omega g t^3 sin(\lambda)$, where $\lambda$ is the colatitude.

**Foucault** (constrained to horizontal plane): $m\ddot{x} = -T(\frac{x}{l}) + 2m\omega\dot{y}cos(\lambda)$, $m\ddot{y} = -T(\frac{y}{l}) - 2m\omega(\dot{x}cos(\lambda) - \dot{z}sin(\lambda))$, $m\ddot{z} = -T(\frac{l-z}{l}) - mg + 2m\omega\dot{y}sin(\lambda)$, $\hat{n} = isin(\omega cos(\lambda)t) + jcos(\omega cos(\lambda)t)$.

**Rotation in plane:** $I = \int r^2 dm$. $\vec{\Omega} = I\vec{\omega}$, $T = \frac{1}{2}I\omega^2$ $\vec{\Lambda} = I\dot{\vec{\omega}}$. **Parallel axis theorem:** $I_A = I_{CM} + mb^2$. **Perpendicular axis theorem:** $I_x = I_y + I_z$. $I_{sphere} = \frac{2}{5}ma^2$. $I_{cylinder} = \frac{1}{2}ma^2$. $I_{plate} = \frac{1}{12}m(a^2 + b^2)$.

**Rotation in space:** $\Omega = \sum m_\mu(r_\mu \times (\omega \times r_\mu)$, $[(r_\mu \times (\omega \times r_\mu)]_x = {\omega_x}^2({y_\mu}^2 + {z_\mu}^2) - \omega_y x_\mu y_\mu - \omega_z x_\mu z_\mu$, $I_{xx} = \int(y^2 + z^2)dm$, $I_{xy} = -\int(xy)dm$, $\mathcal{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix}$ is the inertia tensor. $T = \frac{1}{2}\omega \cdot \Omega$ is kinetic energy. **Principal Axis Theorem:** If $\omega_1, \omega_2, \omega_3$ and $\Omega_1, \Omega_2, \Omega_3$ are the angular velocities and momenta about the principal axis, $\Omega_i = I_i\omega_i$ and $T = \frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_1^3)$. **Ellipsoid of rotation:** Let $\hat{n}$ be a unit vector in the direction of $\hat{\omega}$, $\vec{\omega} = \hat{n}\omega = \omega(icos(\alpha) + jcos(\beta) + kcos(\gamma))$. $T = \frac{1}{2}I\omega^2$ where $I = I_{xx}cos^2(\alpha) + I_{yy}cos^2(\beta) + I_{zz}cos^2(\gamma) + 2I_{xy}cos(\alpha)cos(\beta) + 2I_{xz}cos(\alpha)cos(\gamma) + 2I_{yz}cos(\beta)cos(\gamma)$. $\rho = \frac{\hat{n}}{\sqrt{I}}$ is ellipsoid of revolution. **Rotational symmetry about $s = z$ axis:** $I_s = I_z$, $I = I_x = I_y$. $I\dot{\omega}_x + \omega_y\omega_z(I_s - I) = 0$, $I\dot{\omega}_y + \omega_x\omega_z(I - I_s) = 0$, $I_s\dot{\omega}_z = 0$. $\vec{J}_s = const$, put $\gamma = \frac{I_s - I}{I}\omega_s$; then $\dot{\omega}_x + \gamma\omega_y =$, $\dot{\omega}_y - \gamma\omega_x =$, so $\ddot{\omega}_x + \gamma^2\omega_x = 0$ and $T_p = \frac{2\pi}{\gamma}$. **Precession of Earth:** $T_p = \frac{2\pi I}{\omega_z(I_s - I)} \approx 305 days$. Precession of Disc: $T_p = \frac{2\pi}{\omega_z}$. **Gyroscope:** $J_{x'} = I_{x'}\omega_{x'} = I\dot{\theta}$, $J_{y'} = I\varphi sin(\theta)$, $J_{z'} = I_s S$. $S = \dot{\varphi}cos(\theta) + \dot{\phi}$, $I_s\dot{S} = 0$.

