844: Characterizing polynomials that can be solved

844 I: Rings, Factorization, and the fundamental theorem of Galois theory

§1) The problem of constructing "algebraic closures" of fields
§2) Constructing fields and homomorphisms with Zorn's Lemma
§3) Hilbert's methods for polynomial rings, (without Zorn)
§ [not written: Transcendence degree (Is k[x,y] ≅ k[x,y,z]?)]
§4) $\mathbf{Z}[X]$ is a ufd, (after Gauss)
§5) R[X] is a ufd if R is, (generalizing Gauss' proof)
§6) A Diophantine puzzle
§7) Back to Galois theory: normal and separable extensions
§8) The Fundamental Theorem of Galois Theory via the theorem of the "primitive element"
§9) Galois theory of finite fields

844 II: Identifying (and solving) "solvable" polynomials, eg. solution formulas for cubics and quartics over $\mathbf{Q}$

§10) Polynomials over $\mathbb{Q}$ with solvable group are solvable
§11) The general equation of degree n
§12) Discriminants and the fundamental theorem on symmetric functions
§13) Computing discriminants via "resultants"
§14) "Cardano's formula" for solving a cubic
§15) On the quartic formula and Galois groups of quartics
§16) The Galois group of X^n-1, over $\mathbb{Q}$
§17) A product decomposition for the groups $\mathbf{Z}_n^*$; every finite product of cyclic groups is a Galois group over $\mathbf{Q}$
§18) Fundamental theorem of finite abelian groups, every finite abelian group is a Galois group over $\mathbb{Q}$
§19) Appendix: Summary of proof of Dirichlet's theorem on primes in arithmetic progression, (used in section 17)

844 course notes part 2

§10) Polynomials over $\mathbb{Q}$ with solvable group are solvable by radicals

Last fall we proved that solvability of the Galois group is necessary for a polynomial over $\mathbb{Q}$ to be solvable by radicals; now we prove this condition is also sufficient. The proof works for any field of characteristic zero.

Theorem (Galois): If f in $\mathbb{Q}[X]$ is a polynomial with solvable Galois group, then f is a solvable polynomial. I.e. if all the simple constituents of $G_{\mathbb{Q}}(f)$ are cyclic of prime order, then the splitting field of f lies in a radical extension of $\mathbb{Q}$.

Proof: From the FTGT, we know that to every composition series for the Galois group: $G_{\mathbb{Q}}(f)=H_0 \supset H_1 \supset H_2 \supset \ldots\ldots \supset H_n = \{e\}$, there corresponds a sequence of fields intermediate between $\mathbb{Q}$ and the splitting field L of f: $\mathbb{Q}=L_0 \subset L_1 \subset \ldots\ldots \subset L_n=L$, where L_j = the fixed field of H_j. [Note that because smaller subgroups correspond to bigger subfields, the indices go down for the groups while they go up for the fields.] Then H_j is the Galois group of the extension $L_j \subset L$, and $H_{j+1} \subset H_j$ is the normal subgroup of k-automorphisms of L which restrict to the identity on L_{j+1}. Thus H_j/H_{j+1} is the Galois group of the extension $L_j \subset L_{j+1}$. Since by hypothesis $G_{\mathbb{Q}}(L) = G$ is solvable, each quotient group H_j/H_{j+1} has prime order p_j

We want to add in p_j-th roots of unity for all j, so choose $N=\prod p_j$ and adjoin to L_0 a primitive Nth root of unity ζ. Then consider the sequence of field extensions: $\mathbb{Q}(\zeta)=L_0(\zeta) \subset L_1(\zeta) \subset \ldots \subset L_n(\zeta)=L(\zeta)$. Since $L \subset L(\zeta)$, and $\mathbb{Q}(\zeta)$ is a radical extension of $\mathbb{Q}$, it will suffice to show that each extension $L_j(\zeta) \subset L_{j+1}(\zeta)$ arises by adjoining a single p_j th root β_j, where $(\beta_j)^{p_j} = \alpha_j$ is in $L_j(\zeta)$. First however we must check that by adjoining ζ, we have not changed too greatly the degree of the extension.

Claim: The extension $L_j(\zeta) \subset L_{j+1}(\zeta)$ has degree either p_j or 1.

proof of claim: By assumption, $L_j \subset L_{j+1}$ was Galois of prime degree $p_j = p$ say, so that $L_{j+1} = L_j(\alpha_1,\ldots,\alpha_p)$ where $\{\alpha_1,\ldots,\alpha_p\}$ is the full set of roots of some irreducible polynomial f of degree p, with coeefficients in L_j. Thus $L_j(\zeta) \subset L_{j+1}(\zeta)$ is still a Galois extension since $L_{j+1}(\zeta)=L_j(\zeta,\alpha_1,\ldots,\alpha_p)$, is the splitting field of the same polynomial f.

with coefficients in $L_j \subset L_j(\zeta)$. Since the Galois group is a functor for inclusions of normal extensions, the two inclusions $L_j \subset L_j(\zeta)$,and $L_{j+1} \subset L_{j+1}(\zeta)$, induce a restriction homomorphism of Galois groups $\varphi: G(L_{j+1}(\zeta)/L_j(\zeta)) \to G(L_{j+1}/L_j)$. We claim the restriction homomorphism φ is injective. To see this, recall that in general, if F is the splitting field of a polynomial f over E, then restriction is an injective homomorphism from G(F/E) to Bij($\{\alpha_j\}$), where $\{\alpha_j\}$ is the set of roots of f. In our case, both $L_{j+1}(\zeta)$ and L_{j+1} are splitting fields of the same polynomial f with roots $\{\alpha_1,\ldots,\alpha_p\}$. Thus we have a composition of restrictions
$G(L_{j+1}(\zeta)/L_j(\zeta)) \to G(L_{j+1}/L_j) \to Bij(\{\alpha_1,\ldots,\alpha_p\})$, where both $G(L_{j+1}(\zeta)/L_j(\zeta)) \to Bij(\{\alpha_1,\ldots,\alpha_p\})$ and $G(L_{j+1}/L_j) \to Bij(\{\alpha_1,\ldots,\alpha_p\})$ are injective. Consequently, $G(L_{j+1}(\zeta)/L_j(\zeta)) \to G(L_{j+1}/L_j)$ is also injective. Since $G(L_{j+1}/L_j) \cong \mathbf{Z}_p$, it follows that $G(L_{j+1}(\zeta)/L_j(\zeta))$ is isomorphic to either $\mathbf{Z}_p$ or to {0}, and hence the Galois extension $L_j(\zeta) \subset L_{j+1}(\zeta)$ has degree either $p = p_j$, or 1. **QED Claim.**

Now, since $\mathbb{Q} = L_0 \subset L_0(\zeta) = \mathbb{Q}(\zeta)$ is the splitting field of $X^N - 1$, hence a radical Galois extension, it suffices to prove the following:

Proposition: If an extension $k \subset L$ is Galois of prime degree p, where $\mathbb{Q} \subset k \subset L \subset \mathbb{C}$, and k contains all pth roots of unity, then $L = k(\beta)$ where β^p is an element of k

All the proofs I have seen of this use the following:
Lemma: Given $\mathbb{Q} \subset k \subset L \subset \mathbb{C}$, where $k \subset L$ is Galois of prime degree p, and k contains all pth roots of 1, then for any $\sigma \neq id$ in $G_k(L) \cong \mathbf{Z}_p$, there is an α in L such that $\sigma(\alpha) = \zeta\alpha \neq \alpha$, for some ζ in k.

Proof of the proposition (Assuming the lemma): Since $[L:k] = p$, is prime, all we have to do is find an element α of L, such that α is not in k but α^p is in k. [Then $k(\alpha)$ will be a subfield of L larger than k, hence $[k(\alpha):k] = p$, so $k(\alpha) = L$.]
Remember that an element of L lies in k iff it is fixed by $G = G_k(L)$. Since $G \cong \mathbf{Z}_p$, if $\sigma \neq id$ is any non trivial element of G, then σ generates G, hence an element β of L belongs to k iff $\sigma(\beta) = \beta$. Thus we just need to find α in L such that $\sigma(\alpha) \neq \alpha$, but $\sigma(\alpha^p) = \alpha^p$. If α is chosen as in the lemma above, then we have $\sigma(\alpha) = \zeta\alpha \neq \alpha$, in

particular $\alpha \neq 0$. To see that $\sigma(\alpha P) = \alpha P$, note first that $\sigma P = id$ since $G \cong \mathbb{Z}_p$, hence $\alpha = \sigma P(\alpha) = \zeta P \alpha$, where $\alpha \neq 0$, so $\zeta P = 1$. Then $\sigma(\alpha P) = (\sigma(\alpha))P = (\zeta\alpha)P = \zeta P \alpha P = \alpha P$ **QED** (modulo proving the lemma).

We will give three proofs of the lemma
1st proof of lemma: If you know some linear algebra, the lemma asserts the existence of some "eigenvectors" for σ. To prove they exist, we appeal to a fact about linear transformations, whose proof will be presented later.

Fact: If k is a field of characteristic zero, containing all pth roots of unity, and $\sigma \neq id$ is a linear transformation of a finite dimensional k-vector space V such that $\sigma P = id$, then there is a primitive pth root of unity ζ, and an "eigenvector" α in V such that $\sigma(\alpha) = \zeta\alpha \neq \alpha$.

[Sketch of proof of the Fact: Since $\sigma P = id$, but $\sigma \neq id$, σ satisfies the polynomial $X^P - 1 = 0$, but not $X-1$. Hence the minimal polynomial f of σ over k divides X^P-1, but is not $X-1$, and thus f has a root $\zeta \neq 1$ which is a primitive pth root of unity. By the "Cayley - Hamilton" theorem, σ also satisfies its characteristic polynomial $\chi(X)$, which is thus a multiple of the minimal polynomial f, so that ζ is also a root of χ. It follows that ζ is an eigenvalue of σ, hence some corresponding eigenvector α exists such that $\sigma(\alpha) = \zeta\alpha \neq \alpha$.]

Using this fact, to prove the lemma it suffices to note that if $k \subset L$ is a finite extension, then L is a finite dimensional k vector space and any σ in $G_k(L)$ defines a k- linear transformation of L.
QED for 1st proof of Lemma.

For the second proof of the lemma, we will use an important, but easy, fact about "group characters", which we will prove completely.
Definition: If G is a group and L is a field, a "character of G in L" is a homomorphism $\sigma: G \to L^*$ from G into the multiplicative group L^*.

Sublemma(E. Artin / R. Dedekind): Any set $\{\sigma_1,\ldots,\sigma_n\}$ of distinct characters of a group G in a field L, is L-linearly independent.
proof of sublemma: The statement means if $a_1,\ldots,a_n$ are elements of L which are not all zero, then the map $G \to L$ defined by $\Sigma a_i\sigma_i$ does not take every element of G to 0. It is true for $n = 1$, since then for

every x in G, $\sigma_1(x) \neq 0$, so $a \neq 0$ implies $a\sigma_1(x) \neq 0$.
Now let $n > 1$, and assume

$$(*)\ a_1\sigma_1 + \ldots\ldots + a_n\sigma_n = 0,$$

is a dependency relation, but that no such relation holds with fewer than n characters. Then no a_i can be zero, since if say $a_n = 0$, we would have a dependency relation among $\sigma_1, \ldots \sigma_{n-1}$. Since $\sigma_1 \neq \sigma_2$, there is some z in G with $\sigma_1(z) \neq \sigma_2(z)$. Since for every x in G, we have $0 = a_1\sigma_1(xz) + \ldots + a_n\sigma_n(xz) = a_1\sigma_1(z)\sigma_1(x) + \ldots + a_n\sigma_n(z)\sigma_n(x)$, and since $\sigma(z) \neq 0$, hence

$$(**)\ a_1\sigma_1(z)\sigma_1 + \ldots + a_n\sigma_n(z)\sigma_n = 0,$$

is another dependency relation among the characters $\sigma_1, \ldots \sigma_n$. Now if we multiply the first relation (*) by $\sigma_1(z)$ and subtract the result from the second relation (**), the first terms $a_1\sigma_1(z)\ \sigma_1$ cancel and we have a relation among $\sigma_2, \ldots, \sigma_n$

$$(***)\ a_2\ [\sigma_2(z) - \sigma_1(z)]\ \sigma_2 + \ldots\ldots = 0.$$

Since $\sigma_1(z) \neq \sigma_2(z)$ by choice, and $a_2 \neq 0$, the first coefficient of (***) is not zero, and we have a shorter relation, which is a contradiction.
QED. sublemma.

2nd proof of the lemma: First we want to cook up an element α in L such that $\sigma(\alpha) = \zeta\alpha$. Equivalently we want an α such that $\zeta^{-1}\sigma(\alpha) = \alpha$. So we want an element of L which is left unchanged when we apply $\zeta^{-1}\sigma$ to it. Choose $\zeta \neq 1$ a primitive pth root of unity in k, and recall that $\sigma^p = id$. Then $\tau = \zeta^{-1}\sigma$ is a k linear transformation on L with $\tau^p = id$, hence the elements id, τ, $\tau^2, \ldots, \tau^{p-1}$ form a group whose elements are simply permuted by τ Thus if we add them up we get a linear transformation $\varphi = id + \tau + \tau^2 + \ldots + \tau^{p-1}$ which is left unchanged when we apply τ to it. I.e. $\tau(\varphi) = \tau + \tau^2 + \ldots + \tau^p = \tau + \tau^2 + \ldots + id = \varphi$. Thus if β is any element of L, and $\alpha = \varphi(\beta)$, we get $\tau(\alpha) = \alpha$. Thus $\tau(\alpha) = \zeta^{-1}\sigma(\alpha) = \alpha$, so $\sigma(\alpha) = \zeta\alpha$.
The only thing remaining is to show that we can choose α so that $\zeta\alpha \neq \alpha$. Since $\zeta \neq 1$, this can only happen if $\alpha = 0$., and here is where the sublemma comes in. Since id, $\sigma, \ldots, \sigma^{p-1}$ are distinct characters of the group L^* in L, they are independent over L, so the linear transformation $\varphi = id + \zeta^{-1}\ \sigma + \zeta^{-2}\ \sigma^2 + \ldots + \zeta^{1-p}\ \sigma^{p-1}$ is not zero. Hence we can choose β so that $\alpha = \varphi(\beta) \neq 0$.
QED for 2nd proof of Lemma.

Remark: This proof shows that every pth root of unity is an eigenvalue of σ, a stronger statement than the first proof gave.

Terminology: For ζ a pth root of unity, the expression $\varphi(\beta) = \beta + \zeta\,\sigma(\beta) + \zeta^2\,\sigma^2(\beta) + \ldots + \zeta^{p-1}\,\sigma^{p-1}(\beta)$ is called a "LaGrange resolvent". The only place we needed the Artin/Dedekind result was to insure the existence of a non zero LaGrange resolvent with $\zeta \neq 1$. In our next and last proof, possibly the classical one, we will show directly that some such resolvent must be non zero

3rd proof of the lemma: Let β be any element of L not belonging to k, and $\zeta \neq 1$ any primitive pth root of unity Then $1, \zeta, \ldots, \zeta^{p-1}$ are the distinct pth roots of unity, and all but 1 are primitive. Now we simply form all p of the corresponding LaGrange resolvents, $\alpha_0, \alpha_1, \ldots, \alpha_{p-1}$, and prove at least one of $\alpha_1, \ldots, \alpha_{p-1}$ is non zero by adding them up. I.e.

$$\alpha_0 = \beta + \sigma(\beta) + \sigma^2(\beta) + \ldots + \sigma^{p-1}(\beta)$$

$$\alpha_1 = \beta + \zeta\sigma(\beta) + \zeta^2\,\sigma^2(\beta) + \ldots + \zeta^{p-1}\,\sigma^{p-1}(\beta)$$

$$\alpha_2 = \beta + \zeta^2\sigma(\beta) + \zeta^4\,\sigma^2(\beta) + \ldots + \zeta^{2(p-1)}\,\sigma^{p-1}(\beta)$$

$$\alpha_j = \beta + \zeta^j\sigma(\beta) + \zeta^{2j}\,\sigma^2(\beta) + \ldots + \zeta^{j(p-1)}\,\sigma^{p-1}(\beta)$$

$$\alpha_{p-1} = \beta + \zeta^{p-1}\sigma(\beta) + \zeta^{2(p-1)}\,\sigma^2(\beta) + \ldots + \zeta^{(p-1)(p-1)}\,\sigma^{p-1}(\beta)$$

Observe:
(i) In each row and column to the right of the equal signs, except the first, the coefficients occurring are the full set of distinct pth roots of unity. For example, in the jth row and column, the coefficients are $1, \zeta^j, \zeta^{2j}, \ldots, \zeta^{j(p-1)}$. (Here, j runs from 0 to p-1.)
(ii) By the formula for summing geometric series, the sum of the pth roots of unity $1 + \zeta^j + \zeta^{2j} + \ldots + \zeta^{j(p-1)} = (1-\zeta^{jp})/(1-\zeta^j) = 0$.
(iii) $\sigma(\alpha_0) = \alpha_0$, so α_0 belongs to k
Thus the sum of all the columns is zero except the first, whose sum

is $p\beta$. Hence $\alpha 0+\alpha 1+....+\alpha p-1 = p\beta$. Since p is in k, but β is not, the right side is not in k, hence neither is the left side. Since $\alpha 0$ does belong to k, some α_j with $j \geq 1$ is non zero. For this αj, $\sigma(\alpha j) = \zeta^{-j}\alpha j \neq \alpha j$, which proves the lemma.
QED for 3rd proof of lemma.

Remarks: (i) Note that this proof complements the second proof, since the second proof showed that for each ζ there is a corresponding β in L for which the Lagrange resolvent is non zero, while the third proof shows that for every β in L-k there is a corresponding $\zeta \neq 1$ for which the Lagrange resolvent is non zero. (ii) The theorem we have just proved, that a polynomial over $\mathbb{Q}$ with solvable Galois group G is solvable by radicals, is true for polynomials over all fields of characteristic zero, and even over fields of prime characteristic q, provided q does not occur among the orders of the simple constituents of G. The proof is the same as we have given. To see what goes wrong in the third proof for example, in characteristic p the quantity $p\beta$ would be zero, hence it would lie in k, and we would have no contradiction. As for the first proof, there are matrices of order 2, in characteristic 2, whose eigenvalues are all 1, such as the linear transformation of k^2, where $k = \mathbb{Z}_2$, defined by $(1,0)\mapsto(1,0)$, and $(0,1)\mapsto(1,1)$. In the second proof, primitive pth roots of unity ζ do not exist in characteristic p, since then $(X^p-1) = (X-1)^p$, and 1 is the only pth root of unity.

Application: Fundamental Theorem of Algebra:
Claim: The field $\mathbb{C} = \mathbb{R}(i) \cong \mathbb{R}[X]/(X^2+1)$, is algebraically closed.
We need two facts:
Fact (a): Every polynomial of odd degree over $\mathbb{R}$ has a root in $\mathbb{R}$.
Fact (b): Every element of $\mathbb{C}$ has a square root in $\mathbb{C}$.