**Euler's Equations:** Let $O$ be a principal axis coordinate system fixed in a body, the external torque is $\vec{\Lambda}$. $I_1\dot{\omega}_1 + (I_3 - I_2)\omega_2\omega_3 = \Lambda_1$, $I_2\dot{\omega}_2 + (I_1 - I_3)\omega_1\omega_3 = \Lambda_2$, $I_3\dot{\omega}_3 + (I_2 - I_1)\omega_1\omega_2 = \Lambda_3$ along the principal axes. $\omega \cdot \Omega = c$ is invariant plane. The angular velocity and momentum in terms of the Euler angles $\phi, \theta, \psi$, from $O_{xyz}$ fixed in space to $O_{x'y'z'}$, principal axis, is: $\omega_{x'} = \dot{\phi} sin(\theta) sin(\psi) + \dot{\theta} sin(\psi)$, $\omega_{y'} = \dot{\phi} sin(\theta) cos(\psi) - \dot{\theta} sin(\psi)$, $\omega_{z'} = \dot{\phi} cos(\theta) + \dot{\psi}$, $\phi$ is from $x$ to line of nodes, $\theta$ is from $z$ to $z'$ axis, and, $\psi$ is from line of nodes to $x'$;. $T = \frac{1}{2}(I_1 {\omega_1}^2 + I_2 {\omega_2}^2 + I_3 {\omega_3}^2)$. **Top:** Suppose $\vec{e_3}$ is the axis of top's line of symmetry. $\vec{s} = s\vec{e_3} = \dot{\psi}\vec{e_3}$ . $\Omega = I_1 \omega_1 e_1 + I_2 \omega_2 e_2 + I_3(\omega_3 + s) e_3$, $\Lambda = le_3 \times mg = (\frac{d\Omega}{dt})_F, I_1 = I_2$. $(\frac{d\Omega}{dt})_F = (\frac{d\Omega}{dt})_B + \omega \times \Omega$. $I_1 \dot{\omega_1} + (I_3 - I_2)\omega_2 \omega_3 = mglsin(\theta)$, $I_2 \dot{\omega_2} + (I_1 - I_3)\omega_1 \omega_3 - I_3 \omega_1 s = 0$, $I_3(\dot{\omega_3} + \dot{s}) = 0$. In Euler angles, with $\psi = 0$, this is $\omega_1 = \dot{\theta}$, $\omega_2 = \dot{\psi} sin(\theta)$, $\omega_3 = \dot{\psi} cos(\theta)$. $\dot{\theta}, \dot{\psi}, s$ are angular velocity of precession, nutation and spin.

**Holonomic constraint:** $\phi(q_1, q_2, ..., q_n, t) = 0$. **Generalized coordinates:** $\delta W = \sum_\alpha \Phi_\alpha \delta q_\alpha$, $\Phi_\alpha = \sum \vec{f} \cdot \frac{\partial r}{\partial q_\alpha}$. **Lagrange equations:** $(\frac{d}{dt}) \frac{\partial T}{\partial \dot{q_\alpha}} - \frac{\partial T}{\partial q_\alpha} = \Phi_\alpha$. If the forces are all conservative and $L = T - V$ then $(\frac{d}{dt}) \frac{\partial L}{\partial \dot{q_\alpha}} - \frac{\partial L}{\partial q_\alpha} = 0$. **Generalized momentum:** $p_\alpha = \frac{\partial T}{\partial \dot{q_\alpha}}$. **Hamilton:** $H(p_1, ..., p, q_1, ..., q_n, t) = \sum p_\alpha \dot{q_\alpha} - L$. $\dot{p_\alpha} = -\frac{\partial H}{\partial q_\alpha}$, $\dot{q_\alpha} = \frac{\partial H}{\partial p_\alpha}$. Hamilton Principal: For conservative forces ($H = T + V$), $L = T - V$, $\delta \int_{t_1}^{t_2} L dt = 0$. Note: $H = \sum p_\alpha \dot{q_\alpha} - L$.

## 4.3 Quantum Mechanics

**Formalism:** Let $|i>$ denote base states $< i|j >= \delta_{ij}$. $|\psi >= \sum_i |i >< i|\psi >$, $< \psi|\phi >= \sum_i < \phi|i >< i|\psi >$. $|\psi >= \sum_i |i >< i|\psi >$ evolves under $\hat{A}$ so $|\phi >= \hat{A}|\psi >$ and $< i|\phi >= \sum_j < i|\hat{A}|j >< j|\psi >$, $A_{ij} =< i|\hat{A}|j >$.

**Free particle:** $\Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi(k) e^{i(kx - \frac{\hbar k^2}{2m} t)} dk$ and $\phi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x, 0) e^{-ikx} dx$. For free particle, $\hbar\omega = \frac{\hbar^2 k^2}{2m}$.

$\int_{\mathbb{R}^3} |\psi(\vec{x})| d\vec{x} = 1$. **Spatial operators:** $X\psi = x\psi$, $Y\psi = y\psi$, $Z\psi = z\psi$, $\vec{R} = (X, Y, Z)$. **Momentum operators:** $p_x \psi = \frac{\hbar}{i} \frac{\partial}{\partial x} \psi$, $p_y \psi = \frac{\hbar}{i} \frac{\partial}{\partial y} \psi$, $p_y \psi = \frac{\hbar}{i} \frac{\partial}{\partial z} \psi$, $\vec{P} = (p_x, p_y, p_z)$. **Angular Momentum:** $L_x = yp_z - zp_y$, etc. $< A >= \int \psi^*(\vec{r}) A \psi(\vec{r}) d\vec{r}$, $\Delta A = \sqrt{< A^2 > - < A >^2}$.