These two facts have Galois theoretic formulations as follows:
Lemma A): $\mathbb{R}$ has no non trivial finite extensions of odd degree.
proof: If $\mathbb{R} \subset F$ is a finite extension of odd degree, the primitive element theorem implies F is generated by one element whose minimal irreducible polynomial over $\mathbb{R}$ has odd degree. By fact a) above, this minimal polynomial has a root, hence is irreducible iff it has degree one, so $\mathbb{R} = F$. **QED.**

is $p\beta$. Hence $\alpha_0+\alpha_1+\ldots+\alpha_{p-1} = p\beta$. Since p is in k, but β is not, the right side is not in k, hence neither is the left side Since α_0 does belong to k, some α_j with $j \geq 1$ is non zero. For this α_j, $\sigma(\alpha_j) = \zeta^{-j}\alpha_j \neq \alpha_j$, which proves the lemma.
QED for 3rd proof of lemma.

Remarks: (i) Note that this proof complements the second proof, since the second proof showed that for each ζ there is a corresponding β in L for which the Lagrange resolvent is non zero, while the third proof shows that for every β in L-k there is a corresponding $\zeta \neq 1$ for which the Lagrange resolvent is non zero. (ii) The theorem we have just proved, that a polynomial over $\mathbb{Q}$ with solvable Galois group G is solvable by radicals, is true for polynomials over all fields of characteristic zero, and even over fields of prime characteristic q, provided q does not occur among the orders of the simple constituents of G. The proof is the same as we have given. To see what goes wrong in the third proof for example, in characteristic p the quantity $p\beta$ would be zero, hence it would lie in k, and we would have no contradiction. As for the first proof, there are matrices of order 2, in characteristic 2, whose eigenvalues are all 1, such as the linear transformation of k^2, where $k = \mathbb{Z}_2$, defined by $(1,0)\mapsto(1,0)$, and $(0,1)\mapsto(1,1)$. In the second proof, primitive pth roots of unity ζ do not exist in characteristic p, since then $(X^p-1) = (X-1)^p$, and 1 is the only pth root of unity.

Application: Fundamental Theorem of Algebra:
Claim: The field $\mathbb{C} = \mathbb{R}(i) \cong \mathbb{R}[X]/(X^2+1)$, is algebraically closed.
We need two facts:
Fact (a): Every polynomial of odd degree over $\mathbb{R}$ has a root in $\mathbb{R}$.
Fact (b): Every element of $\mathbb{C}$ has a square root in $\mathbb{C}$.

These two facts have Galois theoretic formulations as follows:
Lemma A): $\mathbb{R}$ has no non trivial finite extensions of odd degree.
proof: If $\mathbb{R} \subset F$ is a finite extension of odd degree, the primitive element theorem implies F is generated by one element whose minimal irreducible polynomial over $\mathbb{R}$ has odd degree. By fact a) above, this minimal polynomial has a root, hence is irreducible iff it has degree one, so $\mathbb{R} = F$ QED.

Cor: Every finite Galois extension of $\mathbb{R}$ has degree 2^n, for $n \geq 0$.
proof: Let $\mathbb{R} \subset K$ be a finite Galois extension, with group G, and $S \subset G$ is a Sylow 2-subgroup. If $F \subset K$ is the fixed field of S, then $[K:F] = \#(S)$, and hence $[F:\mathbb{R}] = \#(G)/\#(S)$ is odd. Thus $\mathbb{R} = F$, $G = S$, and $[K:\mathbb{R}] = \#(S) = 2^n$, where $n \geq 0$. QED.

Lemma B): $\mathbb{C}$ has no quadratic extensions.
proof: If $\mathbb{C} \subset K$ is a quadratic extension, then by our lemma in the proof of Galois' theorem above, since $\mathbb{C}$ contains -1 (a primitive square root of unity) K is generated by the square root of an element α of $\mathbb{C}$. But by fact (b) above, α is in $\mathbb{C}$, so $\mathbb{C} = K$. QED.

Cor: The field $\mathbb{C}$ has no Galois extensions of degree 2^n, for $n > 0$.
proof: If $\mathbb{C} \subset K$ is Galois of degree 2^n, with group G, all simple constituents of G are $\cong \mathbb{Z}_2$. Thus the fundamental theorem of Galois theory implies such an extension decomposes into a tower of quadratic extensions. Since $\mathbb{C}$ has no quadratic extensions $n = 0$. QED.

Corollary: The field $\mathbb{C}$ is algebraically closed.
proof: It suffices to show every finite extension of $\mathbb{C}$ equals $\mathbb{C}$. If $\mathbb{C} \subset L$ is any finite extension, then $\mathbb{R} \subset L$ is also finite and can be enlarged to a finite Galois extension $\mathbb{R} \subset K$, where $[K:\mathbb{R}] = 2^n$, by the corollary of lemma A. Then $\mathbb{C} \subset K$ is also Galois and $[K:\mathbb{C}] = 2^{n-1}$. By the corollary of lemma B, $n = 1$, and $K = L = \mathbb{C}$. QED.

Exercise #107) Prove: **(a)** Every polynomial of odd degree over $\mathbb{R}$ has a root in $\mathbb{R}$, and **(b)** Every element of $\mathbb{C}$ has a square root in $\mathbb{C}$, hence $\mathbb{C}$ has no quadratic extensions (Hint: Use the "intermediate value theorem" from calculus for (a).)

This completes our discussion of the proof of Galois' general theorem on solvability of polynomials with solvable Galois groups. We want to show in the next section how to actually produce solution formulas, for the general equations of degree three and four.

§11) Elementary symmetric polynomials and The Galois group of the "general equation" of degree n

We know now, since every subgroup of S_4 is a solvable group, that

every polynomial over $\mathbb{Q}$ of degree ≤ 4 is solvable by radicals, which means that its solutions lie in a field obtained from $\mathbb{Q}$ by the successive adjunction of square roots and cube roots. Consequently, the solutions of such a polynomial have rational expressions in terms of rational numbers, square roots, and cube roots. But how do we find such expressions explicitly? It would be rather unwieldy if every polynomial of degree three, say, had a different solution formula. We would greatly prefer a "universal" solution formula that works for all polynomials of degree three. In that case we really should be trying to solve, not all *particular* polynomials, but the one "general" polynomial of degree three, the one whose coefficients are letters, i.e. variables.
The general polynomial of degree n is a polynomial $f(X) = X^n+a_1X^{n-1}+\ldots+a_{n-1}X+a_n$, with letters a_i for coefficients, where it is understood that any rational number may be substituted for each letter a_i. Thus the coefficents are themselves independent variables over $\mathbb{Q}$. Hence this is a polynomial over the field $k = \mathbb{Q}(a_1,\ldots,a_n)$ of rational functions in the n independent variables $a_1,\ldots,a_n$ over $\mathbb{Q}$. This field is the fraction field of the polynomial ring $\mathbb{Q}[a_1,\ldots,a_n]$, in those independent variables, and f belongs to the polynomial ring $\mathbb{Q}(a_1,\ldots,a_n)[X] = k[X]$, where X is an element which is transcendental over k. Thus, with reference to $\mathbb{Q}$, X is another variable, independent of the a_i. Now since we want to solve f, our first question is: what is the Galois group of f over k? Since f has degree n over k, and k is a field, we know the Galois group is isomorphic to a subgroup of S_n. We claim that in fact $G(f) \cong S_n$

Theorem: The Galois group of the "general" polynomial $f(X) = X^n+a_1X^{n-1}+\ldots+a_{n-1}X+a_n$, with coefficients in $\mathbb{Q}(a_1,\ldots,a_n)$, (where the a_i are independent transcendentals over $\mathbb{Q}$), is $\cong S_n$.
proof: To see this, first let $\alpha_1,\ldots,\alpha_n$ be roots of f in some extension field of k, and let $L = k(\alpha_1,\ldots,\alpha_n)$ be the splitting field. Our first remark is that $L = \mathbb{Q}(\alpha_1,\ldots,\alpha_n)$. I.e., at first we see that $L = k(\alpha_1,\ldots,\alpha_n) = \mathbb{Q}(a_1,\ldots,a_n,\alpha_1,\ldots,\alpha_n)$, but since $f(X) = X^n+a_1X^{n-1}+\ldots+a_{n-1}X+a_n = \Pi(X-\alpha_i)$, it follows by multiplying out that each a_i is $\pm$ a sum of products of the $\{\alpha_j\}$. In fact $a_n = (-1)^n\Pi\alpha_j$, and $a_1 = -(\Sigma\alpha_j)$, while $a_2 = \Sigma\alpha_i\alpha_j$, summed over all $i<j$. Similarly $a_s = (-1)^s$(the sum of all products of s distinct α's). In any case, each a_s

is in the field generated by the α's over $\mathbb{Q}$, so $L = \mathbb{Q}(\alpha_1, \ldots, \alpha_n)$, as claimed. Thus we want to compute the Galois group of f, i.e. the Galois group of the extension $\mathbb{Q}((a_1,\ldots,a_n) \subset \mathbb{Q}(\alpha_1,\ldots,\alpha_n)$, where the $\{a_j\}$ are independent variables over $\mathbb{Q}$, and the α's are related to the a's as above

Now let's take a dual point of view: consider a set of n independent transcendentals $\beta_1,\ldots,\beta_n$ over $\mathbb{Q}$, and let $L = \mathbb{Q}(\beta_1,\ldots,\beta_n)$. Form the polynomial $g(X) = \prod(X-\beta_j) = X^n - \sigma_1 X^{n-1} + \ldots + (-1)^{n-1}\sigma_{n-1}X + (-1)^n\sigma_n$, where as above, each σ_t is the sum of all products of t distinct β's. I.e. $\sigma_1 = -(\Sigma\beta_j)$, $\sigma_2 = \Sigma\beta_i\beta_j, \ldots$, and $\sigma_n = (-1)^n\prod\beta_j$. These functions σ_t in the variables $\{\beta_j\}$ are called the "elementary symmetric functions" of the β_j, because they are among the simplest expressions which are unchanged by every permutation of the β's, and because they generate all symmetric functions in a sense which is made precise in the "fundamental theorem" in section #13) In particular, the expressions $\sigma_s(\beta_1,\ldots,\beta_n)$ are invariant under the action of the symmetric group S_n on the β's.
Thus the relation between the solutions and the coefficients is the same for g(X) as for f(X), the only difference being. for f we assumed the *coefficients* were independent variables, while for g we assumed the *solutions* were independent variables Now we claim: (i) the Galois group of g(X) is $\cong S_n$, and (ii) f(X) and g(X) have isomorphic Galois groups.
To see (i), recall that the Galois group G(g) is the group of the extension $\mathbb{Q}(\sigma_1,\ldots,\sigma_n) \subset \mathbb{Q}(\beta_1,\ldots,\beta_n)$, and is isomorphic to a subgroup of $Bij(\{\beta_j\}) \cong S_n$. I.e. since every automorphism of $\mathbb{Q}(\beta_1,\ldots,\beta_n)$ which fixes $\mathbb{Q}(\sigma_1,\ldots,\sigma_n)$ must permute the β's, restriction gives a homomorphism $G(g) \to Bij(\{\beta_j\}) \cong S_n$, which we know is injective Thus to prove $G(g) \cong S_n$, we just have to show this restriction homomorphism is surjective, i.e. that every permutation of the β's extends to an automorphism of the extension $\mathbb{Q}(\sigma_1,\ldots,\sigma_n) \subset \mathbb{Q}(\beta_1,\ldots,\beta_n)$. But since the $\{\beta_j\}$ are independent transcendentals over $\mathbb{Q}$, every permutation of the β's extends to an automorphism of the polynomial ring $\mathbb{Q}[\beta_1,\ldots,\beta_n]$, and hence to an automorphism of the fraction field $\mathbb{Q}(\beta_1,\ldots,\beta_n)$. Moreover, since the functions σ_t are invariant under all these permutations, every element of S_n yields an automorphism which fixes the field

$\mathbb{Q}(\sigma_1,\ldots,\sigma_n)$. Hence $G(g) \cong S_n$. QED (i).

Now to prove (ii), that f(X) and g(X) have the same Galois group, we want to show the extensions $\mathbb{Q}(\sigma_1,\ldots,\sigma_n)\subset\mathbb{Q}(\beta_1,\ldots,\beta_n)$ and $\mathbb{Q}(a_1,\ldots,a_n)\subset\mathbb{Q}(\alpha_1,\ldots,\alpha_n)$ are isomorphic. Since these are the splitting fields of the polynomials g and f, it suffices to find an isomorphism between $\mathbb{Q}(a_1,\ldots,a_n)$ and $\mathbb{Q}(\sigma_1,\ldots,\sigma_n)$ which carries the polynomial f(X) into g(X). Thus we need an isomorphism of fields $\mathbb{Q}(a_1,\ldots,a_n) \cong \mathbb{Q}(\sigma_1,\ldots,\sigma_n)$ that carries each coefficient a_i of f into the corresponding coefficient $(-1)^i\sigma_i$ of g. It will follow that there exists an isomorphism of splitting fields $\mathbb{Q}(\alpha_1,\ldots,\alpha_n) \cong \mathbb{Q}(\beta_1,\ldots,\beta_n)$, and since the automorphism group of $\mathbb{Q}(\beta_1,\ldots,\beta_n)$ over $\mathbb{Q}(\sigma_1,\ldots,\sigma_n)$ contains arbitrary permutations of the β's, we can even choose the automorphism $\mathbb{Q}(\alpha_1,\ldots,\alpha_n) \cong \mathbb{Q}(\beta_1,\ldots,\beta_n)$ to carry each α_i into β_i. So finally, to prove the existence of an isomorphism $\mathbb{Q}(a_1,\ldots,a_n) \cong \mathbb{Q}(\sigma_1,\ldots,\sigma_n)$, we just need to show that the $\{(-1)^i\sigma_i\}$ are themselves independent transcendentals over $\mathbb{Q}$. Let's be precise about this:

Definition: A collection of elements $\tau_1,\ldots,\tau_n$ in an extension field of k, are independent transcendentals over k iff the only polynomial $f(T_1,\ldots,T_n)$ in the polynomial ring $k[T_1,\ldots,T_n]$ which vanishes when each T_i is replaced by τ_i, is the trivial polynomial $f = 0$.

Remark: (i) Another way to think of the previous definition is that the $\tau_1,\ldots,\tau_n$ are *algebraically independent* over k, which means simply that any finite set of distinct monomials in the τ's are *linearly* independent over k
(ii) Still another way to think of this definition is that $\tau_1,\ldots,\tau_n$ are independent transcendentals over k iff any for choice of elements $\alpha_1,\ldots,\alpha_n$ in any extension field L over k, there is a unique k homomorphism $k[\tau_1,\ldots,\tau_n]\to L$ taking each τ_i to α_i. (It is because of this point of view, which we think of as "substituting" the α's for the τ's, that we sometimes call such τ's "independent variables" over k.)

Lemma: If $X_1,\ldots,X_n$ are independent transcendentals over k, and σ_j is the jth elementary symmetric function of the X's, then $\{(-1)^j\sigma_j\}j=1,..n$, are also independent transcendentals over k.

proof: Let $F(T_1,\dots,T_n)$ be any polynomial in n variables over k such that $F(-\sigma_1,\dots,(-1)^n\sigma_n) = F(-\sigma_1(X_1,\dots,X_n),\dots,(-1)^n\sigma_n(X_1,\dots,X_n)) = 0$ in $k[X_1,\dots,X_n]$. Then we should get zero by substituting for the X's, any elements of any field extension of k; i.e. we should get zero when substituting α_j for X_j, where the α's are the solutions of the polynomial f over $\mathbb{Q}(a_1,\dots,a_n)$, given above. I.e. then $F(-\sigma_1(\alpha_1,\dots,\alpha_n),\dots,(-1)^n\sigma_n(\alpha_1,\dots,\alpha_n)) = 0$ in $\mathbb{Q}(\alpha_1,\dots,\alpha_n)$. However, since the relation between the solutions and the coefficients of f and g are identical, we know $(-1)^j\sigma_j(\alpha_1,\dots,\alpha_n) = a_j$ for every j. Thus $0 = F(-\sigma_1(\alpha_1,\dots,\alpha_n),\dots,(-1)^n\sigma_n(\alpha_1,\dots,\alpha_n)) = F(a_1,\dots,a_n)$ in $\mathbb{Q}(\alpha_1,\dots,\alpha_n)$. Since we have assumed the elements $\{a_j\}$ were independent transcendentals over $\mathbb{Q}$, the polynomial F they satisfy must be identically zero, and thus by definition, the elements $\{(-1)^j\sigma_j\}$ are also independent transcendentals over $\mathbb{Q}$. [Of course the $\{\sigma_j\}$, without the minus signs, are also independent transcendentals over $\mathbb{Q}$.]
QED lemma.

The previous lemma completes the proof that the Galois group of the general polynomial of degree n is $\cong S_n$. **QED theorem.**

Remarks: (i) Since the Galois group $G(f) = S_n$, in particular G(f) acts transitively on the roots, hence the general polynomial is irreducible.
(ii) This proof appears to work also in characteristic p, in particular the general polynomial f is separable in charac. p, with $G(f) \cong S_n$.
(iii) In spite of the fact that the general polynomial of degree 3 say has Galois group S3, in charac. 3, we do not expect the polynomial to be solvable by radicals, since 3 divides #(S3). In fact, even the general quadratic equation cannot be solved by radicals in charac. 2

Exercise #108) (i) If $k \subset L$ is a finite Galois extension with Galois group G, and if $H \subset G$ is any subgroup of G, prove there exists a finite Galois extension $E \subset F$ whose Galois group is isomorphic to H.
(ii) Prove, if G is any finite group at all, then there is a finite Galois extension $E \subset F$ whose Galois group is isomorphic to G, and that E can be chosen to have any desired characteristic.
Unsolved problem: To determine whether for every finite group G, there is a finite Galois extension $\mathbb{Q} \subset L$ with Galois group $\cong G$. We will present below a proof of the special case where G is abelian.

§12) Discriminants, and the Fundamental Theorem on Symmetric Functions

Although S_n is not solvable for $n > 4$, there is always one normal subgroup of S_n that yields a cyclic quotient, namely $A_n \subset S_n$. This means for the general polynomial of degree n, there is always an intermediate field between the coefficient field and the splitting field that is quadratic over the coefficient field, and (in charac. ≠ 2) which is obtained by adjoining a square root of an expression in the coefficients. In degrees 2, 3, 4, where the general polynomial is solvable by radicals (at least in charac. ≠ 2, 3), the first step in finding a solution formula is to determine explicitly a generator of this first radical extension. Let us assume below that the characteristic is neither 2 nor 3.

If $\alpha_1,\dots,\alpha_n$ are the roots of the general n^{th} degree polynomial f, to find a generator of the fixed field of A_n, we seek a rational expression δ in the α's, which is invariant under even permutations of the α's but not under odd ones. Moreover, since we want to show the extension is radical, we want δ such that $\delta^2 = D$ lies in the field fixed by all permutations. Thus, for any permutation φ we should have $\delta^2 = \varphi(\delta^2) = \varphi(\delta)\varphi(\delta)$, so that δ and $\varphi(\delta)$ have the same square, and yet $\varphi(\delta) \neq \delta$. It follows that we must have $\varphi(\delta) = -\delta$. So we look for an expression δ in the α's which changes sign when we transpose two of the α's. Now, if for some particular polynomial two roots are equal, we would have an expression δ which changes sign when those two roots are transposed, yet also stays the same! The only such field element is zero. Thus we are looking for an expression in the α's which changes sign when two α's are interchanged, and equals zero when two α's are given equal values.

If we still haven't guessed it, we can consider the familiar case of degree $n = 2$, the general (monic) quadratic equation x^2+bx+c, whose roots are $\alpha_1 = (1/2)(-b + [(b^2-4c)^{1/2}])$, $\alpha_2 = (1/2)(-b - [(b^2-4c)^{1/2}])$. These two roots are equal precisely when $(b^2-4c)^{1/2} = 0$. It follows that good candidates for δ, D are $\delta = (b^2-4c)^{1/2}$, and $D = \delta^2 = b^2-4c$. We see also that $\delta = \alpha_1-\alpha_2$. What more logical way is there to get an element which vanishes when two roots are equal, than to subtract the two roots from each other! Note that δ also changes sign when we transpose the two roots.

Now let's find an analogous element δ for a monic polynomial of degree $n=3$, whose roots are $\alpha_1,\alpha_2,\alpha_3$. We need to involve the differences $(\alpha_1-\alpha_2)$, $(\alpha_1-\alpha_3)$, and $(\alpha_2-\alpha_3)$, in such a way that the element δ vanishes when any two of the roots are equal, so a plausible candidate for δ is the "difference product" $\delta = (\alpha_1-\alpha_2)(\alpha_1-\alpha_3)(\alpha_2-\alpha_3)$. Then $D = \delta^2 = (\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2$. The squared element $D = \delta^2$ (lying in the *coefficient field*) is called the "discriminant" of the polynomial f

Definition: If f is a monic polynomial over a field k, with roots $\alpha_1,\ldots,\alpha_n$ in some splitting field, the <u>discriminant</u> of f is the expression $D = \Pi(\alpha_i-\alpha_j)^2$, where the product runs over all i,j with $1 \le i < j \le n$.

Remarks: (i) The discriminant $D = \delta^2$ is always an element of the coefficient field k In fact, for the case of the (separable) general polynomial, D is fixed by the entire Galois group S_n, hence D belongs to the coefficient field. Since this holds for the general polynomial it holds for all particular polynomials We will see later that D is actually an integral polynomial in the coefficients of f, and will even find an explicit expression for D in low degrees.
(ii) In charac. p=2, note that δ itself lies in the coefficient field of the general polynomial. I.e. a transposition takes δ to $-\delta$, but in characteristic 2, we have $\delta = -\delta$. Thus δ is fixed by the Galois group of the (separable) general polynomial, hence lies in the coefficient field. In particular δ does not provide a generator of the splitting field of the general quadratic polynomial in charac. 2.
(iii) We sometimes encounter the similar expression $\Pi(\alpha_i-\alpha_j)$, where i,j run over all pairs with $i \neq j$. However this expression is slightly different from D since it contains the factors $(\alpha_i-\alpha_j)(\alpha_j-\alpha_i)$ in place of $(\alpha_i-\alpha_j)^2$ Hence this new expression is not D but $(-1)^{n(n-1)/2}D$, a fact which is not noticed in some books.

Now recall that a "solvable" polynomial f is one whose roots lie in a field obtained in stages by adjoining generators, each of which has a power lying in the previous subfield. Hence in order to find a formula for the roots, involving only the coefficients of f, we could express the appropriate power of each of these generators in terms

of the generator of the previous subfield, until we get down to the coefficient field. Hence the last step in every case would seem to be expressing the square of the quadratic generator $\delta^2 = D$ in terms of the coefficients. (A formula for the solutions can also be obtained directly, without the intermediate step of finding one for the discriminant, but having a formula for the discriminant is also useful later for computing Galois groups of special cubics [cf. ex. #110 below].) Computing the discriminant formula is not an entirely trivial task, and offers a convincing demonstration of not just how clever and insightful the mathematicians of old were, but also how diligent and skillful they were at performing lengthy and arduous calculations accurately. I know about four different approaches to this general calculation which I will describe. Fortunately today one can also appeal to a computer algebra program like Mathematica or Maple, to make carrying these calculations out much less odious. Finally, in the special case of the discriminant of a cubic with no quadratic term, we will give a relatively easy determinant calculation using the "resultant", and another even easier calculation pointed out by Robert Varley.

Theorem: (i) If $\delta^2 = D$ is the discriminant of the polynomial f, there is an explicit computational procedure for expressing D as a polynomial in the coefficients of f. In fact, for each n, there is a universal formula for D, which holds for all f of degree n.
(ii) If n = 2, and $f = X^2+bX+c$, then $D = b^2-4c$.
(iii) If n = 3 and $f = X^3+pX+q$, then $D = -4p^3 - 27q^2$.
(iv) If n = 3, and $f = X^3+pX^2+qX+r$, then
$$D = p^2q^2 - 4q^3 - 4p^3r + 18pqr - 27r^2.$$
proof: Consider the general polynomial f of degree n. We know the discriminant D is fixed by all elements of the Galois group, hence lies in the coefficient field, so D is the quotient of some pair of polynomials in the coefficients of f. But we want to express D as one polynomial in the coefficients, and we want to give an explicit procedure for producing that polynomial. We proceed as follows. By definition $D = \Pi(\alpha_i-\alpha_j)^2$, product over all i,j: $1 \leq i < j \leq n$, where the α's, the roots of f, are independent transcendentals over the base field. Permuting the α's merely permutes the factors of D in some way, hence leaves D fixed. Thus D is a "symmetric polynomial" in the α's, in the sense of the following definition.

Definition: A "symmetric polynomial" in n variables $X_1,\ldots,X_n$ over a field k, is a polynomial in $k[X_1,\ldots,X_n]$ which is left fixed by every permutation of $X_1,\ldots,X_n$.

Definition: The j^{th} "elementary symmetric function" of the variables $X_1,\ldots,X_n$, is $\sigma_j = \sigma_j(X_1,\ldots,X_n) = \Sigma X_{i(1)}\ldots X_{i(j)}$, the sum of all products of j different variables among the $X_1,\ldots,X_n$.
[In particular, $\sigma_1 = X_1+\ldots+X_n$, $\sigma_2 = \Sigma\, X_iX_j$, for all $1 \le i < j \le n$, and $\sigma_n = \Pi X_i$.]

Since, up to sign, the coefficients of f are just the elementary symmetric functions in the roots, part (1) of the theorem above follows from the so called "fundamental theorem on symmetric functions", proved next. We will give two proofs of this result.
Proposition: Every "symmetric" polynomial in $k[X_1,\ldots,X_n]$ belongs to the subring $k[\sigma_1,\ldots,\sigma_n] \subset k[X_1,\ldots,X_n]$. Moreover, there is an effective procedure for expressing a given symmetric polynomial f, as a polynomial in the $\{\sigma_j\}$.
First proof: The method is to order the monomials in the polynomial f "lexicographically", and then show how to reduce the "degree" of the leading monomial of f by subtracting an explicit monomial in the σ_j. This process may introduce new monomials in the X_i into f, but only ones of lower degree. Since this process lowers the degree of the leading monomial at each step, and there are only a finite number of monomials having degree less than any given one, this process must stop in a finite number of steps. The number of steps is at most the number of monomials having degree $\le$ the degree of the original leading term of f, (which can be larger than the number of terms in the original polynomial f).

Recall that in the "lexicographical ordering" on monomials in the variables $X_1,\ldots,X_n$, we have $X_1^{i1}X_2^{i2}\ldots X_n^{in} > X_1^{j1}X_2^{j2}\ldots X_n^{jn}$ iff the first non zero integer in the sequence $(i_1-j_1), (i_2-j_2),\ldots,(i_n-j_n)$ is positive. The "leading term" of a polynomial in $X_1,\ldots,X_n$ is the term whose monomial is largest in the lexicographical ordering.
Lemma: If f is a symmetric polynomial with leading monomial $m = X_1^{i1}X_2^{i2}\ldots X_n^{in}$, then $i_1 \ge i_2 \ldots \ge i_n$
proof: Since f is symmetric, f must also contain the monomial $\sigma(m)$

for every permutation σ Thus if $\alpha < \beta$, and σ is the transposition $(\alpha\beta)$, then f contains the monomial $\sigma(m) = X_1^{i1}X_2^{i2}....X_\alpha^{i\beta}....X_\beta^{i\alpha}....X_n^{in}$. Since m is the leading term, we must have $m \geq \sigma(m)$, i.e. $i_\alpha - i_\beta \geq 0$ QED.

Exercise #109) Prove:
(i) The leading term of the product of two polynomials, in the lexicographical ordering, is the product of their leading terms.
(ii) The leading term of σ_j^r is $(X_1X_2....X_j)^r$.
(iii) The leading term of $\sigma_1^{a1}\sigma_2^{a2}....\sigma_n^{an}$ is $(X_1^{a1+a2+....+an})(X_2^{a2+a3+....+an})(X_3^{a3+....+an})(.....)(X_n^{an})$.

As a consequence of the previous lemma and exercise, if $t = cX_1^{i1}X_2^{i2}..X_n^{in}$ is the leading term of a symmetric polynomial f, then $g = f - c(\sigma_1^{i1-i2})(\sigma_2^{i2-i3})....(\sigma_{n-1}^{in-1-in})(\sigma_n^{in})$ is also symmetric and has smaller leading term than f. Thus if we repeat this process, to reduce the leading term of g, and so on, we will eventually get zero, at which point we have expressed f as a polynomial in the σ_j. **QED for the first proof of fund. thm. on symmetric functions, hence also for part (i) of Theorem.**

First proof of general cubic discriminant formula:
Remarks: Using the algorithm from the proof of the previous proposition to express $D = (X-Y)^2(X-Z)^2(Y-Z)^2$ in terms of $\sigma_1 = X+Y+Z$, and $\sigma_2 = XY+XZ+YZ$, and $\sigma_3 = XYZ$, is a bit tedious when done by hand. So let's use Mathematica to do the calculations. Care must be exercised however, since even computers can give wrong answers when the data is entered in a form they do not recognize. In particular, one needs to separate the factors being multiplied like this: (X Y + X Z + Y Z), since if two multiplied letters are too close together, the program treats them as one symbol XY and gives results like $X\,XY + X^2\,Y$, instead of $2X^2Y$. Indeed, unless sufficient spaces are used generously to separate almost everything, the calculation can yield simply the wrong answer, with no warning. When entered correctly, the whole calculation took five steps, each computation taking 4 or 5 seconds on a Macintosh IIsi with 17 megs of built in RAM.

Step one: Begin with $D = \delta^2 = (X-Y)^2(X-Z)^2(Y-Z)^2$, and notice that in the lexicographical ordering, the leading term of (X-Y) is X, the leading of (X-Z) is X, and that of (Y-Z) is Y. Since the leading term of a product is the product of the leading terms, by exercise #109) above, the leading term of D is X^4Y^2. Then the algorithm in the proof of the Prop. tells us to subtract from D the monomial $\sigma_1^2\sigma_2^2$ $= (X+Y+Z)^2(XY+XZ+YZ)^2$.

We evaluate $D - \sigma_1^2\sigma_2^2$ in Mathematica by entering the command Expand[(X-Y)^2(X-Z)^2(Y-Z)^2 - (X+Y+Z)^2(XY+XZ+YZ)^2], (but separating the symbols by more spaces), and hitting the two keys "command" and "return". The result of this calculation has 13 terms, the leading one being $-4X^4YZ$.

Thus the next calculation is $D - \sigma_1^2\sigma_2^2 + 4\sigma_1^3\sigma_3$. I.e.
Step two: In Mathematica one can re enter the result of a previous calculation by typing %, instead of retyping everything, so this time we enter Expand[% + 4 (X + Y + Z)^3 (X Y Z)], spaced like this, then hit command/return. Result: 20 terms; $-4X^3Y^3 \pm \ldots$..

Step three: Expand[% + 4 (X Y + X Z + Y Z)^3]; 17 terms; $18X^3Y^2Z \pm \ldots$...

Step four: Expand[% -18(X+Y+Z)(X Y+X Z+Y Z)(X Y Z)] (with more spaces between symbols); 1 term, $-27X^2Y^2Z^2$.

I can do this one, but Mathematica gives:
Step five: Expand[% +27(X Y Z)^2] = 0.

After combining the calculations we have
$D - \sigma_1^2\sigma_2^2 + 4\sigma_1^3\sigma_3 + 4\sigma_2^3 - 18\sigma_1\sigma_2\sigma_3 + 27\sigma_3^2 = 0.$
Hence, $D = \sigma_1^2\sigma_2^2 - 4\sigma_1^3\sigma_3 - 4\sigma_2^3 + 18\sigma_1\sigma_2\sigma_3 - 27\sigma_3^2.$

Of course we aren't quite done, since the σ's differ from the coefficients of f by minus signs I.e. if $f = X^3+pX^2+qX+r$, as in the theorem, then $\sigma_1 = -p$, $\sigma_2 = q$, $\sigma_3 = -r$, so in fact the minus signs all cancel, and we get $D = p^2q^2 - 4p^3r - 4q^3 + 18pqr - 27r^2$
QED, first proof of discr. formula for general cubic.

Second approach to fund thm. of symmetric functions:
We give a second, faster algorithm for expressing symmetric polynomials in terms of elementary ones, and use it on D, again with the help of Mathematica. This method uses induction on the number of variables, so it is most useful when n = 3, since computations for n = 2 are easy, making the inductive step simple.

Let g be a symmetric polynomial in n variables, and set $\tilde{g}(X_1,\ldots,X_{n-1}) = g(X_1,\ldots,X_{n-1},0)$. Then $\tilde{g}$ is symmetric in n-1 variables, hence by induction can be written as a polynomial $\tilde{h}(\tilde{\sigma})$ in the elementary symmetric functions $\tilde{\sigma}_j$ of n-1 variables. Note $\tilde{\sigma}_j(X_1,\ldots,X_{n-1}) = \sigma_j(X_1,\ldots,X_{n-1},0)$. Now put $G = g - \tilde{h}(\sigma)$. By construction, G is symmetric in n variables and $G(X_1,\ldots,X_{n-1},0) = 0$. Thus X_n divides G, and thus by symmetry all X_j divide G. Thus $G = \sigma_n \cdot g_1$, where the degree of every term of g_1 is three less than the degrees of the terms of g. Now repeat this procedure on g_1.
QED 2nd proof of fund thm on symm fncs.

2nd computation of D for general cubic:
Since the new algorithm lowers the degree by three each time, it should only take two steps to compute the sextic discriminant of a cubic, as opposed to the five steps of the previous algorithm. But we do have to compute the inductive steps as well.
Step one: In $D = (X-Y)^2(X-Z)^2(Y-Z)^2$, set Z=0, getting $\tilde{D} = (X-Y)^2X^2Y^2$ = (by hand) $((X+Y)^2 - 4XY)(XY)^2 = [\tilde{\sigma}_1^2 - 4\tilde{\sigma}_2](\tilde{\sigma}_2)^2 = \tilde{\sigma}_1^2\tilde{\sigma}_2^2 - 4\tilde{\sigma}_2^3$
[Note we already have two of the terms from our previous algorithm.]
Now consider $E = D - \sigma_1^2\sigma_2^2 + 4\sigma_2^3$, which we calculate on Mathematica, getting $E = (XYZ)h = \sigma_3 h$, where h has 10 terms. In fact, $h(X,Y,Z) = -4X^3+6X^2Y + 6XY^2 -4Y^3 + Z(\ldots..)$, so setting Z=0 in h, gives $\tilde{h}(X,Y) = -4X^3+6X^2Y + 6XY^2 -4Y^3$ = (by hand) $-4((X+Y)^3 - 3(XY)(X+Y)) + 6(XY)(X+Y) = -4(X+Y)^3 + 18(XY)(X+Y) = -4\tilde{\sigma}_1^3 + 18\tilde{\sigma}_1\tilde{\sigma}_2$.
Again on Mathematica, we compute $h + 4\sigma_1^3 - 18\sigma_1\sigma_2 = -27XYZ = -27\sigma_3$. Thus $h = -27\sigma_3 - 4\sigma_1^3 + 18\sigma_1\sigma_2$, and
$E = \sigma_3 h = -27\sigma_3^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3$.

Then $D = E + \sigma_1^2\sigma_2^2 - 4\sigma_2^3 =$
$-27\sigma_3^2 - 4\sigma_1^3\sigma_3 + 18\sigma_1\sigma_2\sigma_3 + \sigma_1^2\sigma_2^2 - 4\sigma_2^3 =$
$-27r^2 - 4p^3r + 18pqr + p^2q^2 - 4q^3$, as desired.
QED, second proof of discr. formula for general cubic.

Exercise #110)
(i) Express $(X-Y)^2$ as a polynomial in $\sigma_1 = X+Y$, and $\sigma_2 = XY$.
[QED for part (ii) of the discriminant theorem above.]
(ii) Express $X^2+Y^2+Z^2$ as a polynomial in $\sigma_1 = X+Y+Z$, $\sigma_2 = XY+XZ+YZ$, and $\sigma_3 = XYZ$.
(iii) Express $X^3+Y^3+Z^3$ as a polynomial in $\sigma_1 = X+Y+Z$, $\sigma_2 = XY+XZ+YZ$, and $\sigma_3 = XYZ$.
(iv) Express $X^4+Y^4+Z^4$ as a polynomial in $\sigma_1 = X+Y+Z$, $\sigma_2 = XY+XZ+YZ$, and $\sigma_3 = XYZ$.

Third derivation of gen'l cubic discrim. formula:
(Sketch, using Van der Monde's determinant).
Consider $D = \delta^2 = (\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2$, the discriminant of the general polynomial of degree 3, with roots $\alpha_1,\alpha_2,\alpha_3$. Then δ is the determinant of the 3×3 matrix A, whose jth column is $(1, \alpha_j, \alpha_j^2)$, and therefore δ is also the determinant of A^t, the transpose of A. Thus $D = \delta^2 = \det(AA^t) =$
$\pi_0\pi_2\pi_4 + 2\pi_1\pi_2\pi_3 - \pi_2^3 - \pi_0\pi_3^2 - \pi_1^2\pi_4$, where
$\pi_i = \alpha_1^i+\alpha_2^i+\alpha_3^i$. Now using exercise #99), and the fact that if $f = X^3+pX^2+qX+r$, then $p = -\sigma_1$, $q = \sigma_2$, $r = -\sigma_3$, we can get part (iv) of the discriminant theorem, by substituting and expanding.
QED, third proof of gen'l cubic discrim. formula.

Since the previous arguments depended on use of Mathematica, we give a direct argument for reduced cubics, due to Robert Varley.
Derivation of the special cubic discriminant(Varley):
If $f = X^3+pX+q$, we know by the fundamental theorem on symmetric functions that D(f) can be written as a polynomial in p, q with integer coefficients, and we want to calculate it. Let's ask ourselves what this polynomial must look like. Observe that $D = (\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2$ is homogeneous of degree six in the roots α, while $p = \sigma_2 = (\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3)$ is homogeneous of

degree two, and $q = -(\alpha_1\alpha_2\alpha_3)$ is homogeneous of degree three. It follows that the only possible monomials in p, q which can occur are p^3 and q^2, so that D must equal a polynomial of form ap^3+bq^2. Now we just need to compute a and b, which we can do by calculating D for a couple of convenient cubics. We need to choose two simple monic cubics whose roots add to zero, such as $f = (X-1)(X+1)X = X^3-X$, and $g = (X-1)^2(X+2) = X^3 -3X+2$. Since g has two equal roots, we know $D(g) = 0 = ap^3+bq^2 = a(-3)^3 + b(2)^2 = -27a+4b$. Since the roots of f are 1,-1,0, we compute $D(f) =(1)^2(1)^2(2)^2=4 = ap^3+bq^2 = a(-1)^3 = -a$. Thus $a = -4$. From $D(f) = 0 = -27a+4b$, then $b = -27$. QED.