Elements of state space are denoted: $| >$ and $(\phi, \psi) =< \phi|\psi >$, physically observable quantities are described by hermitian operators acting on state space: $A|\psi >$, each observable quantity is an eigenvalue of the hermitian operator.

**Postulate 1:** Associated with any *isolated* physical system is a complex vector space, $V$ with an inner product called a state space. The system is completely described by $v \in V$.

**Postulate 2:** The evolution of a closed quantum system is described by a unitary transformation on the state: $|\psi(t_2) >= U|\psi(t_1) >$. **Postulate 2':** The evolution of a closed quantum system is described by Schroedinger's equation $i\hbar \frac{\partial|\psi>}{\partial t} = H|\psi >$.

**Postulate 3:** Quantum measurements are described by a collection of measurement operators, $\{M_m\}$ that act on the state space. If $|\psi >$ is the state immediately before the measurement, the probability that the event $m$ occurs is $< \psi|M_m^\dagger M_m|\psi >$ and the state after the measurement is given by $\frac{M_m|\psi>}{\sqrt{<\psi|M_m^\dagger M_m|\psi>}}$ and $M_m$ satisfies $\sum_m M_m^\dagger M_m = I$. A projective measurement on an observable with spectral decomposition, $M = \sum_m m P_m$, results in one of the $m$ values as possible outcomes. $\Delta(C)\Delta(D) \geq \frac{<\psi|[C,D]|\psi>}{2}$.

**Postulate 4:** The state space of a composite system is the tensor product of the state spaces of the component systems. If we number the systems $1, 2, \ldots, n$, and system $i$ is in the prepared state $|\psi_i >$ then the joint state is $|\psi_1 > \otimes |\psi_2 > \otimes \ldots \otimes |\psi_n >$.

A set of gates is said to be a set of **universal quantum gates** if any unitary operator can be approximated to arbitrary accuracy by a quantum circuit involving only those gates. The Hadamard, CNOT, phase and $\frac{\pi}{8}$ gates form a universal set.

For harmonic oscillator: $H = \frac{p^2}{2m} + k\frac{x^2}{2}$. Another hamiltonian: $H = \frac{1}{2m}(p - \frac{q}{c}A)^2 + V(R) + q\phi - \frac{q}{mc}S \cdot B, B = \nabla \times A$. Simultaneously observable quantities commute. Independence and uncertainty: $[R_j, P_k] = i\hbar\delta_{jk}, [R_j, R_k] = 0$.

**Feynman Postulates:** If there is no spin or polarization: (1) $< x|s >= a + bi$. $Pr($ particle arrives at $x|$ particle leaves $s) = | < x|s > |^2$. (2) $< x|s >_{both} =< x|s >_1 + < x|s >_2$. (3) $< x|s >_{via\ 1} =< x|1 >< 1|s >$. $a$ is probability light scattered at 1 arrives at $D_1$ and $b$ that it arrives at $D_2$ ($a >> b$). $< \vec{r_2}|\vec{r_1} >= \frac{A}{r_{12}}e^{i\frac{\vec{p}\cdot\vec{r_{12}}}{\hbar}}$; get $p$ relativistically by $(pc)^2 = E^2 - (m_0c^2)^2$ or non-relativistically as $E = \frac{p^2}{2m}$. Rules for outcomes: (1)If final states are distinguishable, add probabilities *not amplitudes* for indistiguisible processes leading to the same final state add *amplitudes*; (2) use complete description of isolates system. Outcomes of scattering with indistinuishable particles always interfere: add amplitudes for Bosons, subtract for Fermions. $P_n(Bose) = n!P_n(different)$. Treat metal conduction as noninteracting Fermion gas.

For H, $V(r) = -\frac{q^2}{r}$. $i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi - V(r)$. $\psi(r,t) = e^{\frac{i}{\hbar}Et}$. If $f(x,u,z) = g(r(x,y,z), \theta(x,y,z), \psi(x,y,z))$, $f_{xx} = g_{rr}(r_x)^2 + g_{\theta\theta}(\theta_x)^2 + g_{\psi\psi}(\psi_x)^2 + 2(g_{r\theta}r_x\theta_x + g_{r\psi}r_x\psi_x + g_{\psi\theta}\psi_x\theta_x) + g_r r_{xx} + g_\theta\theta_{xx} + g_\psi\psi_{xx}$. $\nabla^2 f = \frac{1}{r^2}(r^2 g_r)_r + \frac{1}{rsin(\theta)}(sin(\theta)f_\theta)_\theta + \frac{1}{r^2 sin^2(\theta)}f_{\psi\psi}$. Use this to solve Schrodinger Equation for Hydrogen.

## 4.4 Thermodynamics and Statistical Mechanics

**First law:** $\Delta Q$: heat into system, If $\Delta W$: work on system, $\Delta E$: increase in energy of system then $\Delta Q + \Delta W = \Delta E$. For ideal gas, $PV = E = \frac{2}{3}N < \frac{mv^2}{2} >= nRT = NkT$. In general, $PV = (\gamma - 1)U$ ($\gamma = \frac{5}{3}$ for ideal gas). $(\frac{\partial U}{\partial T})_V = C_v = \frac{3}{2}R$, $(\frac{\partial U}{\partial T})_p = C_p = C_v + R$, for adiabatic process: $pV^\gamma = c$, $\gamma = \frac{C_P}{C_V}$. **Second Law:** It is impossible to build a cyclic engine that converts thermal energy completely into mechanical work. **Carnot Process:** $1 \rightarrow 2$: isothermal at $T_H$, $2 \rightarrow 3$: adiabatic add $Q_H$, $3 \rightarrow 4$: isothermal at $T_C$, $4 \rightarrow 1$: adiabatic add $Q_C$. $e = 1 - \frac{T_C}{T_H}$. $S = \int \frac{dQ}{T} \geq 0$. $S = Nk_B ln(\Omega)$. In irreversible process, entropy increases, at $T = 0$, $S = 0$. For reversible process, $S = \frac{Q_1}{T_1} = \frac{Q_2}{T_2}$, $W = Q_1 - Q_2 = Q_1(1 - \frac{T_2}{T_1})$; $eff = \frac{W}{Q_1} = \frac{T_2 - T_1}{T_1}$. $S = kln(W)$.

For **monatomic gas**, $P = \frac{2}{3}U = \gamma - 1 = \frac{2}{3} < mv^2 >= \frac{3}{2}kT$. In a mixture at constant temperature with two species 1 and 2, $n_1 < m_1{v_1}^2 >= n_2 < m_2{v_2}^2 >$ but considering two particles with relative velocity $w$ with velocity of enter of mass $v_{CM}$ we can argue at equilibrium that $< w \cdot V_{CM} >= 0$ so $n_1 = n_2$ (Avogadro's hypothesis. For **photon gas**, $PV = N < p \cdot v > /3$ so $\gamma = \frac{4}{3}$. For **diatomic gas**: $\gamma = \frac{9}{7}$.

**Atmosphere:** $\frac{dn}{dh} = -\frac{mg}{kT}$, $n = n_0 e^{-PE/kT}$ and $\frac{n_{>u}}{n{<u}} = e^{-KE/kt}$. **Evaporation model:** $W$ is binding energy of liquid, $n$ is density of vapor, $1/V_a$ is density of liquid then $nV_a = e^{-W/kT}$. **Chemical kinetics:** $\frac{n_A n_B}{n_{AB}} = ce^{W/kt}$. **Diffusion:** Average time to collision is $\frac{1}{n_0}\int_0^\infty t\frac{N(t)dt}{\tau}$, $N = N_0 e^{t/\tau}$. **Mean Free Path:** $l = \tau v$. **Thermal conductivity:** $\frac{1}{A}\frac{dQ}{dt} = -\kappa\frac{dT}{dz}$, $\kappa = \frac{knlv}{\gamma - 1}$ if $MFP <<$ container.

**Maxwell Distribution:** $F_{MB} = N(\frac{m}{2\pi kT})^{\frac{3}{2}}e^{-m(v_x^2+v_y^2+v_z^2)/(2kT)}$, the frequency of a particle around $v$; $dn_\nu = F_{MB}g(q)dq$. $v_{rms} = \sqrt{\frac{3kT}{m}}$.
**Bose-Einstein Distribution (Bosons):** $F_{BE} = (e^\alpha e^{E_i/kT} - 1)^{-1}$, $\alpha$ is type specific 0 for photon.
**Fermi-Dirac Distributions (Fermions):** $F_{FD} = (e^{(E_i - E_f)/kT} + 1)^{-1}$, $E_f$ is the Fermi energy. $C_V = \frac{1}{N}(\frac{\partial E}{\partial T})_V$ approximately $3R$ for many solids.