In the next section we use a generalization of the discriminant to obtain a calculation for the general cubic that can be done by hand.

§13) The Resultant, and its formula as a determinant.

In this section we sketch another approach to calculating discriminants using the determinantal formula for a related invariant, the "resultant" of two polynomials. Recall that for a monic polynomial $f = X^n+a_1X^{n-1}+..+a_{n-1}X+a_n$, with roots $\alpha_1,....,\alpha_n$, the discriminant is defined as $D(f) = \Pi(\alpha_i-\alpha_j)^2$, product over all $i<j$ Thus $D(f) = 0$ iff for some $i\neq j$ we have $\alpha_i=\alpha_j$, iff the roots of f are not all distinct, iff f has some "multiple" roots, iff f, f' have a common root. We also know that D(f) is a "universal" polynomial in the coefficients of f Now if f,g are two monic polynomials over a field k, it is of interest to have a way to measure whether they have a root in common. If the roots of f are $\alpha_1, \ldots,\alpha_n$, and those of g are $\beta_1,\ldots,\beta_m$, one way to measure this is to introduce the "resultant" of f,g, denoted R(f,g), and defined by the difference product $\Pi(\alpha_i-\beta_j)$, for all i,j Moreover, since a polynomial has a multiple root iff it has a root in common with its derivative, there should be a connection between the discriminant of f and the resultant of f and f'. Since f, f' are not usually both monic, we need to extend the definition of resultant, and of discriminant, to non monic polynomials. We do so as follows.

Definition: If $f = a_0X^n+a_1X^{n-1}+..+a_{n-1}X+a_n$, is a polynomial with coefficients in a field k, with roots $\alpha_1,....,\alpha_n$ in some extension field, and if $g = b_0X^m+b_1X^{m-1}+....+b_{m-1}X+b_m$, with roots $\beta_1,....,\beta_m$, then

we define the discriminant of f as $D(f) = a_0^{2n-2} \Pi(\alpha_i-\alpha_j)^2$, for all $i<j$, and the resultant of f,g as $R(f,g) = a_0^m b_0^n \Pi(\alpha_i-\beta_j)$, for all i,j.

Remarks: The constants in these definitions do not affect the vanishing properties of these quantities, as long as $a_0 b_0 \neq 0$, and have the following advantages: the factor in front of D(f) makes the discriminant of aX^2+bX+c come out as b^2-4ac, instead of $(b^2-4ac)/a^2$, and the factors in front of R(f,g) make the determinant formula below come out simpler, and hence easier to remember.

The determinant formula for the resultant is based on the following:
Lemma: For f,g as in the previous definition, non constant polynomials, the following statements are equivalent:
(i) f,g have a common root in some extension field of k;
(ii) the gcd of f,g in k[X] has degree ≥ 1;
(iii) the lcm of (A,B) in k[X] has degree $\leq m+n-1$;
(iv) there are non zero polynomials A, B in k[X], with $\deg(A) \leq m-1$, $\deg(B) \leq n-1$, such that $Af = Bg$;
(v) the following set of polynomials
$(x^{m-1}f, x^{m-2}f, \ldots\ldots x^2f, xf, f, x^{n-1}g, x^{n-2}g, \ldots\ldots, x^2g, xg, g)$
is k - linearly dependent in k[X];
(vi) the coefficient vectors $(a_0, a_1, \ldots, a_n, 0, \ldots\ldots, 0)$,
$(0, a_0, a_1, \ldots, a_n, 0, \ldots\ldots, 0)$, $\ldots\ldots\ldots$, $(0, \ldots\ldots, 0, a_0, a_1, \ldots, a_n)$,
$(b_0, b_1, \ldots\ldots, b_m, 0, \ldots\ldots, 0)$, $(0, b_0, b_1, \ldots\ldots, b_m, 0, \ldots\ldots, 0)$, $\ldots\ldots\ldots\ldots$,
$(0, 0, \ldots\ldots\ldots, 0\ b_0, b_1, \ldots\ldots, b_m)$, of the polynomials in (v) are linearly dependent in k^{n+m};
(vii) the determinant of the matrix whose rows are the coefficient vectors in (vi) is zero.
proof: This is a straightforward exercise. QED.

This leads to the following "Sylvester's" formula for the resultant:
Prop: For f, g as in the previous definition, the resultant R(f,g) equals the determinant of the following $(m+)\times(m+n)$ matrix, where there are m rows of α's and n rows of β's, and where spaces not containing a's or b's are filled in by zeroes:

$a_0\ a_1 \ldots\ldots\ldots\ldots\ldots\ldots a_n$
$0\quad a_0\ a_1 \ldots\ldots\ldots\ldots\ldots\ldots a_n$

$$
\begin{vmatrix}
. \ldots\ldots\ldots\ldots a_0 & a_1 & \ldots\ldots\ldots & a_n & & = R(f,g) \\
b_0 & b_1 & \ldots\ldots\ldots & . & b_m & \\
0 & b_0 & b_1 & \ldots\ldots\ldots\ldots & b_m & \\
\\
. . & \ldots\ldots\ldots\ldots & b_0 & b_1 & \ldots\ldots\ldots\ldots & b_m
\end{vmatrix}
$$

Proof: Sketch: Consider the case of the general polynomials f,g where the α's and β's are variables. Let S in $\mathbb{Z}[\ldots,a_i,\ldots,b_j,.]$ stand for the determinant above. If we factor out a_0 from the first m rows, and b_0 from the last n rows of S, we see $S = a_0^m b_0^n \cdot$ (polynomial in the a_i/a_0's and b_j/b_0's), and since the a_i/a_0's and b_j/b_0's, up to sign, are elementary symmetric functions of the α's and β's, S can be expanded as $a_0^m b_0^n \cdot$(polynomial in the α's and β's). Then since S vanishes when f,g have a common root, S is a multiple of $a_0^m b_0^n \prod_{i,j} (\alpha_i - \beta_j) = R(f,g)$, in $\mathbb{Q}[a,b]$. Moreover, S is homogeneous of degree m in the a's and of degree n in the b's, and when expanded in terms of a's and b's, it contains the (product of the main diagonal) term $a_0^m b_0^n$. We will prove the same properties for the resultant $R = R(f,g) = a_0^m b_0^n \prod_{i,j}(\alpha_i - \beta_j)$ Since R is symmetric in both the α's and the β's (separately), R can at least be written as some polynomial in the a's and b's. (I.e. $\prod_{i,j}(\alpha_i - \beta_j)$ can be written as a polynomial in the a_i/a_0, and b_j/b_0, and then the factors a_0^m, b_0^n allow clearing denominators.)

Since $g(X) = b_0 \prod_j (X - \beta_j)$, one gets $\prod_i g(\alpha_i) = b_0^n \prod_{i,j} (\alpha_i - \beta_j)$. Hence $R(f,g) = a_0^m b_0^n \prod_{i,j}(\alpha_i - \beta_j) = a_0^m \prod_i g(\alpha_i)$. Since g is homogeneous linear in its coefficients, the b's, this shows that R is homogeneous of degree n in the b's. Similarly, $f(X) = a_0 \prod_i (X - \alpha_i)$, so $\prod_j f(\beta_j) = a_0^m \prod_{i,j} (\beta_j - \alpha_i) = (-1)^{nm} a_0^m \prod_{i,j} (\alpha_i - \beta_j)$. Thus $R = (-1)^{nm} b_0^n \prod_j f(\beta_j)$, so R is homogeneous of degree m in the a's. Now that we have shown both R,S are both polynomials in the a's and b's, and homogeneous of the same degrees, and that R divides S in $\mathbb{Q}[a,b]$, it follows that S is a rational multiple of R.

Now we know S contains the term $a_0^m b_m^n$ (with coefficient 1). Hence it suffices to compute the coefficient of this monomial in R. Since we saw above that $R = a_0^m \prod_i g(\alpha_i) = a_0^m \prod_i (b_0 \alpha_i^m + \ldots + b_m)$, we see that the product of the constant terms of the factors in $\prod$ is

$b_m{}^n$, hence this expansion contains exactly the term $a_0{}^m b_m{}^n$. It follows that R(f,g) equals the determinant S. QED

Finally one can prove, from the equation $f = a_0\Pi(X-\alpha_i)$, and using the product rule for the derivative, the following connection between the resultant and the discriminant:
Prop: With f as in the previous definition, we have
$R(f,f') = a_0{}^{2n-1}\Pi(\alpha_i-\alpha_j)$, for $i \neq j$,
$= (-1)^{n(n-1)/2}\, a_0{}^{2n-1}\, \Pi(\alpha_i - \alpha_j)^2$, for $i < j$,
$= (-1)^{n(n-1)/2}\, a_0\, D(f)$.

Corollary: If f is monic, i.e. $a_0 = 1$, then $R(f,f') = (-1)^{n(n-1)/2}\, D(f)$.

Corollary: (i) If $f = X^3+pX+q$, then D(f) is *minus* the determinant of the 5×5 matrix:

$$\begin{vmatrix} 1 & 0 & p & q & 0 \\ 0 & 1 & 0 & p & q \\ 3 & 0 & p & 0 & 0 \\ 0 & 3 & 0 & p & 0 \\ 0 & 0 & 3 & 0 & p \end{vmatrix} = 4p^3 + 27q^2.$$

(ii) If $f = X^3+pX^2+qX+r$, then D(f) is *minus* the determinant of the 5×5 matrix:

$$\begin{vmatrix} 1 & p & q & r & 0 \\ 0 & 1 & p & q & r \\ 3 & 2p & q & 0 & 0 \\ 0 & 3 & 2p & q & 0 \\ 0 & 0 & 3 & 2p & q \end{vmatrix} = -p^2q^2 + 4q^3 + 4p^3r - 18pqr + 27r^2.$$

Remarks: In some geometry books the discriminant is defined as R(f,f'). This is advantageous for projective geometry since D(f) then vanishes also when f actually has degree lower than n, which reveals that the homogeneous polynomial of degree n, of which f is the affine version, has a root "at infinity" in common with f'. Eg. if n=2, this makes the discriminant of $aX^2 + bX + c$ equal $-a(b^2-4ac)$, which vanishes also when a=0. On the other hand, this is disadvantageous for Galois theorists, since we like the fact that our old discriminant has a square root in the coefficient field precisely

when the Galois group is contained in the subgroup $A_n \subset S_n$.

Exercise #111) (i) Assume f is a separable polynomial over a field k of characteristic $\neq 2$, with distinct roots $\alpha_1,\ldots,\alpha_n$ in some splitting field, and discriminant $D = \Pi(\alpha_i-\alpha_j)^2$. Prove D has a square root in k iff the Galois group of f is contained in A_n.
(ii) Why would part (i) fail if we defined D(f) as $R(f,f')$?
(iii) If f is an irreducible cubic over a field of characteristic $\neq 2$ and f is separable, prove $G(f) \cong \mathbb{Z}_3$ iff D(f) has a square root in k, and $G(f) \cong S_3$ iff not.
(iv) Compute G(f) for $f = X^3 - X + 1$ and $k = \mathbb{Z}_3$.
(v) Compute G(f) for $f = X^3 - X + 1$ and $k = \mathbb{Z}_5$.
(vi) Compute G(f) for $X^3 - 3X - 1$, over $\mathbb{Q}$.
(vii) If $f = X^3 - \alpha X^2 + (\alpha-3)X + 1$, where α is in $\mathbb{Q}$, prove that f is irreducible over $\mathbb{Q}$, and $D(f) = (\alpha^2-3\alpha+9)^2$; hence that $G(f) \cong A_3$.

§14) "Cardano's" formula for the roots of a cubic
The solution of the general cubic equation was apparently found first by Scipione del Ferro around 1515, rediscovered by Niccolo Fontana (or "Tartaglia" = the stammerer) about 1535, and revealed by him to Girolamo Cardano, who published it, with due credit to both del Ferro and Fontana, although evidently not with Fontana's permission, in his book "Ars Magna" in 1545. Since priority is usually established most firmly by publication, these formulas have become known by Cardano's name. We will see how a solution of the cubic can be derived fairly easily using the methods of Galois theory, and then we will compare the solution to that of Fontana/Cardano.
First consider the general monic cubic equation $Y^3+aY^2+bY+c = 0$. We know the sum of the roots Y equals $s_1 = -a$. Since there are three roots, if we were to subtract $(s_1)/3$ from each root, the new roots would have sum zero. Hence if we denote the new roots by X, where $X = Y-(s_1)/3 = Y + (a/3)$, and substitute $Y = X - a/3$, the new equation $(X - a/3)^3 + a(X - a/3)^2 + b(X - a/3) + c = 0$, becomes $X^3 + pX + q = 0$. If we can solve this simpler equation for X, then the solution of the original equation is $Y = X - a/3$. Thus we may consider only the equation $f(X) = X^3 + pX + q = 0$.

Recall the **Galois - theoretic point of view:**
(i) Exhibit the splitting field of f as a (good) radical extension of the coefficient field, obtained in stages by adjoining pth roots, for all the primes p occurring in the composition series for the Galois group G.
(ii) Then use LaGrange resolvents formed from the generators of the corresponding quotient groups of G to actually compute candidates for these pth roots β.
(iii) Express all these pth roots in terms of the coefficients of f, using the symmetries of G.
(iv) We have already observed that the solutions α of the polynomial can be obtained by adding the resolvents β.

So let $k = \mathbb{Q}(\omega,p,q)$ be the field generated over $\mathbb{Q}$ by the coefficients p,q of f, and a primitive cube root ω of unity, $\omega^2 + \omega + 1 = 0$. If $\alpha_1,\alpha_2,\alpha_3$ are the roots of f, the splitting field is $L = k(\alpha_1,\alpha_2,\alpha_3)$. If f is the general polynomial of its form, i.e. if p,q, are variables, then $D(f) = -4p^3-27q^2$, is not the square of an element of k (since p^3 is not a square mod (q)). Thus by ex. #99, the Galois group of f is $S_3 \cong Bij(\{\alpha_1,\alpha_2,\alpha_3\})$. The composition series for G is $\{e\} \subset A_3 \subset S_3$, and thus there are two stages to building up the splitting field L, $k \subset F \subset L$. Here $F = k(\delta)$ is the quadratic extension generated by $\delta = D^{1/2}$, and $L = F(\beta)$, where $\beta = \alpha + \omega\cdot\tau(\alpha) + \omega^2\cdot\tau^2(\alpha)$, is a LaGrange resolvent, τ is a generator of $A_3 \cong$ the Galois group of $F \subset L$, and α is a suitable element of L. We choose $\alpha = \alpha_1$, one of the roots, since we will eventually want to solve for the roots in terms of β, and choose $\tau = (\alpha_1\alpha_2\alpha_3)$. Our third, explicit proof of Galois' theorem showed that at least one of the two following LaGrange resolvents will serve as β either $\beta_1 = \alpha_1 + \omega\cdot\alpha_2 + \omega^2\cdot\alpha_3$, or $\beta_2 = \alpha_1 + \omega^2\cdot\alpha_2 + \omega\cdot\alpha_3$. (Since we will see just below that β_1 and β_2 are Galois conjugates, they are both non zero, hence either would serve as β.) We also saw how to solve for α_1 in terms of these β's, since if we define $\beta_0 = \alpha_1 + \alpha_2 + \alpha_3$, then $1 + \omega + \omega^2 = 0 = \beta_0$. Thus we get $\beta_1 + \beta_2 = 3\alpha_1$, so that $\alpha_1 = (1/3)\beta_1 + (1/3)\beta_2$. Since we have expressed α_1 in terms of the β's, we would be done if we could express the β's in terms of the coefficients p, q.
Now it followed from the construction of the LaGrange resolvents that both β_1^3 and β_2^3 belong to F, (because they are fixed by τ), and thus are quadratic over k. Therefore, in order to express them

in terms of p,q, it suffices to find the quadratic equation over k they satisfy. This equation can be found by Galois theory. At the beginning of our study of fields, we noticed that the roots of a quadratic equation over $\mathbb{R}$ are complex conjugates of each other, and the coefficients of the equation are (up to sign) just the sum and the product of the roots. In the present case, we recover the coefficients of the quadratic equation for β_1^3 by adding and multiplying β_1^3 with its Galois conjugate in F. Since the Galois group of F/k is just $S_3/A_3 \cong \mathbb{Z}_2$, and is generated by the coset of any odd permutation, we may represent a generator by the transposition $\nu = (\alpha_2\alpha_3)$. Then the Galois conjugate of β_1 is $\nu(\beta_1) = \beta_2$. Thus the equation over k satisfied by β_1^3 is just $T^2 - (\beta_1^3+\beta_2^3)T + (\beta_1\beta_2)^3$.

Let's calculate the coefficients of that quadratic equation. We know if we express them in terms of the α's that they must be symmetric, and hence there is an algorithm for writing them in terms of p, q but we hope it will be easy to do by hand.
First we get $\beta_1\beta_2 = (\alpha_1 + \omega\cdot\alpha_2 + \omega^2\cdot\alpha_3)(\alpha_1 + \omega^2\cdot\alpha_2 + \omega\cdot\alpha_3) = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + (\omega^2+\omega)(\alpha_1\alpha_2) + (\omega^2+\omega)(\alpha_1\alpha_3) + (\omega^2+\omega)(\alpha_2\alpha_3)$, and if we note that $\omega^2+\omega+1=0$ implies $(\omega^2+\omega) = -1$, this becomes $\beta_1\beta_2 = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 - (\alpha_1\alpha_2) - (\alpha_1\alpha_3) - (\alpha_2\alpha_3) = \sigma_1^2 - 2\sigma_2 - \sigma_2 = \sigma_1^2 - 3\sigma_2$. Since $\sigma_1 = \alpha_1 + \alpha_2 + \alpha_3 = 0$, and $\sigma_2 = p$, this is just $\beta_1\beta_2 = -3p$. Hence $\beta_1^3\beta_2^3 = -27p^3$.

Now for $\beta_1^3 + \beta_2^3$, we simplify a bit first: $\beta_1^3 + \beta_2^3 = (\beta_1+\beta_2)^3 - 3\beta_1\beta_2(\beta_1+\beta_2)$. Then we calculate $\beta_1+\beta_2 = 2\alpha_1+(\omega^2+\omega)\alpha_2+(\omega^2+\omega)\alpha_3 = 2\alpha_1-\alpha_2-\alpha_3 = 3\alpha_1$, where at the last step we added $0 = \alpha_1+\alpha_2+\alpha_3$ to the expression.
Substituting $3\alpha_1$ for $\beta_1+\beta_2$, and $-3p$ for $\beta_1\beta_2$ in the expression found above gives $\beta_1^3 + \beta_2^3 = (3\alpha_1)^3 - 3(-3p)(3\alpha_1) = 27\alpha_1(\alpha_1^2 + p)$.
Since $p = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3$, we get
$(\alpha_1^2 + p) = (\alpha_1^2 + \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3) = (\alpha_1[\alpha_1 + \alpha_2 + \alpha_3] + \alpha_2\alpha_3) = \alpha_2\alpha_3$ Thus
$\beta_1^3 + \beta_2^3 = 27\alpha_1\alpha_2\alpha_3 = 27\sigma_3(\alpha) = -27q$.
The quadratic equation satisfied by β_1^3 and β_2^3 is:
$T^2 + (27q)\,T - 27p^3$.