**Conductor:** half filled conduction band. **Insulator:** filled conduction band large gap $\approx 5ev$. **Semiconductor:** filled conduction band small gap $\approx 1ev$ which can be overcome by thermal excitation. **Electron mobility:** $\mu = \frac{v_d}{E}$, $v_d$ is drift velocity. **Fine constant:** $\frac{ke^2}{\hbar c} \approx \frac{1}{137}$. **Josephson junction** is two superconductors separated by thin $\approx 1nm$ insulator; if there is no potential difference, electrons tunnel and we get dc, if dc potential is applied, we get ac with $f \approx \frac{2eV}{h}$.

| Type | Family 1 | Family 2 | Family 3 |
|---|---|---|---|
| Quark | Up ($u$) | Charm ($c$) | Top ($t$) |
| Quark | Down ($d$) | Strange ($s$) | Bottom ($b$) |
| Lepton | electron neutrino ($\nu_e$) | muon neutrino ($\nu_\mu$) | Tau neutrino ($\nu_\tau$) |
| Lepton | electron ($e$) | muon ($\mu$) | Tau ($\tau$) |

Figure 4.4: Matter Particles - Fermions - not including antiparticles

## 4.5 More Quantum

**Polar Decomposition:** Let $A$ be a linear operator on $V$. Then there is a unitary operator $U$ and positive operators $J, K$: $A = UJ = KU$. $J = \sqrt{A^T A}$, $K = \sqrt{AA^T}$.

**Singular Value Decomposition:** Let $A$ be a square matrix the $\exists U, V$ and a diagonal matrix $D$ with non-negative entries such that $A = UDV$. Entries of $D$ are called singular values.

**Standard Model:** Quantized force fields materialize as particles. Matter particles: **Fermions** (half-integral spins). Force particles: **Bosons** (integral spins). $u$: $q = +\frac{2}{3}$, $m = 2Mev$; $d$: $q = -\frac{1}{3}$, $m = 5Mev$; $c$: $q = +\frac{2}{3}$, $m = 1.25Gev$; $s$: $q = -\frac{1}{3}$, $m = 95Mev$; $t$: $q = +\frac{2}{3}$, $m = 171Mev$; $b$: $q = -\frac{1}{3}$, $m = 4.2Gev$.
$\nu_e$: $q = 0$; $\nu_\mu$: $q = 0$; $\nu_\tau$: $q = 0$. $e$: $-1$, $m = .511Mev$; $\mu$: $-1$, $m = 106Mev$; $\tau$: $0$, $m = 1.78Gev$. Bosons: Photon $\gamma$ - EM Force: $q = 0$, $m = 0$; Gluons - Strong Force: $q = 0$, $m = 0$; Z - weak force: $q = 0$, $m = 91Gev$; $W^+, W^-$ - weak force: $q = 0$, $m = 80.4Gev$; Higgs ($H$): $q = 0$, $114Gev < m < 192Gev$; Graviton - gravity: $q = 0$, $m = 0$.

**Hall Effect:** In metal or semiconductor, imagine a thin ($2D$) slab, $z$-up, $x$-across, $y$-back, with an electric field, $\vec{E}_y$, back, current $\vec{j}_x$ across. Turn on a magnetic field $\vec{B}_z$, and the charges move to the back until equilibrium caused by electrostatic build-up when $B_z v_x = E_y$, then Hall resistance is $R_H = \frac{E_y}{B_z} j_x$, $j_x = v_x N_q$. At low temperature ($< 30mK$), a quantum effect appears: $R_H$ grows monotonically with $\vec{B}_z$ and is quantized by $\frac{1}{n}\frac{h}{e^2}$; this IQHE is evident in a GaAs-GaAlAs hetero-juncture. The magnetic field shifts the Landau Levels. The diagonal resistance $R_{xx}$ is at times 0 when the **Fermi energy** of the electrons lies between the **Landau Levels** freezing out scattering. (The Fermi energy, $E_F$, is the energy of the fermion composite at $0K$.) When the mobility of the electrons is high, additional plateaus (corresponding to $R_{xx} = 0$) appear; this is due to electron interaction giving rise to fractional charge like quasi-particles; this is the FQHE. Unlike IQHE, the FQHE gives rise to non-Abelian statistics in the gapped degenerate states.

**Laughlin wave function:** $\Phi^m(z_1, \ldots, z_n) = \prod_{i<j}(z_i - z_j)^m e^{-\frac{1}{4l^2}\sum |z_i|^2}$. **Moore-Reed:** $\Phi^m(z_1, \ldots, z_n) = \prod_{i<j}(z_i - z_j)^m e^{-\frac{1}{4l^2}\sum |z_i|^2} Pf(\frac{1}{z_i - z_j})$. Energy spectrum of 2DEG breaks into allowed states $E_n = (n + \frac{1}{2})\hbar\omega_c$ in $B$ field (Landau levels). When chemical potential lies in Landau bands, material is metallic. Otherwise localized states materialize adding electrons only add and subtract localized states, no currents flow and system is **incompressible**. **Magnetic Length:** $l_B = \sqrt{\frac{\hbar}{eB}}$; within $l_B$ of the edge, they form quasi-1D channels. Because there is no back-scattering, $R_{xx} = 0$. **Hidden subgroup:** $G^{abelian} \geq H$, $f : G \rightarrow X$ hides $H$ if $f : G/H \leftrightarrow X$.

A set of gates is said to be a set of **Universal Quantum gates** if any unitary operator can be approximated to arbitrary accuracy by a quantum circuit involving only those gates. The Hadamard, CNOT, phase and $\frac{\pi}{8}$ gates form a universal set.

**Quantum Ion Trap System.** The qubits are representations of the hyperfine nuclear spin states at the lowest vibrational modes (phonons) of trapped atoms. Arbitrary transforms are constructed with laser pulses using Jaynes Cummings. Qubits interact via shared phonon state. Initial state preparation involves cooling atoms by trapping and optical pumping to their lowest motional ground and hyperfine state. The measurement is the measurement of the population of hyperfine states. The phonon lifetimes are short and the atoms are difficult to prepare. For NMR, the coupling is weak and difficult to control.

**Spintronics:** Spintronics exploits the intrinsic spin of electrons and its associated magnetic moment, in addition to its fundamental electronic charge, in solid-state devices. Electrons are spin-1/2 fermions and constitute a two-state system with spin "up" and spin "down". To make a spintronic device, the primary requirements are to have a system that can generate a current of spin polarized electrons comprising more of one spin species – up or down – than the other (called a spin injector), and a separate system that is sensitive to the spin polarization of the electrons (spin detector). Manipulation of the electron spin during transport between injector and detector (especially in semiconductors) via spin precession can be accomplished using real external magnetic fields or effective fields caused by spin-orbit interaction. Spin polarization in non-magnetic materials can be achieved either through the Zeeman effect in large magnetic fields and low temperatures, or by non-equilibrium methods. In the latter case, the non-equilibrium polarization will decay over a timescale called the "spin lifetime". Spin lifetimes of conduction electrons in metals are relatively short (typically less than 1 nanosecond) but in semiconductors the lifetimes can be very long (microseconds at low temperatures), especially when the electrons are isolated in local trapping potentials (for instance, at impurities, where lifetimes can be milliseconds).

**Quantum error correcting conditions:** Suppose $C$ is a quantum code and $P$ is a projection operator onto $C$. Suppose $\mathcal{E}$ is a quantum operator with measurements $E_i$. A necessary and sufficient condition for the existence of an error correction operator $\mathcal{R}$ is $PE_i^\dagger + E_j P = \alpha_{ij}$.

A **qubit** is a two dimensional space $|\psi >= a|0> +b|1>$ such that $|a|^2 + |b|^2 = 1$.

Let $R_1, R_2, \ldots, R_N$ be trajectories in $3+1$ dimensional space from $t_i$ to $t_f$. $\psi(\vec{r_1}, \vec{r_2}) \rightarrow e^{i\theta}$. Normally, $\theta$ can either be 0 or $\pi$, if $\theta$ is arbitrary, this describes an *anyon.* Non-abelian anyons are associated with higher dimensional representations of the braid group. This can occur when there is a set of $g$ degenerate states with particles are fixed $R_1, \ldots, R_N$. If $\{\psi_\alpha\}$ is an orthonormal basis and $\psi_\alpha \rightarrow [\rho(\sigma_1)]_{\alpha\beta}\psi_\beta$. It is non-abelian if $[\rho(\sigma_1)]_{\alpha\beta}[\rho(\sigma_2)]_{\beta\gamma} \neq [\rho(\sigma_2)]_{\alpha\beta}[\rho(\sigma_1)]_{\beta\gamma}$.