Then the quadratic formula yields:
$\beta_1^3 = (1/2)\{-27q + [(27q)^2 + 4(27p^3)]^{1/2}\}$
$\beta_2^3 = (1/2)\{-27q - [(27q)^2 + 4(27p^3)]^{1/2}\}$.

Thus β_1 is one of the three cube roots of the first expression and $\beta_2 = -3p/\beta_1$. Finally we have $\alpha_1 = (1/3)\beta_1 + (1/3)\beta_2$. The three choices of a cube root of the expression above yield the three roots $\alpha_1,\alpha_2,\alpha_3$. If we simplfy the expression a little we notice that $[(27q)^2 + 4(27p^3)]^{1/2} = (-27D)^{1/2}$ so that
$\beta_1^3, \beta_2^3 = (-27q)/2 \pm (3/2)(-3D)^{1/2}$, and the three roots of f are

$\alpha_i =$
$(1/3)\{[(-27q)/2 + (3/2)(-3D)^{1/2}]^{1/3} + [(-27q)/2 - (3/2)(-3D)^{1/2}]^{1/3}\}$,
where the first cube root is chosen arbitrarily from the three choices available, and the second cube root is chosen so that the product of the two cube roots is $-3p$.
We can simplify the formula a bit more, by putting the factor of 1/3 under the cube root sign, and the other fractions under the square root sign, to get the following version of:

<u>The cubic formula of del Ferro, Fontana, Cardano:</u>
The solutions of $X^3 + pX + q = 0$, are given by:

$$\alpha_i = (-q/2 + [-D/108]^{1/2})^{1/3} + (-q/2 - [-D/108]^{1/2})^{1/3},$$

where $D = -4p^3-27q^2$, and the product of the two cube roots is $-p/3$. Choosing the complex cube root of the first term in all three possible ways yields the three solutions $\alpha_1,\alpha_2,\alpha_3$.

We give two interesting examples of using this equation:
(i) $X^3 - 12X + 16 = 0$ Here the solution $X = 2$ would seem to be the simplest, but since $D = 0$, the formula yields $(-8)^{1/3} + (-8)^{1/3} = -4$ Since the polynomial factors as $(X-2)(X-2)(X+4)$, we see that -4 is a solution. However the other two solutions, both $X = 2$, can apparently be obtained only by substituting complex cube roots of -8 into the solution formula. Phenomena like this were somewhat disturbing to early workers who had not yet fully accepted even negative numbers, much less complex numbers. Indeed, one can

show that a cubic with three real roots, irreducible over $\mathbb{Q}$, *cannot* be solved entirely by *real* radicals.

(ii) $X^3+X-2 = 0$. Here $X = 1$ is a real solution, and one can show there is only one real solution of this equation. Therefore the real answer given by Cardano's formulas, namely
$X = (1 + (2/3)\sqrt{(7/3)})^{1/3} + (1 - (2/3)\sqrt{(7/3)})^{1/3}$, must equal 1!!.

Exercise #112)
(i) Show that Cardano's formulas do yield all three solutions of $X^3 - 12X + 16 = 0$.
(ii) Show that Cardano's formula is correct for X^3+X-2, i.e. show $(1 + (2/3)\sqrt{(7/3)})^{1/3} + (1 - (2/3)\sqrt{(7/3)})^{1/3} = 1$.

Finally, if we compare with the original solution as given to Cardano by Fontana, in a short poem, Fontana's prescription for solving $X^3 + pX = r$, was simply to find two numbers a,b such that $a-b = r$, and $27ab = p^3$. Then the desired solution is $X = a^{1/3} - b^{1/3}$. This is the same as our solution if we put $-r = q$, $27a = \beta_1^3$, and $-27b = \beta_2^3$. I e then $a-b = r$ iff $\beta_1^3 + \beta_2^3 = -27q$, and $27ab = p^3$ iff $\beta_1^3\beta_2^3 = -27p^3$, exactly the conditions we found above which if satisfied by β_1^3 and β_2^3 implied that our cubic has solution $\alpha = (1/3)(\beta_1 + \beta_2)$. So Fontana's prescription was a description of the roots of the "resolvent quadratic" associated to the cubic, and how to use those roots to express roots of the cubic.

It is usual to describe Fontana's solution of the cubic via a prescription for transforming the cubic equation into the corresponding quadratic equation as follows: To solve $X^3+pX+q = 0$, put $X = u+v$, obtaining $0 = (u+v)^3 + p(u+v) + q$
$= u^3 + v^3 + q + (3uv + p)(u+v)$. Observe that this equals zero if simultaneously $(3uv + p) = 0 = u^3 + v^3 + q$. To solve this pair of equations we substitute the formula $v = (-p/3u)$ obtained from solving the first equation, into the second equation $0 = u^3 + v^3 + q$, getting $u^3 + (-p/3u)^3 + q = 0$. Multiplying by u^3 and rearranging gives $u^6 + qu^3 - p^3/27 = 0$, which is quadratic in u^3. Note: If we set $u = \beta/3$, this quadratic equation becomes $(\beta^3)^2 + (27q)\beta^3 - 27p^3$,

exactly the quadratic equation derived above for our β_1^3, β_2^3. Hence $u = (1/3)\beta_1$ and $v = (1/3)\beta_2$, and the roots of the cubic are $X = u+v = \alpha = (1/3)\beta_1 + v = (1/3)\beta_2$, as expected. The amazing thing about this last prescription for solving a cubic is that to use it we only need one insight: namely, the solution is the sum of two other auxiliary quantities. Apparently del Ferro discovered this was the right way to proceed, perhaps sensing that it might be easier to solve two equations in two unknowns, than to solve the original equation in one unknown. In this discovery, that a general cubic could always be solved by adding the cube roots of the two solutions of an auxiliary quadratic equation, we can see the germ of the decomposition theory of Galois for normal field extensions. It just took 300 years to realize it fully.

§15) A few remarks on solving quartic polynomials

The composition series $\{e\} \subset A_3 \subset S_3$ for the group of the general cubic f, with quotients $A_3 \cong \mathbf{Z}_3$, $S_3/A_3 \cong \mathbf{Z}_2$, mirrored a decomposition of the splitting field L of f into the tower $k \subset F \subset L$, where k is the coefficient field, and F is the fixed field of A_3. This led us to attempt the solution of the cubic in two stages, first we constructed some A_3 - invariant combinations of the roots of f (our β_1^3 and β_2^3 above), which generated F. Then since we knew the Galois group $S_3/A_3 \cong \mathbf{Z}_2$ of F over k, we could find the quadratic equation over k satisfied by these generators of F. Using the quadratic formula, we could express the generators of F in terms of the coefficients of f. Finally we knew how to express the roots of f in terms of the generators of F.

To solve a quartic f in an analogous way, in as few steps as possible, the normal series $\{e\} \subset V \subset S_4$, where $V \cong \mathbf{Z}_2 \times \mathbf{Z}_2$ and $S_4/V \cong S_3$, suggests considering the tower of fields $k \subset F \subset L$, where k is the coefficient field of f, L is the splitting field, and F is the fixed field of V. We can then look for some V - invariant combinations Θ_i of the roots of f, which generate F. Since the Galois group of F is $S_4/V \cong S_3$, we expect to find an irreducible cubic over k satisfied by the Θ_i. Using the Cardano formulas just derived above, we can then write the Θ_i in terms of the coefficients of f, and finally try to express the roots of f in terms of the Θ_i. We refer the reader to the excellent

discussion in Van der Waerden, which we summarize below.

Denote the roots of $f = X^4 + pX^2 + qX + r = 0$ by $\alpha_1,\alpha_2,\alpha_3,\alpha_4$, the "coefficient field" by $k = \mathbb{Q}(i,p,q)$, the splitting field by L, and the Galois group $G = G(L/k) \cong S_4 \cong Bij(\{\alpha_i\})$. Then the normal subgroup $S_4 \supset V = \{e, (12)(34), (13)(24), (14)(23)\}$, and the following three elements of L are V- (but not S_4) invariant:
$\Theta_1 = (\alpha_1+\alpha_2)(\alpha_3+\alpha_4)$, $\Theta_2 = (\alpha_1+\alpha_3)(\alpha_2+\alpha_4)$, $\Theta_3 = (\alpha_1+\alpha_4)(\alpha_2+\alpha_3)$.
Consequently the Θ_i belong to F, the fixed field of V. Since no other element of G fixes all three of the Θ_i, they cannot belong to the fixed field of any larger group, hence they generate F. The Galois group of F/k is $G/V \cong S_3$. Since $\{\Theta_1,\Theta_2,\Theta_3\}$ form a G/V -orbit, they satisfy the following irreducible monic cubic over k: $g(U) = \prod(U-\Theta_i)$. We call g the "resolvent cubic" of f. Thus $F = k(\Theta_1,\Theta_2\Theta_3)$ is the splitting field of g. The coefficients of g as usual are ± the elementary symmetric functions applied to the Θ_j, i.e. $g(T) = U^3 - s_1(\Theta)U^2 + s_2(\Theta)U - s_3(\Theta)$.

We can express the s_j in terms of the α_j and then the coefficients of f, with the following results:
$s_1(\Theta) = 2\sigma_2(\alpha) = 2p$,
$s_2(\Theta) = \sigma_2{}^2(\alpha) + \sigma_1(\alpha)\sigma_3(\alpha) - 4\sigma_4(\alpha) = p^2-4r$,
$s_3(\Theta) = \sigma_1(\alpha)\sigma_2(\alpha)\sigma_3(\alpha) - \sigma_1{}^2(\alpha)\sigma_4(\alpha) - \sigma_3{}^2(\alpha) = -q^2$.

Thus the resolvent cubic is $U^3 - 2pU^2 + (p^2-4r)U + q^2 = 0$.
One can then use Cardano/Fontana's procedure to solve this for the Θ_j. We will see how to recover the roots α_i from the Θ_j. Since $\alpha_1+\alpha_2+\alpha_3+\alpha_4=0$, it follows that $\Theta_1 = (\alpha_1+\alpha_2)(\alpha_3+\alpha_4) = -(\alpha_1+\alpha_2)^2$, and thus $(\alpha_1+\alpha_2) = (-\Theta_1)^{1/2}$. Similarly, $(\alpha_3+\alpha_4) = -(-\Theta_1)^{1/2}$, $(\alpha_1+\alpha_3) = (-\Theta_2)^{1/2}$, $(\alpha_2+\alpha_4) = -(-\Theta_2)^{1/2}$, $(\alpha_1+\alpha_4) = (-\Theta_3)^{1/2}$, $(\alpha_2+\alpha_3) = -(-\Theta_3)^{1/2}$. Thus we have the following formulas:

$2\alpha_1 = (-\Theta_1)^{1/2} + (-\Theta_2)^{1/2} + (-\Theta_3)^{1/2}$,
$2\alpha_2 = (-\Theta_1)^{1/2} - (-\Theta_2)^{1/2} - (-\Theta_3)^{1/2}$,
$2\alpha_3 = -(-\Theta_1)^{1/2} + (-\Theta_2)^{1/2} - (-\Theta_3)^{1/2}$,
$2\alpha_4 = -(-\Theta_1)^{1/2} - (-\Theta_2)^{1/2} + (-\Theta_3)^{1/2}$
If one wants to compute the discriminant of the quartic, one can check that the squared difference of any two roots of the associated

cubic equals the product of the squared differences of two pairs of roots of the quartic. Thus the quartic and its resolvent cubic have the same discriminant, and we can compute the discriminant of the (special) quartic from the formula for the (general) cubic

Galois groups of quartic polynomials.
Since we can reduce the Galois field theory of the general quartic polynomial to that of a cubic polynomial, we can do the same for particular quartic polynomials. I.e. if $V \subset S_4$ is the normal subgroup of order four mentioned above, we can reduce the computation of the Galois group of a quartic f with group $G \subset S_4$, to that of its resolvent cubic, by studying the decomposition $k \subset F \subset L$ of the splitting field L, where F is the subfield corresponding to $G \cap V$.

I.e. suppose f is a separable irreducible quartic over a field k, with roots $\alpha_1,\alpha_2,\alpha_3,\alpha_4$, splitting field $k \subset L$, and Galois group $G \subset S_4$. Then $(V \cap G) = H \subset G$ is a normal subgroup, with fixed field $F \subset L$, and $k \subset F$ is a Galois extension with group $G/H \subset S_4/V \cong S_3$. Then the degree [F:k] divides $\#(S_3) = 6$, so is either 1,2,3, or 6. Just knowing the degree of the subfield $k \subset F$ determines the Galois group also of the splitting field $k \subset L$ almost completely, as follows (from Hungerford):

Theorem: If V is the normal subgroup of order 4 in S_4, f is an irreducible separable quartic over the field k, with splitting field $k \subset L$, and Galois group $G \subset S_4$, and if $k \subset F \subset L$ is the intermediate field corresponding to $G \cap V$, then
(i) $[F : k] = 6$ iff $G \cong S_4$,
(ii) $[F : k] = 3$ iff $G \cong A_4$,
(iii) $[F : k] = 1$ iff $G \cong V$,
(iv) $[F : k] = 2$, and f is irreducible over F, iff $G \cong D_4$,
(v) $[F : k] = 2$ and f is reducible over F, iff $G \cong \mathbf{Z}_4$.
proof: We know G is a transitive subgroup of S_4, and that $\#(G)$ is a multiple of 4. One can check that the only such subgroups are S_4, A_4, V, D_4 (there are three of these), and $\mathbf{Z}_4$ (and three of these). V is a subgroup of all of these except $\mathbf{Z}_4$, and it intersects $\mathbf{Z}_4$ in two elements. We know by Artin's lemma that in all cases L is Galois over $(G \cap V)$, of degree $\#(G \cap V)$, which equals 2 when $G \cong \mathbf{Z}_4$, and is otherwise 4. Thus $[L:k] = 2[F:k]$ if $G \cong \mathbf{Z}_4$, and $[L:k] = 4[F:k]$ otherwise. Thus $[F:k] = 6$ if $G \cong S_4$, $[F:k] = 3$ if $G \cong A_4$, $[F:k] = 1$ if $G \cong V$, and

$[F:k] = 2$ if either $G \cong D_4$ or if $G \cong \mathbf{Z}_4$. Thus reading backwards we have proved (i), (ii), and (iii).
Moreover, if $[F:k] = 2$ and f remains irreducible over F, then $[L:F] \geq 4$, so $[L:k] \geq 8$, and we must have $G \cong D_4$. Conversely, if $G \cong D_4$, then $V \subset G$ so V is the Galois group of $F \subset L$. Since we know the elements of $V = \{e, (12)(34), (13)(24), (14)(23)\}$, we see that the V - orbit of α_1 say, is $\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$. Thus the irreducible polynomial of α_1 over F has degree four, i.e. f remains irreducible over F. QED.

Just for convenience, we introduce another candidate for "the" resolvent cubic of a quartic. Since the resolvent cubic is the irreducible monic equation satisfied over k by three convenient generators of F which form a G-orbit, there may be several of them. If a general quartic f has roots $\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$, then F can be generated over the coefficient field by either the elements Θ_i used above, or equally well, by the elements
$\Omega_1 = \alpha_1\alpha_2 + \alpha_3\alpha_4$, $\Omega_2 = \alpha_1\alpha_3 + \alpha_2\alpha_4$, $\Omega_3 = \alpha_1\alpha_4 + \alpha_2\alpha_3$.

FACTS: (i) If f is the general quartic $f = X^4+aX^3+bX^2+cX+d$, the new resolvent cubic becomes $\Pi(T-\Omega_i) =$ $g(T) = T^3 - bT^2 + (ac-4d)T + (4bd - a^2d - c^2)$.

(ii) If f is the special quartic $f = X^4 + aX^2 + bX + c$, then the new resolvent cubic is $g(T) = T^3 - aX^2 - 4cX - (b^2-4ac)$, or $T^3 - aX^2 - 4cX + 4ac-b^2$, as you prefer. One of these last two may be easier to remember in time of need, than the Θ version above.

Remarks: (i) Essentially the same arguments as for the general quartic, show that for any particular quartic $f = X^4 + aX^2 + bX + c$, over k, with group $G \subset S_4$ and splitting field L, if $H = V \cap G$ then F = fixed field of $H = k(\Omega_1,\Omega_2,\Omega_3)$ = splitting field of the resolvent cubic $g(T)$, where $g(T) = T^3 - aX^2 - 4cX - (b^2-4ac)$.
(ii) According to the previous theorem, we can compute the Galois group G of a quartic f by first computing the group $G \cap V$ of the resolvent cubic. This is easy using the various formulas we have derived. I.e. first write down the resolvent cubic g. If g splits completely in k, $G \cap V \cong \{0\}$, if g has an irreducible quadratic factor, $G \cap V \cong \mathbf{Z}_2$. If g is irreducible, use the cubic discriminant formula to

determine whether $G \cap V \cong S_3$ or A_3. Since $[F:k] = \#(G \cap V)$, the theorem gives G.

Examples: (i) If $f = X^4 + X^3 + X^2 + X + 1$, we know by the trick of putting $X = (Y+1)$ and using Eisenstein, that f is irreducible over $\mathbb{Q}$. By the formulas above, the resolvent cubic is $T^3 - T^2 - 3T + 2$, which factors as $T^3 - T^2 - 3T + 2 = (T-2)(T^2+T-1)$. The rational root theorem tells us the only possible rational roots of the quadratic factor are ± 1, neither of which works. Thus the degree of the splitting field of the cubic over $\mathbb{Q}$ is 2, and the group of f is $G_{\mathbb{Q}}(f) \cong \mathbb{Z}_4$. (We will see below that if $f = X^{p-1} + X^{p-2} + \ldots + X + 1$, then $G_{\mathbb{Q}}(f) \cong \mathbb{Z}_{p-1}$, if p is prime.)
(ii) If $f = X^4 + 1$, the resolvent $T^3 - 4T = T(T-2)(T+2)$ splits in $\mathbb{Q}$. Thus $G_{\mathbb{Q}}(f) \cong V \cong \mathbb{Z}_2 + \mathbb{Z}_2$. (Note: the splitting field of f is the same as the splitting field of X^8-1, and that $\mathbb{Z}_2 + \mathbb{Z}_2 \cong \mathbb{Z}_8^*$. We will see below that the group of X^n-1 is always $\cong \mathbb{Z}_n^*$. Compare with ex. (i) where $n = 5$.) It is also easy to compute this group directly
(iii) If $f = X^4 - 2$, f is irreducible by Eisenstein, and the resolvent cubic is $T^3+8T = T(T^2+8)$. The group of this cubic is $\mathbb{Z}_2$, since T^2+8 has no rational roots, so $G_{\mathbb{Q}}(f)$ is either D_4 or $\mathbb{Z}_4$. One can check that f remains irreducible over $\mathbb{Q}([-8]^{1/2}) = \mathbb{Q}(i \cdot 2^{1/2})$, (since $2^{1/2}$ is not in $\mathbb{Q}(i \cdot 2^{1/2})$), so $G_{\mathbb{Q}}(f) \cong D_4$, as we computed directly before.
(iv) If $f = X^4 + X^2 + X + 1$, (no cubic term), the resolvent cubic is $T^3 - T^2 - 4T + 3$, which is irreducible (± 1, ± 3 aren't roots), with discriminant 257, which is not a square in $\mathbb{Q}$. Thus, the resolvent has group S_3, hence $G_{\mathbb{Q}}(f) \cong S_4$.
(v) If $F = 4X^4 + 4X + 3$, the monic version is $f = X^4 + X + 3/4$, and the resolvent is thus $T^3 - 3T - 1$. The discriminant of this irreducible cubic is 81, a square in $\mathbb{Q}$, so the group of the resolvent is $\mathbb{Z}_3$, and the group of f is $G_{\mathbb{Q}}(f) \cong A_4$.
(vi) If $f = X^4 + X^2 + X + 1$, as in (iv), but the base field is $k = \mathbb{Z}_5$, then f is still irreducible. Over finite fields we know all Galois groups are cyclic, so the only possibility is $G_k(f) \cong \mathbb{Z}_4$.
(vii) If k is a finite field and f is any irreducible polynomial over k of degree n, then we claim $G(f) \cong \mathbb{Z}_n$. Since we already know a finite extension of a finite field has cyclic Galois group, we just need to show the splitting field of f has degree n over k, i.e. that the field L

obtained by adjoining one root of f yields the splitting field of f. However we also know a finite extension of a finite field is always normal, hence f (which has one root in L) splits in L. QED.

Exercise #113) problem #10, Hungerford, p. 277.
Exercise #114) problem #6, Hungerford, p.301.
Exercise #115) problem #8, Hungerford, p.301.
Exercise #116) problem #9, Hungerford, p.301.
Exercise #117) Let k be a field of characteristic $p > 0$.
(i) Prove that if a is in k, and if $r = p^t$ with $t \geq 1$, then $X^r - a = (X-\alpha)^r$, for some α in an extension of k. [Hint: Then the map $\varphi: L \to L$ defined by $\varphi(\alpha) = \alpha^r$ is injective, so $L \supset k$ contains at most one root of $X^r - a$.]
(ii) If f is an irreducible polynomial over k, and if f has a multiple root in some splitting field, show that the only terms in f which can have non zero coefficients are those of form X^s where p divides s. Deduce that $f(X) = g(X^r)$ where $r = p^t$, for some $t \geq 1$, where g(X) is irreducible and has no repeated roots.
(iii) If f is irreducible over k, prove that all roots of f have the same multiplicity.

§16) The Galois groups $G_{\mathbb{Q}}(X^n-1)$ of "cyclotomic fields".

It is of fundamental interest for number theorists to know as much as possible about finite extensions of the most basic of fields, the rational numbers. How complicated can a finite extension of $\mathbb{Q}$ be? One easier question would seem to be: What are the possible Galois groups of extensions of $\mathbb{Q}$? We have already seen (in homework) that if we allow the base field to be *arbitrary*, then any finite group can occur as a Galois group. If the base field is restricted to be $\mathbb{Q}$, then it is an open question to determine exactly which Galois groups can occur. Our last theorem this quarter will be to show that at least all finite abelian groups occur as Galois groups over $\mathbb{Q}$.

Our approach to the proof of the theorem will be to find a big family of abelian extensions of $\mathbb{Q}$, and then attempt to prove that every finite abelian group occurs as a quotient of one of these Galois groups. Since every quotient of a Galois group is again a Galois group with the same base field (proof?), we will be done. Recalling some of our

work from the fall, one place to look for abelian extensions of $\mathbb{Q}$, is at the groups of equations of the form X^n-1. We already showed in the fall that the group $G_{\mathbb{Q}}(X^n-1)$ is isomorphic to a subgroup of $(\mathbb{Z}_n)^*$, the multiplicative group of units of the ring $\mathbb{Z}_n$. Now we will go back and complete that argument by showing $G_{\mathbb{Q}}(X^n-1) \cong (\mathbb{Z}_n)^*$. The key ingredient will be the proof that the "cyclotomic" polynomials are all irreducible in characteristic zero. (This is false in positive characteristic.)

Remark on "primitive roots of unity": In characteristic zero, (and also in characteristic p where p does not divide n), the polynomial X^n-1 has n distinct roots (in some splitting field) which form a cyclic multiplicative group of order n. A generator ζ of this group is called a "primitive" nth root of unity. If we define a primitive nth root of unity to be a non zero element in an extension of k which has order n, then k has primitive nth roots of unity iff charac(k) = 0, or if charac(k) = $p > 0$ and p does not divide n.

Cyclotomic polynomials: If the base field is $\mathbb{Q}$, and $n \geq 1$, define $\Phi_n = \prod(X-\zeta)$, product over all primitive nth roots ζ of unity. Φ_n is called the "nth cyclotomic polynomial". Note: degree $\Phi_n = \varphi(n) =$ #{positive integers relatively prime to, and less than, n}, where φ is the Euler "phi function".

Examples:
$\Phi_1 = X-1$; $\Phi_2 = X+1$; $\Phi_3 = X^2+X+1$; $\Phi_4 = X^2+1$, $\Phi_5 = X^4+X^3+X^2+X+1$.

It appears that these polynomials actually have integer coefficients, which is one of the first things we want to prove.
Note: $X^2-1 = \Phi_1\Phi_2$; $X^3-1 = \Phi_1\Phi_3$; $X^4-1 = \Phi_1\Phi_2\Phi_4$; $X^5-1 = \Phi_1\Phi_5$.

Lemma: For all n, $X^n-1 = \prod\Phi_r$, product over all $r \geq 1$ dividing n.
proof: Note that the roots of Φ_r are precisely the elements of order r in the group $\mathbb{C}^*$, and the roots of X^n-1 are the elements of $\mathbb{C}^*$ whose order divides n. Hence the left and right sides have the same roots, and the same degree, and both sides are monic. QED.

Corollary: For all n, Φ_n has integer coefficients

proof: This holds for n=1. Assume it holds for all $r < n$. From the formula in the lemma, we have $X^n - 1 = h(X) \cdot \Phi_n$, where $h(X) = \prod' \Phi_r$, product for all $r < n$, r dividing n. By induction h(X) has integer coefficients and is monic, hence primitive. Since $X^n - 1$ is primitive, Φ_n is also primitive, in particular Φ_n has integer coefficients. **QED**.

Irreducibility of cyclotomic polynomials over Q

Proposition: For all n, Φ_n is irreducible over $\mathbb{Q}$.

Remark: We already know this result when $n = p$ is prime, and then $\Phi_p = X^{p-1} + X^{p-2} + \ldots\ldots + X + 1$.

Corollary: The Galois group $G_{\mathbb{Q}}(X^n - 1) \cong (\mathbb{Z}_n^*, \cdot)$.

proof: We know $G \subset \mathbb{Z}_n^*$ is a subgroup, so it suffices to show that $\#(G) = \varphi(n) = \#(\mathbb{Z}_n^*)$. If $\mathbb{Q} \subset F_n$ is the splitting field for $X^n - 1$, $F_n = \mathbb{Q}(\zeta)$ for any primitive nth root of unity ζ, so it suffices to show the minimal polynomial of ζ over $\mathbb{Q}$ has degree $= \varphi(n)$. Hence the corollary follows from the proposition. **QED**.

proof of prop: Let ζ be a primitive nth root of unity with irreducible primitive (not necessarily monic) polynomial f over $\mathbb{Q}$. We claim $\Phi_n = f$. To prove this it suffices to show every primitive nth root of unity is a root of f. Now every primitive nth root of unity has form ζ^r where $\gcd(r,n) = 1$, and if $r = \prod p_i$, then $\zeta^r = (\zeta^{p_1})^{p_2 \cdots p_s} = ((\zeta^{p_1})^{p_2})^{p_3 \cdots p_s} = \ldots$etc. Hence it sufices to prove the:

Lemma: If ζ is a primitive nth root of unity, f an irreducible primitive polynomial in $\mathbb{Z}[X]$ such that $f(\zeta) = 0$, and p is a prime not dividing n, then $f(\zeta^p) = 0$.

proof: Let g be an irreducible primitive polynomial of ζ^p over $\mathbb{Q}$. If f and g have a common root, then each is a primitive irreducible polynomial for that root, hence $f = \pm g$. Then $f(\zeta^p) = \pm g(\zeta^p) = 0$, proving the lemma. Hence we may assume f,g have no common root, so that $\gcd(f,g) = 1$. Then since every root of unity is a root of $X^n - 1$, the product fg divides $X^n - 1$, so $X^n - 1 = f(X)g(X)h(X)$ for some primitive polynomial h. We know all nth roots of unity are distinct over $\mathbb{Q}$, but in fact the same holds modulo p since p does not divide n; i.e. the only root mod p of $(X^n - 1)' = nX^{n-1}$ is zero, not a root of $X^n - 1$, mod p. Now reduce the equation $X^n - 1 = f(X)g(X)h(X)$ mod p. Since the left hand side still has degree n, and no multiple roots, f,g

each have the same degree mod p that they have over $\mathbb{Q}$, so they do both have roots in $\mathbb{Z}_p$, but f and g cannot have a common root mod p. We will show in a moment however, that every root of f mod p is also a root of g. First observe the following:

Sublemma: For any g in $\mathbb{Z}[X]$, $g(X^p) = (g(X))^p$, mod p.
proof: $g(X) = X^a \pm X^b \pm \ldots \pm X^c$, for some exponents a,b,...,c, not necessarily all different. Thus $(g(X))^p = (X^a \pm X^b \pm \ldots \pm X^c)^p = (X^a)^p \pm (X^b)^p \pm \ldots \pm (X^c)^p = (X^p)^a \pm (X^p)^b \pm \ldots \pm (X^p)^c = g(X^p)$. **QED.**

This gives a contradiction as follows. Since $g(\zeta^p) = 0$, ζ satisfies $g(X^p)$, so f(X) divides $g(X^p)$. Thus $g(X^p) = f(X)m(X)$ in $\mathbb{Z}[X]$. Reducing mod p, we get $g(X^p) = (g(X))^p = f(X)m(X)$ in $\mathbb{Z}_p[X]$. Since degree f is still positive mod p, f has a root mod p, which is thus also a root of g. This contradicts the conclusion reached just before the sublemma. **QED for the Lemma and the Prop., i.e. Φ_n is irred.**

In the next section we study the groups $\mathbb{Z}_n^*$ a bit more to see just how many abelian groups are quotients of groups of this form.

§17) A product decomposition for the groups $\mathbb{Z}_n^*$

We know the groups $\mathbb{Z}_n^*$ are abelian, and we would like to classify them further. The simplest abelian groups are the cyclic ones, and we know that the product of abelian groups is abelian. This lets us construct a large array of abelian groups, by taking arbitrary products of cyclic groups. How many different abelian groups do we get this way? Eventually we will show that all finite abelian groups are isomorphic to products of cyclic groups, but for now we want to prove a partial decomposition result for the groups $\mathbb{Z}_n^*$. First we recall the definition of a finite product of abelian groups.

Definition: (i) If A, B are abelian groups, their "product" $A \times B$ is the group whose underlying set is the Cartesian product $A \times B$ = {all ordered pairs (a,b) with a in A, b in B}, and with group operation defined "componentwise"; i.e. $(a_1, b_1)\cdot(a_2, b_2) = (a_1a_2, b_1b_2)$
(ii) More generally, if $A_1,\ldots,A_n$ is a finite collection of abelian groups, we define their product as the set $\prod_i A_i = A_1 \times A_2 \times \ldots \times A_n$, with

componentwise group operation.

Proposition: If $n = ab$, where $\gcd(a,b) = 1$, then $\mathbf{Z}_n^* \cong \mathbf{Z}_a^* \times \mathbf{Z}_b^*$.

Lemma: (Chinese Remainder Theorem, CRT): If $a, b > 1$, and $\gcd(a,b) = 1$, then for any integers s,t, there is an integer m such that $m \equiv s \pmod a$, and $m \equiv t \pmod b$
proof: This says the map $\varphi: \mathbf{Z} \to \mathbf{Z}_a \times \mathbf{Z}_b$ where $\varphi(m) = ([m]_a, [m]_b)$, is surjective. If we note that $\varphi(m) = 0$ iff both a,b divide m, iff ab divides m, we see that there is an induced map $\tilde{\varphi}: \mathbf{Z}_{ab} \to \mathbf{Z}_a \times \mathbf{Z}_b$ which is injective. Since source and target have the same finite cardinality, $\tilde{\varphi}$ is also surjective. Then φ is surjective too. QED.

Remark: The map $\tilde{\varphi}$ is a ring map, hence a ring isomorphism.

proof of prop: Since $\tilde{\varphi}$ is a ring isomorphism $\mathbf{Z}_n \cong \mathbf{Z}_a \times \mathbf{Z}_b$, it induces an isomorphism of units $\mathbf{Z}_n^* \cong (\mathbf{Z}_a \times \mathbf{Z}_b)^* \cong \mathbf{Z}_a^* \times \mathbf{Z}_b^*$
QED prop.

Remark on notation: We use $\mathbf{Z}_n$, $\mathbf{Z}/(n)$, and $\mathbf{Z}/n$ interchangeably.

Corollary: If $n = \prod_i p_i^{r_i}$, all p_i distinct primes, and if $n_i = p_i^{r_i}$, then $\mathbf{Z}_n^* \cong \prod_i (\mathbf{Z}/n_i)^*$.

Corollary: If $n = \prod_i p_i$, all p_i *distinct* primes, then $\mathbf{Z}_n^* \cong \prod_i (\mathbf{Z}/p_i)^* \cong \prod_i (\mathbf{Z}/(p_i - 1))$.

The previous corollary shows at least if n has no repeated prime factors, then $\mathbf{Z}_n^*$ is isomorphic to a product of finite cyclic groups If we appeal to the famous theorem of Dirichlet, we can now prove that every finite product of finite cyclic groups occurs as a Galois group over $\mathbf{Q}$.

Theorem of Dirichlet: If $m > 0$, and $\gcd(m,a) = 1$, then there are infinitely many primes p in $\mathbf{Z}$ such that $p \equiv a \pmod m$.
Corollary: If $m > 0$, there are infinitely many primes $p \equiv 1 \pmod m$.
Remark: We include below a summary of Serre's account of the proof from his book A Course in Arithmetic.

Corollary: Every finite product of finite cyclic groups is a quotient of a group of form $\mathbf{Z}_n^*$ for some n.
proof: Let $G = \Pi_i (\mathbf{Z}/n_i)$, (where $\mathbf{Z}/r$ denotes $\mathbf{Z}_r$). Choose $p_1 \equiv 1 (\mathrm{mod}\ n_1)$. Then choose $p_2 > p_1$ such that $p_2 \equiv 1 (\mathrm{mod}\ n_2)$ Continue so that for all i, $p_i \equiv 1 (\mathrm{mod}\ n_i)$ and $p_i > p_{i-1}$ These choices are possible by Dirichlet's theorem Then for $n = \Pi_i p_i$, since n has no repeated prime factors, $\mathbf{Z}_n^* \cong \Pi_i (\mathbf{Z}/p_i)^* \cong \Pi(\mathbf{Z}/(p_i-1))$. Now $p_i \equiv 1 (\mathrm{mod}\ n_i)$ implies $p_i - 1 \equiv 0 (\mathrm{mod}\ n_i)$ so for each i, $p_i - 1 = n_i m_i$, for some $m_i \geq 1$. Then consider the subproduct $\Pi(\mathbf{Z}/m_i) \subset \Pi(\mathbf{Z}/(p_i-1)) \cong \mathbf{Z}_n^*$, with quotient $\Pi(\mathbf{Z}/n_i) = G$. (We embed each $(\mathbf{Z}/m_i)$ as a subgroup of $(\mathbf{Z}/(p_i-1)) \cong (\mathbf{Z}/n_i m_i)$ by the map $[x] \mapsto [n_i x]$, and thus $\Pi(\mathbf{Z}/m_i)$ embeds in $\Pi_i (\mathbf{Z}/(p_i-1))$ as a subproduct by $(\ldots,[x_i],\ldots) \mapsto (\ldots,[n_i x_i],\ldots)$.) This solves our problem. **QED**.

Corollary: Every finite product G of finite cyclic groups is a Galois group over $\mathbb{Q}$; in fact $G \cong G_{\mathbb{Q}}(L)$ where $L \subset F_n$ is a subfield of some cyclotomic field $F_n = \mathbb{Q}(\zeta)$, where ζ is a primitive nth root of 1.
proof: If $G = \Pi(\mathbf{Z}/n_i)$, choose n as in the previous corollary so that $G \cong (\mathbf{Z}_n^*)/H$ for some subgroup $H \subset \mathbf{Z}_n^*$. Then G is the Galois group of the fixed field of H, in the splitting field F_n of $X^n - 1$ **QED**.

Remark: The Chinese remainder theorem says precisely that a decomposition of an abelian group as a product of cyclic groups is far from unique. I.e. if $\{p_i\}$ is a collection of distinct primes, and $n = \Pi p_i$, we proved that $\mathbf{Z}/n \cong \Pi(\mathbf{Z}/p_i)$ Thus on the left we have a product with one cyclic factor, while on the right we have a product with a large number of different cyclic factors. Moreover, every distinct partitioning of the factors $\{p_i\}$ into r disjoint sets yields a different product representation of $\mathbf{Z}/n$, with r cyclic factors.

We finish this topic in the next section by proving that in fact every finite abelian group is isomorphic to a product of cyclic groups. We also give a condition which makes a decomposition unique.

§18) Every finite abelian group G is isomorphic to a product of cyclic groups, hence is a Galois group over $\mathbb{Q}$.
Theorem: If G is any finite abelian group, there exists a finite

normal extension $\mathbb{Q} \subset L$, such that $G(L/\mathbb{Q}) \cong G$. In fact there is a positive integer n such that $L \subset F_n$ = splitting field of $X^n - 1$.

Definition: An extension is called "abelian" iff its Galois group is abelian. Note that every subfield of an abelian extension of $\mathbb{Q}$ is also Galois (and abelian) over $\mathbb{Q}$.

Remark: We already know every cyclotomic extension of $\mathbb{Q}$ is an abelian extension of $\mathbb{Q}$, and thus so is every subfield. A famous theorem of Kronecker says that conversely every abelian extension of $\mathbb{Q}$ is a subfield of a cyclotomic extension, so cyclotomic extensions are in a sense the only ones that yield abelian extensions. This shows some of the complexity of the study of roots of unity, i.e. all finite abelian groups occur as Galois groups of subfields of cyclotomic extensions.

Remarks: (i) Shafarevich published some 40 years ago a difficult proof that every finite solvable group G occurs as the Galois group of some extension of $\mathbb{Q}$. He recently pointed out an error in his proof related to the prime 2, and gave indications of how to repair it.
(ii) A finite group is said to be "nilpotent" iff it is a direct product of p-groups for various p. Any finite abelian group is nilpotent as we will show below, and a finite nilpotent group is solvable. On the other hand D_4 is nilpotent but not abelian, and S_3 is solvable but not nilpotent. Thus the implications abelian $\Rightarrow$ nilpotent $\Rightarrow$ solvable which hold for finite groups, are not reversible. The special case of Shafarevich's result where G is a p-group, i.e. of order p^n, with p an odd prime, seems to have been proved beyond doubt, and implies that every "nilpotent" group of odd order is a Galois group over $\mathbb{Q}$.