**Filling factor:** Ratio of electrons to magnetic flux quanta. For FQHE: $\nu = \frac{1}{R_H}\frac{h}{e^2}$ or $\sigma_H = \nu\frac{e^2}{h}$. For composite fermions with $p$-filled Landau levels, $\nu = \frac{p}{2p+1}$. $\frac{1}{3}$ state is fully spin polarized. **Luttinger Liquid:** Interacting electrons in a one dimensional conductor. **Fermi Energy:** The Fermi energy is the energy of the highest occupied quantum state in a system of fermions at absolute zero temperature. A **quantum dot** is a semiconductor whose excitons are confined in all three spatial dimensions. A **quantum well** is a semiconductor whose excitons are confined in two spatial dimensions. A **quantum wire** is a semiconductor whose excitons are confined in one spatial dimension. **Spin Polarization** is the degree to which the intrinsic angular momentum of elementary particles, is aligned with a given direction. This property is related to the magnetic moment, of conduction electrons in ferromagnetic metals giving rise to spin polarized currents. It may also apply to beams of particles, produced for particular aims, such as polarized neutron scattering or muon spin spectroscopy. Spin polarization of electrons or of nuclei, often called simply magnetization, is also produced by the application of a magnetic field is used to produce an induction signal in electron spin resonance (ESR or EPR) and in nuclear magnetic resonance (NMR).

**Aharonov-Bohm:** The AharonovBohm effect is a quantum mechanical phenomenon by which a charged particle is affected by electromagnetic fields in regions from which the particle is excluded. Such effects are predicted to arise from both magnetic fields and electric fields, but the magnetic version has been easier to observe. According to AharonovBohm, the knowledge of the classical electromagnetic field acting locally on a particle is not sufficient to predict its quantum-mechanical behavior. In the case of the AharonovBohm solenoid effect, the wave function of a charged particle passing around a long solenoid experiences a phase shift as a result of the enclosed magnetic field, despite the magnetic field being zero in the region through which the particle passes. This phase shift has been observed experimentally by its effect on interference fringes. There are also magnetic AharonovBohm effects on bound energies and scattering cross sections, but these cases have not been experimentally tested. If $\vec{B} = \nabla \times \vec{A} = 0$, $\varphi = \frac{q}{\hbar}\int_P \vec{A} \cdot dx$. $\Delta\varphi$ is determined by $\Phi$ through the area between two paths; $\Delta\varphi = 2\pi k$ for superconductor through closed loop. Existance of monopole forces $E, B$ to be quantized. The **Coulomb Blockade** is the increased resistance at small bias voltages of an electronic device comprising at least one low-capacitance tunnel junction. **Magnetic quantization:** $\Phi_0 = \frac{h}{2e} \approx 2 \times 10^{-15} Wb$; measured by Josephson effect. **Berry Phase:** Phase acquired in cyclic adiabatic process; measured through interference experiment.

**Ising Model:** Spin coupling: $E = -\sum_{i,j} J_{ij} S_i S_j$. One dimensional: $E = \sum_i S_i S_{i+1}$. Two dimensional: $E = -\sum_{i,j} S_{i,j} S_{i,j+1} + S_{i,j} S_{i+1,j}$. Magnetic field breaks the symmetry. Computational model: (1) Pick random site, (2) flip spin and calculate $\Delta E$, (3) if $\Delta E < 0$, accept, (4) if $\Delta E > 0$ accept with probability $e^{-\beta \Delta E}$.

**Cauchy-Schwartz:** $< \phi|\psi > \leq < \phi|\phi >< \psi|\psi >$. $T_{ij} =< u_i|T|u_j >$ then $T = \sum_{ij} T_{ij} |u_i >< u_j|$. **Continuous version of inner product:** $< \phi|\psi >= \int \phi^* \psi dx$. If $|\phi >= \sum_i c_i |u_i >$ then $\psi >\rightarrow \begin{pmatrix} c_1 \\ c_2 \\ \ldots \\ c_n \end{pmatrix} = \begin{pmatrix} < u_1|\psi > \\ < u_2|\psi > \\ \ldots \\ < u_n|\psi > \end{pmatrix}$. If $|\psi >= \alpha|0 > + \beta|1 >$ then $| < 0|\psi > |^2 = |\alpha|^2$. **Projection operator:** $P_m = \sum_{i=1}^m |u_i >< u_i|$. **Observables:** Hermitian operators on state vectors.

$|\psi(t) >= (\alpha(t), \beta(t))^T$. $H|\psi >= \begin{pmatrix} \omega_1 & \omega_2 \\ \omega_2 & \omega_1 \end{pmatrix} \begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = i\hbar \frac{\partial |\psi>}{\partial t} = \begin{pmatrix} \frac{d\alpha(t)}{dt} \\ \frac{d\beta(t)}{dt} \end{pmatrix}$, $\alpha(t) = e^{i\frac{\omega_1 t}{\hbar}} cos(i\frac{\omega_2 t}{\hbar}))$.

$< w|T^\dagger|v >=< v|T|w >^*$, $[X, P] = i\hbar$. Finding similarity: (1) find eigenvalues/eigenvectors, (2) normalize eigenvectors, $v_i$, (3) $S^{-1} = (v_1, \ldots, v_n)$.