It follows from the last corollary of the previous section that to prove the theorem above, we only need to prove the **fundamental theorem of finite abelian groups**, that every finite abelian group is isomorphic to a product of cyclic groups. We will prove later in this section the following more precise statement:

Theorem(FTFAG): Let G be a finite abelian group of order ≥ 2. Then there is a unique sequence of integers $n_1, n_2, \ldots, n_r$, all ≥ 2, such that for all i, n_i divides n_{i+1}, and such that $G \cong \prod(\mathbb{Z}/n_i)$.

Remarks: The whole difficulty in proving this proposition is present already in deciding the following question: If A is a finite abelian group and $(x) \subset A$ is a cyclic group such that the quotient $A/(x) \cong (z)$ is also cyclic, when is $A \cong$ the product $(x) \times (z)$? Note that this is not entirely obvious. For example, if $A = \mathbf{Z}_6$, and $x = [2]$ generates $(x) \cong \mathbf{Z}_3$, then $A/(x) \cong \mathbf{Z}_2$, and indeed $A \cong \mathbf{Z}_2 \times \mathbf{Z}_3$, by the CRT. On the other hand if $A = \mathbf{Z}_8$, and $x = [2]$ generates $(x) \cong \mathbf{Z}_4$, then $A/(x) \cong \mathbf{Z}_2$, but A is not isomorphic to the product $\mathbf{Z}_2 \times \mathbf{Z}_4$, since $\mathbf{Z}_8$ contains elements of order 8, but $\mathbf{Z}_2 \times \mathbf{Z}_4$ does not.
The key to recognizing a group A which is $\cong$ to a product $B \times C$, is finding the appropriate subgroups isomorphic to B and C inside of A. I.e. if $A = B \times C$, then A contains the subgroups $B \times \{1\} \cong B$ and $\{1\} \times C \cong C$. In our first example above, $\mathbf{Z}_6$ contains the subgroups $\{0,2,4\} \cong \mathbf{Z}_3$, and $\{0,3\} \cong \mathbf{Z}_2$. But be careful! In our second example, $\mathbf{Z}_8$ also contains the subgroups $\{0,4\} \cong \mathbf{Z}_2$ and $\{0,2,4,6\} \cong \mathbf{Z}_4$. The difference is that in the first case the two subgroups were almost disjoint, intersecting only in $\{0\}$, while in the second example they intersect in $\{0,4\} \cong \mathbf{Z}_2$. Since in $B \times C$, the subgroups $B \times \{1\}$ and $\{1\} \times C$ intersect only in $\{1\} \times \{1\}$, the good case is when the subgroups have only the identity in common. In cases where this is easy to check, decomposition theorems are easy, as in the following exercise.

Exercise *118) **(i)** If B,C are subgroups of a finite abelian group A such that $\#(A) = \#(B)\#(C)$, and $B \cap C = \{1\}$, prove the map φ: $B \times C \to A$, $\varphi(x,y) = xy$, is an injective homomorphism, hence an isomorphism.
(ii) If A is a finite abelian group, and B,C are subgroups of relatively prime order in A such that $\#(A) = \#(B)\#(C)$, prove $A \cong B \times C$.
(iii) If A is a finite abelian group, prove A is $\cong$ to the product of its Sylow p-subgroups, [i.e. A is nilpotent].

In view of the previous exercise, the hard part is to recognize when a group A is isomorphic to $\mathbf{Z}_2 \times \mathbf{Z}_4$, or to $\mathbf{Z}_9 \times \mathbf{Z}_{35}$, or in general to $B \times C$, where $\#(B)$ and $\#(C)$ are not relatively prime. So let's assume the decomposition theorem for finite abelian groups is true. Then how would we go about proving it? Induction seems like a natural tool, so we need to figure out how to split off the first factor. I.e., if A is a finite abelian group that is not cyclic, how do we even prove that it splits up as a product at all? If we could do that, we could

say by induction that the two factors were both isomorphic to products of cyclic groups and we would be done. The examples above have shown that A can have a cyclic subgroup $(x) \subset A$ such that $A/(x) \cong (y)$ is cyclic, and yet A may not be isomorphic to $(x) \times (y)$. Nonetheless if the theorem is true, and A is finite, abelian, and not cyclic, it must be possible to choose a cyclic subgroup of A which will split off as a factor. If we look at the statement of the theorem above, we see that there is a "largest" factor, the one $\mathbf{Z}/n_r$ whose order is a multiple of the order of every other factor. This reflects the fact we already know, that there is an element of A whose order is a multiple of the order of every other element. Such elements can be seen in the product decomposition to be those whose last component is a generator of the factor $\mathbf{Z}/n_r$. For example, an element ([x],[y]) of $A = \mathbf{Z}_9 \times \mathbf{Z}_{18}$ whose second entry [y] = [1], will have order 18, the maximum possible, and every element of A has order dividing 18. Since an element x of maximal order in A always exists, this gives us somewhere to start. So it seems plausible to try to show a cyclic subgroup (x) generated by an element of maximal order should split off from A.

The following case is the simplest possible:
Conjecture: Let A be a finite abelian group, and let x in A have maximal order. If $A/(x) \cong (\beta)$ is cyclic, then $A \cong (x) \times (\beta)$.

To prove this we need to find a copy of the group (β) inside of A, and one which intersects (x) trivially. If we consider the canonical map $\varphi : A \to A/(x) \cong (\beta)$, then $(x) = \ker(\varphi)$, so a subgroup of A which is isomorphic to (β), and which intersects (x) trivially, is precisely a subgroup (α) which maps isomorphically to (β) via the restricted map $\varphi : (\alpha) \to (\beta)$. This means we need to find an element α of A such that $\varphi(\alpha) = \beta$, and such that $\mathrm{ord}(\alpha) = \mathrm{ord}(\beta)$.

The following lemma is therefore the central ingredient of the proof of the decomposition theorem for finite abelian groups.

Lemma 1: Let A be a finite abelian group, and φ: A$\to(\beta)$ a surjective homomorphism to a cyclic group (β). If $\ker(\varphi) = (x)\subset A$, is cyclic and generated by an element x of maximal order in A, then there is an element α of A such that $\varphi(\alpha) = \beta$, and $\text{ord}(\alpha) = \text{ord}(\beta)$.
proof: Let $\text{ord}(x) = m$, and $\text{ord}(\beta) = n$. We already know the maximal order of an element of A is the l.c.m. of the orders of all elments of A. Thus if γ is any element of A, $\gamma^m = 1$. If moreover $\varphi(\gamma)=\beta$, then $1 = \varphi(1) = \varphi(\gamma^m) = \beta^m$, so n|m also.
Now the full set of preimages of β are the elements of form $\alpha = \gamma\cdot x^s$, for any s. So we want to find an integer s such that $\text{ord}(\gamma\cdot x^s) = n$. Note that $(\gamma\cdot x^s)^n = 1$ iff $\gamma^n = (x^{sn})^{-1} = (x^{-1})^{sn}$. In particular γ^n would have to equal a power of x. Now $\varphi(\gamma^n) = \beta^n = 1$, so γ^n is in $\ker(\varphi) = (x)$. Ie. $\gamma^n = x^r$ for some r. The following claim is the main point of this lemma.

Claim: If $\gamma^n = x^r$, [where $n = \text{ord}(\beta)$, and $\varphi(\gamma)=\beta$], then n divides r.
proof of claim: We know $\gamma^m = 1$, that m = an for some a, and that $\gamma^n = x^r$. Then $1 = \gamma^m = \gamma^{an} = x^{ar}$ Thus m=an|ar, so that indeed n|r. **QED for Claim**

If r is chosen so that $\gamma^n = x^r$, then by the claim there exists b such that r = bn. Then put $\alpha = \gamma\cdot x^{-b}$, so that $\varphi(\alpha) = \beta$, and $\alpha^n = \gamma^n\cdot x^{-bn} = \gamma^n x^{-r} = 1$. Thus $\text{ord}(\alpha)|n$, but we already know that n divides $\text{ord}(\alpha)$ whenever $\varphi(\alpha) = \beta$. So $\text{ord}(\alpha) = n$. Thus α is the element sought in the Lemma. **QED lemma 1**.

Next we present the language used to describe isomorphisms of finite products of abelian groups. Homomorphisms into and out of such products are easy to describe in terms of homomorphisms into or out of the factors. We summarize these elementary properties in the next lemma.

Lemma 2: (i) If A,B, C are abelian groups and φ:A$\to$C, ψ:B$\to$C are homomorphisms out of A,B, there is a unique homomorphism F:A$\times$B$\to$C defined by $F(a,b) = \varphi(a)\psi(b)$. If we regard A $\cong$ A$\times\{1\}$ and B $\cong \{1\}\times$B, as subgroups of A$\times$B via the canonical injections $a\mapsto(a,1)$ and $b\mapsto(1,b)$, then F is the unique homomorphism A$\times$B$\to$C which restricts to φ on A, and to ψ on B

(ii) More generally, if $A_1,\ldots,A_n$ is a finite collection of abelian groups and if for each i, $\varphi_i:A_i\to C$ is a homomorphism out of A_i, there is a unique homomorphism $F:\Pi A_i\to C$ defined by $F(a_1,\ldots,a_n) = \Pi\varphi_i(a_i)$. If we regard $A_i \cong \{1\}\times \{1\}\times A_i\times\{1\}\times\ldots\{1\}$, as a subgroup of $\Pi_i A_i$ via the canonical injection $a_i\mapsto(1,\ldots,1,a_i,1,\ldots,1)$, then F is the unique homomorphism that restricts, on each A_i, to φ_i.
(iii) If A,B,C are abelian groups and $\varphi\ C\to A$, $\psi:C\to B$ are homomorphisms into A,B, there is a unique homomorphism $F:C\to A\times B$ such that $F(z) = (\varphi(z),\psi(z))$. If A and B are regarded as quotients of $A\times B$ via the canonical projections π_A: $A\times B\to A$, π_B: $A\times B\to B$, then F is the unique homomorphism whose compositions with these projections are $(\pi_A)\circ F = \varphi$ and $(\pi_B)\circ F = \psi$.
(iv) More generally, if $A_1,\ldots,A_n$ is a finite collection of abelian groups and if for each i, $\varphi_i:C\to A_i$ is a homomorphism into A_i, there is a unique homomorphism $F:C\to\Pi A_i$ defined by $F(a_1,\ldots,a_n) = (\varphi_1(a_1),\ldots,\varphi_n(a_n))$. If we regard A_i as a quotient of $\Pi_i A_i$ via the canonical projection π_i: $(a_1,\ldots,a_i,\ldots,a_n)\mapsto a_i$, then F is the unique homomorphism whose composition with each π_i, is $\pi_i\circ F = \varphi_i$.
proof: This is easily verified. No ideas are required.
QED lemma 2.

Now we formalize the splitting principle described in the discussion just above lemma 1.

Lemma 3: Assume $\varphi:G\to H$ is a homomorphism of abelian groups, which has a "right inverse", i.e. suppose there is a homomorphism $\psi:H\to G$ such that $\varphi\circ\psi = id_H$. Then $G\cong H\times\ker(\varphi)$. [We say the homomorphism ψ "splits" G.]
proof: To prove two groups are isomorphic, by definition we need to find homomorphisms in both directions that are mutually inverse. With the given assumptions, it is easy to find a homomorphism $f: H\times\ker(\varphi)\to G$. By lemma 2, we need a pair of homomorphisms $H\to G$, and $\ker(\varphi)\to G$. Such maps are immediately at hand, $\psi:H\to G$ and the inclusion homomorphism $j:\ker(\varphi)\subset G$. The resulting map $f: H\times\ker(\varphi)\to G$ takes (β,γ) to $f(\beta,\gamma) = \psi(\beta)\cdot\gamma$. It is perhaps not as immediate to find a homomorphism in the other direction, but with a little experimentation we can cook up $g: G\to H\times\ker(\varphi)$, by setting $g(z) = (\varphi(z), [\psi(\varphi(z))]^{-1}\cdot z)$. [I.e. $\psi\circ\varphi$ is not the identity on G, because ψ need not take $\varphi(z)$ back to z, but ψ takes $\varphi(z)$ back to something that

maps to $\varphi(z)$, so z and $\psi(\varphi(z))$ both map to the same object in H, namely to $\varphi(z)$; hence $[\psi(\varphi(z))]^{-1}\cdot z$ maps to 1, and thus belongs to $\ker(\varphi)$.]
Now we claim that f, g are mutually inverse. Assume z is in G and compute $(f\circ g)(z) = f(\varphi(z), [\psi(\varphi(z))]^{-1}\cdot z) = \psi(\varphi(z))\cdot[\psi(\varphi(z))]^{-1}\cdot z = z$. In the other direction, if z is in H and w in $\ker(\varphi)$, then $\varphi(w) = 1$, so $(g\circ f)(z,w) = g(\psi(z)\cdot w) = (\varphi(\psi(z)\cdot w), [\psi(\varphi(\psi(z)\cdot w))]^{-1}\cdot\psi(z)\cdot w) = (\varphi(\psi(z))\cdot\varphi(w), [\psi(z)]^{-1}\psi(z)\cdot w) = (z,w)$ **QED lemma 3.**

Now we can deduce the main result:

Theorem(FTFAG): Let G be a finite abelian group of order ≥ 2. Then there is a unique sequence of integers $n_1, n_2, \ldots, n_r$, all ≥ 2, such that for all i, n_i divides n_{i+1}, and such that $G \cong \Pi_i\ (\mathbf{Z}/n_i)$.
proof of existence: The theorem holds for $\#(G) = 2$, so we may assume the theorem holds for all non trivial abelian groups of order less than $\#(G)$. Recall that if $m > 1$ is the l.c.m of the orders of the elements of G, we have proved that G contains an element x of order m. Consider the abelian quotient group $H = G/(x)$, of order $\#(G)/m$. If H is trivial then $G \cong (x) \cong \mathbf{Z}_m$ is cyclic and we are done. If H is non trivial, then by induction we may assume $H \cong \Pi_i\ (\beta_i) \cong \Pi_i\ (\mathbf{Z}/n_i)$, $i = 1, \ldots, r$, where $(\beta_i)\subset H$ is a cyclic group generated by β_i, $\text{order}(\beta_i) = n_i$, and n_i divides n_{i+1}, for all i.
Look at the canonical surjective map $\varphi: G \to H$. We claim that φ has a right inverse, i.e. that there is a homomorphism $\psi: H \to G$ such that $\varphi\circ\psi = \text{id}_H$. By lemma 3 above, this would prove that $G \cong \mathbf{Z}_m \times H \cong \mathbf{Z}_m \times (\Pi_i\ \mathbf{Z}/n_i)$. By lemma 2 (ii) above, the only thing needed to construct such a ψ, is to prove that for each i, there is an element α_i in G such that $\varphi(\alpha_i) = \beta_i$ and $\text{order}(\alpha_i) = \text{order}(\beta_i) = n_i$. [I.e. given such α_i, we can then define ψ uniquely by setting $\psi(\beta_i) = \alpha_i$, getting a homomorphism such that $(\varphi\circ\psi)(\beta_i) = \beta_i$, for all i. Then $\varphi\circ\psi$ would agree with id_H on a set of generators for H, hence we would have $\varphi\circ\psi = \text{id}_H$.]
Now, for each i, consider the abelian subgroup $A_i = \varphi^{-1}((\beta_i)) = \varphi^{-1}(\mathbf{Z}/n_i) \subset G$. Then the restriction $\varphi: A_i \to (\beta_i)$ is a surjective homomorphism to a cyclic group (β_i), with $\ker(\varphi) = (x)$ cyclic and generated by an element x of maximal order in G, hence also maximal in A_i. Then by lemma 1 above there is an element α_i in A_i

such that $\varphi(\alpha_j) = \beta_j$ and $ord(\alpha_j) = ord(\beta_j) = n_j$ Moreover, by the first part of the argument for lemma 1, every n_j divides m. Hence $G \cong H \times \mathbf{Z}_m \cong \mathbf{Z}/n_1 \times \ldots\ldots \times \mathbf{Z}/n_r \times \mathbf{Z}/m$, where $n_1|n_2|\ldots|n_r|m$. QED.

proof of uniqueness (sketch):
Remarks: The uniqueness proof seems easier using the preliminary decomposition into a product of Sylow subgroups, proved in exercise#107) above In practice, the CRT makes this a useful way to find the "standard" decomposition, (the one described in the proposition above), of an arbitrarily given product of cyclic groups.

First we prove uniqueness of the standard decomposition of an abelian p-group
Lemma: If $A \cong \mathbf{Z}/p^{r_1} \times \mathbf{Z}/p^{r_2} \times \ldots\ldots \times \mathbf{Z}/p^{r_n}$ is a standard cyclic decomposition of a non trivial abelian p-group A, i.e one such that $1 \leq r_i \leq r_{i+1}$ for all i, then the sequence of exponents $r_1, r_2, \ldots, r_n$ is uniquely determined by the isomorphism class of the group A.
proof: In this proof we will write the operation in A additively, as in $\mathbf{Z}_m$. We use induction on the order of A Since the lemma is trivial for $\#(A) = p$, we may assume the result holds for all groups of order less than $\#(A)$.

We will use two auxiliary subgroups of A in the proof:
Definition: (i) The "p-torsion" of A, denoted A(p), is the subgroup of elements x of A such that $px = 0$; (or, if we were using multiplicative notation, such that $x^p = 1$)
(ii) The subgroup pA = {elements of form px for x in A}.

Remark: The subgroups A(p), and pA, are simply the kernel and the image respectively of the homomorphism $\varphi: A \to A$, $\varphi(x) = px$
Thus both are determined up to isomorphism by the isomorphism class of A.

We can reconstruct the sequence of integers in the decomposition of A, from the subgroups A(p) and pA as follows: let us denote the ith factor of A by $A_i = \mathbf{Z}/p^{r_i}$.
Then observe:
(i) Since an element of a product is trivial iff each entry is trivial, and is a multiple of p iff each entry is so, we have $A(p) \cong$

$A_1(p)\times....\times A_n(p)$, and $pA \cong pA_1\times \ \ \times pA_n$.
(ii) Since px is a multiple of p^r iff x is a multiple of p^{r-1}, for each i we have $A_i(p)$ = {multiples of p^{ri-1} in A_i} $\cong \mathbf{Z}_p$.
(iii) Consequently, $A(p) \cong (\mathbf{Z}_p)^n$, hence $\#(A(p))$ reveals the number n of factor groups in the decomposition of A.
(iv) Since pA_i = {multiples of p in A_i} $\cong \mathbf{Z}p^{ri-1}$, we have $pA \cong \mathbf{Z}/p^{r1-1} \times \mathbf{Z}/p^{r2-1} \times.....\times \mathbf{Z}/p^{rn-1}$ Since $\#(pA) < \#(A)$, by induction the exponents r_i-1 in this sequence which are greater than zero, are determined by the isomorphism class of pA, hence by that of A.
(v) Combining (iii) and (iv), we recover both the sequence of exponents r_i greater than one, and the number which equal one, hence all exponents. (The number of r_i equal to one, plus the number which are greater than one, equals n.) **QED lemma.**

Remark: One can avoid induction in this lemma by using the subgroups of "p^s-torsion" = $A(p^s)$ = {all x in A such that $p^sx = 0$}. Then, $\#(A(p)) = p^n$ iff n = number of factors in the decomposition of A with exponent $r_i \geq 1$. Also $\#(A(p^2)) \leq p^{2n}$, and $\#(A(p^2)) = p^{n+k}$ iff k = number of factors with exponent $r_i \geq 2$. Next $\#(A(p^3)) \leq p^{n+2k}$, and $\#(A(p^3)) = p^{n+k+m}$ iff m = number of factors with exponent $r_i \geq 3$ Etc.....