**Hadamard Gate:** $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. **Pauli matricies:** $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. $\vec{L} = \vec{r} \times \vec{p}$, $L_x = yp_z - zp_y$, $[L_x, L_y] = i\hbar L_z$. $(\Delta A)^2 (\Delta B)^2 \geq (\frac{<A,B>}{2i})^2$.

**Degeneracy (duplicate eigenvalues):** Suppose $A$ has $g_m$ degenerate states then $Prob(\lambda_m) = \sum_{i=1}^{g_m} | < a_m^i|\psi > |^2$.

**Tensor:** $\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix}$.

**Density Operator:** $\rho = \sum_i p_i |\psi_i >< \psi_i|$. If system is in a pure stare $Tr(\rho) = 1$; if system is in a mixed state $Tr(\rho) < 1$. **Klein Gordon and Dirac:** $\frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} - \frac{\partial^2 \varphi}{\partial x^2} = \frac{(mc)^2}{\hbar^2} \varphi$, $i\hbar \frac{\partial \varphi}{\partial t} = -i\hbar c \alpha \cdot \nabla \psi + \beta mc^2 \psi$. **Canonical momentum density:** $\pi(x) = \frac{\partial L}{\partial \dot{\varphi}}$.

$L = T - V$, $S = \int L dt$, $H(p, q) = \sum_i p_i \dot{q}_i - L$. $p = \frac{\partial L}{\partial \dot{q}}$, $F = \frac{\partial L}{\partial q}$ and $\dot{p} = F$.

**Symmetries:** EM ($U(2)$), Weak ($SU(2)$), Strong ($SU(3)$). A **Lie group** (1) depends on parameters $\theta_1, \ldots, \theta_n$ and (2) derivatives with respect to group parameters exist. The diffeomorphism group of a Lie group acts transitively on the Lie group. $g(\theta)_{\theta=0} = e$, $\frac{\partial g(\theta_1, \ldots, \theta_n)}{\partial \theta_i}_{\theta_j = 0} = iX_j$ are the generators. $[X_i, X_j] = if_{ijk} X_k$ is group algebra. Consider $R_x(\zeta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\zeta) & sin(\zeta) \\ 0 & -sin(\zeta) & cos(\zeta) \end{pmatrix}$, $R_y(\phi) = \begin{pmatrix} cos(\phi) & 0 & sin(\phi) \\ 0 & 1 & 0 \\ -sin(\phi) & 0 & cos(\phi) \end{pmatrix}$, and $R_z(\theta) = \begin{pmatrix} cos(\theta) & sin(\theta) & 0 \\ -sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Unitary: $[U, H] = 0$. $SU(2)$ has 3 generators and $SU(3)$ has 8.
**Noether:** If $T(s)$ is a transformation $T(s) : q \mapsto q(s)$ and $\frac{\partial L(q(s)}{\dot{q}(s)} = 0$, then $C = p \frac{\partial q(s)}{\partial s}$ is a conserved quantity.

**Model for electron flow in crystal:** Let $C_n$ be the wave function at site $n$ in a linear array of molecules in a lattice each separated by a distance $b$. $i\hbar \frac{\partial C_{n-1}}{\partial t} = E_0 C_{n-1} - AC_n - AC_{n-2}$. $C_n = a_n(x) e^{-i(e/\hbar)t}$ and $a_n(x) = e^{ikx}$. Substituting,$E = E_0 + A(e^{-ikx} + e^{ikx})$. Using $E_0 = 2A$ and $cos(t) \approx 1 - t^2/2$ for small $t$, we get $E = \hbar\omega = \frac{Ab^2k^2}{2}$ so $\frac{d\omega}{dk} = \frac{2Ab^2}{\hbar} k$. If $E$ is different, say $E_0 + F$ at site 0, we get backscattering or trapping depending on the sign of $F$.

Energy in conduction band $\approx E_0 + \alpha k^2$. $N_n N_p = ce^{E_{gap}/(kT)}$. $E_{gap,Ger} \approx .72ev$, $E_{gap,Si} \approx 1.1ev$; at room temperature, $kT \approx \frac{1}{40} ev$.

**Semiconductor junctions:** $\frac{N_p(p-side)}{N_p(n-side)} = e^{-\frac{q_p V}{kT}}$.

$v_{drift} = \frac{q_n \mathcal{E} \tau_n}{m_n}$, yielding the Ohm law: $\vec{j} = \frac{N q_n^2 \tau_n}{m_n} \vec{\mathcal{E}}$. For Hall effect, $\vec{E_{tr}} = -\vec{v}_{drift} \times \vec{B} = -\frac{1}{qN} B \vec{j}$, $R_H = \frac{1}{qN}$.