Completion of the uniqueness proof: We have proved the sequence of exponents occurring in the standard decomposition of each Sylow p-subgroup of a finite abelian group A is determined by the isomorphism class of A. One can show also that the standard decomposition of A and that of its Sylow subgroups determine each other We indicate how this is done in an example as follows:

Suppose $A \cong \mathbf{Z}_4\times\mathbf{Z}_4\times\mathbf{Z}_{20}\times\mathbf{Z}_{300}\times\mathbf{Z}_{1800}$ is the standard decomposition of A. Then by the CRT, we can decompose each factor into cyclic Sylow subgroups: $\mathbf{Z}_4 = \mathbf{Z}_4$, $\mathbf{Z}_{20} \cong \mathbf{Z}_4\times\mathbf{Z}_5$, $\mathbf{Z}_{300} \cong \mathbf{Z}_4\times\mathbf{Z}_3\times\mathbf{Z}_{25}$, $\mathbf{Z}_{1800} \cong \mathbf{Z}_8\times\mathbf{Z}_9\times\mathbf{Z}_{25}$. Then the Sylow subgroups of A decompose as the products of those of the factors: $A_2 \cong \mathbf{Z}_4\times\mathbf{Z}_4\times\mathbf{Z}_4\times\mathbf{Z}_4\times\mathbf{Z}_8$; $A_3 \cong \mathbf{Z}_3\times\mathbf{Z}_9$; $A_5 \cong \mathbf{Z}_5\times\mathbf{Z}_{25}\times\mathbf{Z}_{25}$. But now we can also go back to the decomposition of A as follows: first consider each Sylow group to have the same number of factors by extending with trivial factors. Eg. consider $A_3 = \{1\}\times\{1\}\times\{1\}\times\mathbf{Z}_3\times\mathbf{Z}_9$. Then for each Sylow subgroup,

order the cardinalities of the factors by size. A_2: (4,4,4,4,8); A_5: (1,1,5,25,25); A_3: (1,1,1,3,9). Now multiply these sequences together term by term, to get (4·1·1, 4·1·1, 4·5·1, 4·25·3, 8·25·9) = (4,4,20,300,1800), and we recover the sequence of factors of the original standard decomposition of A. Satisfy yourself that there is no other way to combine these same powers of these primes to get a sequence of integers $n_1,n_2,\ldots,n_r$, such that $n_i|n_{i+1}$
QED for uniqueness.

Exercise #119)
(i) Find the standard decomposition of each of the following abelian groups of order 864, and decide which of them are isomorphic:
$G_1 = \mathbf{Z}_{24}\times\mathbf{Z}_{36}$, $G_2 = \mathbf{Z}_{18}\times\mathbf{Z}_{48}$, $G_3 = \mathbf{Z}_2\times\mathbf{Z}_3\times\mathbf{Z}_{144}$,
$G_4 = \mathbf{Z}_3\times\mathbf{Z}_{288}$, $G_5 = \mathbf{Z}_8\times\mathbf{Z}_9\times\mathbf{Z}_{12}$.
(ii) Do the same for the groups $\mathbf{Z}_n^*$, for all n, $2 \le n \le 20$.
(iii) How many different abelian groups, up to isomorphism, are there of order 90? 105? 108? 128? 864?

Example: Cyclic decomposition of the groups $\mathbf{Z}_n^*$
We can actually decompose the groups $\mathbf{Z}_n^*$ explicitly into a product of computable cyclic factors, using the CRT and a couple of results from number theory. The point is first to decompose n into prime powers, say $n = 2^r\times\Pi_i\, p_i^{r_i}$, where the p_i are distinct odd primes. Then by CRT, we have a ring isomorphism $\mathbf{Z}_n \cong \mathbf{Z}_{2^r} \times \Pi_i\, (\mathbf{Z}/p_i^{r_i})$, and an isomorphism of unit groups $\mathbf{Z}_n^* \cong (\mathbf{Z}_{2^r})^* \times \Pi_i\, (\mathbf{Z}/p_i^{r_i})^*$. Hence it suffices to decompose each of these factors explicitly. In fact these factors are already almost cyclic. More precisely:

Theorem: (i) If p is an odd prime, $(\mathbf{Z}/p^r)^* \cong \mathbf{Z}/[(1-1/p)p^r]$ is cyclic, (ii) $(\mathbf{Z}_{2^r})^*$ is cyclic for $r = 1, 2$, and $\cong (\mathbf{Z}_2 \times \mathbf{Z}_{2^{r-2}})$ if $r \ge 3$.
Proof (sketch) (i):
Lemma: If a generates $\mathbf{Z}_p^*$ where p is prime, then either a has order p-1 or (p-1)p in $(\mathbf{Z}_{p^2})^*$.
proof: The element a has to have order dividing (p-1)p, hence either k where k divides p-1, or else kp where k divides p-1. But by Fermat's little theorem, $a^{(p-1)p} = a^{p-1} \bmod p$, so the only way either of those orders can occur, if a generates $\mathbf{Z}_p^*$, is if k = p-1.
QED.

Lemma: If a generates $\mathbf{Z}_p^*$ where p is prime, and if a has order p-1 in $(\mathbf{Z}_p2)^*$, then a+p has order (p-1)p, and hence generates $(\mathbf{Z}_p2)^*$.
proof: Use binomial theorem to expand $(a+p)^{p-1}$ and show you don't get 1 mod p^2. QED.

Corollary: If a generates $\mathbf{Z}_p^*$ where p is prime, then either a or a+p generates $(\mathbf{Z}_p2)^*$. In particular, $(\mathbf{Z}_p2)^*$ is cyclic.

Lemma: If a generates $(\mathbf{Z}_p2)^*$ where p is an odd prime, then a generates $(\mathbf{Z}_pk)^*$ for all $k > 2$ also.
proof: We need to show that a has order $(p-1)p^{k-1}$ mod p^k, assuming this true for $k = 2$. Do it by induction on k. QED.
QED. for part (i), **Theorem**.

Lemma:
i) Every element of $(\mathbf{Z}2k)^*$ is annihilated by 2^{k-2}, mod 2^k.
ii) The number 5 has order 2^{k-2}, mod 2^k.
Proof: Use induction on k. QED.

Lemma: If an abelian group (G,.) has order 2^{k-1} and an element y of order 2^{k-2}, and an element of order 2 which is not a power of y, then G is a product of the two cyclic groups generated by x and y.
proof: The map taking (n,m) to $x^n y^m$ is an isomorphism from $\mathbf{Z}_2 \times \mathbf{Z}_2k\text{-}2$ to G. QED.

Corollary: If $k \geq 2$, then $(\mathbf{Z}2k)^* \cong \mathbf{Z}_2 \times \mathbf{Z}_2k\text{-}2$.
proof: We know that 5 is an element of order 2^{k-2}, so we have to find an element of order 2 which is not a power of 5. First, since the quotient of the big group by the subgroup generated by 5, has order two, any element will have its square in the subgroup generated by 5. Then note that any element not a power of five will have its square equal to an even power of five, or else 5 will have a square root, and then this square root will generate the whole group, which is known not to be cyclic. Then modifying the element whose square is a power of five by half that power of five, we get the element of order two we want, which is not a power of five.
QED. part (ii), **Theorem**.

Using the previous theorem we can find the standard product decomposition of any $(\mathbf{Z}_n)^*$. For instance if n = 65, then $\mathbf{Z}_{65} \cong \mathbf{Z}_5 \times \mathbf{Z}_{13}$, so $(\mathbf{Z}_{65})^* \cong (\mathbf{Z}_5)^* \times (\mathbf{Z}_{13})^* \cong \mathbf{Z}_4 \times \mathbf{Z}_{12}$. This time we are lucky and the decomposition is already the standard one.
If n = 176, then 176 = (16)(11) = $2^4 \cdot 11$, so $(\mathbf{Z}_{176})^* \cong (\mathbf{Z}_{16})^* \times (\mathbf{Z}_{11})^* \cong (\mathbf{Z}_2 \times \mathbf{Z}_4) \times (\mathbf{Z}_{10})$. This decomposition is not the standard one, but we can now carry out the procedure described above for recovering the standard decomposition, by passing first to the product of Sylow subgroups. I.e. $(\mathbf{Z}_2 \times \mathbf{Z}_4) \times (\mathbf{Z}_{10}) \cong (\mathbf{Z}_2 \times \mathbf{Z}_4) \times (\mathbf{Z}_2 \times \mathbf{Z}_5)$, So the primes are 2,5 and the sequences of prime powers are (2,2,4), and (1,1,5). Multiplying these sequences together gives (2,2,20), so the standard decomposition is $\mathbf{Z}_2 \times \mathbf{Z}_2 \times \mathbf{Z}_{20}$.
If n = 1800 = $(2^3)(3^2)(5^2)$, then $(\mathbf{Z}_{1800})^* \cong (\mathbf{Z}_{23})^* \times (\mathbf{Z}_{32})^* \times (\mathbf{Z}_{52})^* \cong (\mathbf{Z}_2 \times \mathbf{Z}_2) \times (\mathbf{Z}_6) \times (\mathbf{Z}_{20}) \cong (\mathbf{Z}_2 \times \mathbf{Z}_2) \times (\mathbf{Z}_2 \times \mathbf{Z}_3) \times (\mathbf{Z}_4 \times \mathbf{Z}_5) \cong \mathbf{Z}_2 \times \mathbf{Z}_2 \times \mathbf{Z}_2 \times \mathbf{Z}_4 \times \mathbf{Z}_3 \times \mathbf{Z}_5$. Thus the sequences of prime powers are (2,2,2,4), (1,1,1,3), and (1,1,1,5). Multiplying them together we get (2,2,2,60), so the standard decomposition is $(\mathbf{Z}_{1800})^* \cong \mathbf{Z}_2 \times \mathbf{Z}_2 \times \mathbf{Z}_2 \times \mathbf{Z}_{60}$.
If n= 1729 = (7)(13)(19), then $(\mathbf{Z}_{1729})^* \cong \mathbf{Z}_6 \times \mathbf{Z}_{12} \times \mathbf{Z}_{18} \cong \mathbf{Z}_6 \times \mathbf{Z}_6 \times \mathbf{Z}_{36}$.

Exercise #120) Find the standard decomposition of each of these:
(i) $\mathbf{Z}_{237}^*$,
(ii) $\mathbf{Z}_{360}^*$,
(iii) $\mathbf{Z}_{924}^*$.

General Remarks on "products" versus "sums":
(i) Strictly speaking there are two different general concepts, product and sum, for groups. The property we have stated in (i) and (ii) of lemma 2 above, characterizing maps out of the object, is the defining property of a "sum", while the property in (iii) and (iv) of that lemma, characterizing maps into the object, is the defining property of a "product". Since A×B satisfies both properties, it is actually both a product and a sum, of A and B, in the category of abelian groups. Now when the same object satisfies the properties both of a sum and a product, which do you call it? We have decided to call it, truthfully, a product, but take advantage of the fact that it also acts like a sum. Later, for reasons discussed next, we will have to distinguish the two concepts more carefully.

(ii) The construction of sums differs from that of products for non abelian groups, as well as for infinite collections of abelian groups. The property in part (i) of lemma 2 for instance would not be true if C is allowed to be non abelian, since $\varphi(a)$ and $\psi(b)$ might not commute in C, whereas for (a,b) in A×B we must have $\varphi(a)\psi(b) = F(a,1)F(1,b) = F(a,b) = F(1,b)F(a,1) = \psi(b)\varphi(a)$. If A= B= $\mathbf{Z}$, and C is non abelian, the group that has property (i) in the lemma would be, not A×B, but the free group on two generators. Even if C is abelian, the definition of F in part (ii) of lemma 2 would not make sense for an infinite collection of groups A_i, and homomorphisms φ_i, since you cannot multiply together an infinite collection of values $\Pi\varphi_i(a_i)$ in C. In that case, the group having the property in the lemma would be, not $\Pi_i A_i$ but the subgroup of $\Pi_i A_i$ consisting of elements $(\ldots,a_i,\ldots)$ such that $a_i = 1$ for all but a finite number of entries.

§19) Summary of Serre's account of Dirichlet's Theorem:
Introduction: Except for 2 and 5, all primes end in 1, 3, 7 or 9. We know there are infinitely many primes, and we could ask if there are infinitely many that end in each of those four integers. This can be phrased as: are there an infinite number of primes p such that $p \equiv a \pmod{10}$, for every a such that gcd(a,10) = 1? The answer is yes, and the idea for the proof is to show a prime is "equally likely" to have one ending as another, i.e. that given any choice a among the four numbers 1,3,7,9, the proportion of all primes $\leq n$, and ending in a, approaches 1/4 as $n \to \infty$. The proof below generalizes the fact that $\Sigma\, p^{-1}$ diverges. A fancy way to say this is that $g(s) = \Sigma\, p^{-s} \sim \log(1/(s-1))$; i.e. as $s \to 1^+$, g(s) approaches infinity like $\log(1/(s-1))$ Recall this gives a proof that there are infinitely many primes, since otherwise $\Sigma\, p^{-s}$ as a finite sum of exponential functions, would be finite everywhere, in particular at s = 1
Suppose A is a subset of primes consisting say of "half" of all primes in some sense. Then we might expect that the sum $g_A(s) = \Sigma_A\, p^{-s}$ would only go to infinity "half as fast" as the full sum $\Sigma\, p^{-s}$, i.e. we might expect that $g_A(s) \sim (1/2)\log(1/(s-1))$. Let's turn this intuition around and make this a definition. I.e. A consists of "half" of all primes if the quotient $g_A(s)/[(1/2)\log(1/(s-1))]$ approaches 1, as $s \to 1^+$, and we write this as $g_A(s) \sim (1/2)\log(1/(s-1))$
Definition: More generally if A is any subset of primes, we say that

A has <u>density</u> k, where $0 \le k \le 1$, if $\Sigma_A\, p^{-s} \sim k \log(1/(s-1))$.
Easy Remark: If density(A) > 0, then A is infinite.

Dirichlet's theorem: Given $m \ge 2$, and $a > 0$ relatively prime to m, if $P_a = \{p : p \equiv a \pmod m\}$, then Density $(P_a) = 1/\varphi(m) > 0$. In particular there infinitely many primes congruent to a mod m.
Outline of proof:
We must show that $g_a(s) = \Sigma_{p \equiv a \pmod m}\, p^{-s} \sim (1/\varphi(m))\log(1/(s-1))$, or equivalently, that $\varphi(m) g_a(s) \sim \log(1/(s-1))$.

Step One: Write $g_a(s)$ as a linear combination of easier functions
I.e. If $\chi : \mathbf{Z}_m^* \to \mathbf{C}^*$ is a homomorphism, define χ on $\mathbb{N}$ by setting $\chi(n) = \chi(|n|)$ if gcd(n,m) = 1, and by $\chi(n) = 0$ otherwise. Then:
Lemma 9: $\varphi(m)\, g_a(s) = \Sigma_\chi\ \chi(a^{-1})\, f_\chi(s)$, where $f_\chi(s) = \Sigma_p\, \chi(p)/p^s$.
proof: This follows from the "orthogonality relations".

It is easy to see $f_1(s) \sim \log(1/(s-1))$, since by Euler's formula, if $\mathrm{Re}(s) > 0$, $f_1(s) = \Pi\, 1/(1-p^{-s})\ \Pi_{p|m}\,(1-p^{-s}) = \zeta(s)\ \Pi_{p|m}\,(1-p^{-s}) \sim \log(1/(s-1))$
Hence the next lemma would finish the proof:
Lemma 8: If $\chi \ne 1$, $f_\chi(s)$ remains bounded as $s \to 1^+$.

Step two: proof of lemma 8:
The trick is to write f_χ as a difference of two bounded functions.
Lemma: If $L_\chi(s) = \Sigma\ \chi(n)/n^s$, then
$f_\chi(s) = \log(L_\chi(s)) - \Sigma_{p,k\ge 2}\ \chi(p)/kp^{ks}$.
proof: This goes by expanding L_χ as an Euler product, and plugging into the Taylor series for $\log(1/(1-t)) = \Sigma\ t^k/k$.
The boundedness of the second function on the right hand side in this lemma is elementary, in fact $\Sigma_{p,k\ge 2}\ |\chi(p)|/kp^{ks} \le \Sigma_{p,k\ge 2}\ 1/p^{ks}$
$\le \Sigma_p\,(\Sigma_{k\ge 2}\ 1/p^{ks}) = \Sigma_p\,(1/p^s(p^s-1)) \le \Sigma_p\,(1/p(p-1))$
$\le \Sigma_n\,(1/n(n-1)) \le 1$.

The hard part is to prove $\log(L_\chi(s))$ is bounded as $s \to 1^+$, but this would follow from knowing that $L_\chi(1) \ne 0$, if $\chi \ne 1$, since the log of a non zero number is finite.

Step Three: Analysis of $\zeta_m(s) = \Pi_\chi L_\chi(s)$.
Lemma: The function $\zeta_m(s) = \Pi_\chi L_\chi(s)$ is given by a Dirichlet series with ≥ 0 coefficients, and diverges for $s = 1/\varphi(m)$.
proof: $\zeta_m(s)$ has a product expansion convergent in $Re(s) > 1$, as follows $\zeta_m(s) = \Pi_\chi L_\chi(s) = \Pi_\chi \Pi_p 1/(1-\chi(p)p^{-s}) = \Pi_p \Pi_\chi [1/(1-\chi(p)p^{-s})]$.
Sub Lemma: $\Pi_w (1-wT) = (1-T^n)$, where the product is over all nth roots of 1, i.e. over all $w \in U_n$.

Hence $\zeta_m(s) = \Pi_p \Pi_\chi [1/(1-\chi(p)p^{-s})] = \Pi_{p \nmid m} 1/(1-p^{-f(p)s})^{g(p)}$, by the sublemma, where $f(p)g(p) = \varphi(m)$. Hence if we expand these factors as geometric series, and multiply out, we get an ordinary Dirichlet series, with positive integer coefficients, converging at least in $Re(s) > 1$. However this series can be explicitly seen to dominate the series $\Sigma' n^{-\varphi(m)s}$, summed over those n with $\gcd(n,m) = 1$, which in turn dominates $\Sigma_{p \nmid m} p^{-\varphi(m)s}$, which we know diverges at $s = 1/(\varphi(m)$ **QED**

Now we are ready to prove the main result:
Step Four: Proof that $L_\chi(1) \neq 0$, when $\chi \neq 1$.
Lemma: $\zeta(s) = 1/(s-1) + \psi(s)$, in $Re(s) > 0$, where ψ is holomorphic
proof: This is Prop. 10 in Serre.

Corollary: $L_\chi(1) \neq 0$, when $\chi \neq 1$.
proof: Otherwise, since we know every L_χ except L_1 is holomorphic for $Re(s) > 0$, and that L_1 has, like $\zeta(s)$, a simple pole at $s = 1$ (by the Lemma just above), the function $\zeta_m(s)$ would be holomorphic for $Re(s) > 0$. Since the Dirichlet series for $\zeta_m(s)$ has real non negative coefficients, it would follow that the series itself must converge for $Re(s) > 0$, by Prop 7, but we have just shown this series to diverge at $1/\varphi(m) > 0$. **QED**