Mat-1.2991 Discrete mathematics
Appendix II

G. Gripenberg

Aalto University

September 30, 2013

Normal subgroups and quotients

Suppose that G is a group and H is a subgroup of G.

- *$aH = Ha$ for all $a \in G$ if and only if $axa^{-1} \in H$ for all $a \in G$ and $x \in H$.*
  Why? Assume that $a \in G$ and $x \in H$. Now $ax \in aH$ so if $aH = Ha$ there is a $y \in H$ such that $ax = ya$. But then $axa^{-1} = y \in H$. If on the other hand $axa^{-1} = y \in H$ then $ax = ya$ so that $aH \subset Ha$. But we can consider a^{-1} instead of a and get $a^{-1}H \subset Ha^{-1}$ from which it follows that $Ha = aa^{-1}Ha \subset aHa^{-1}a = aH$ as well.
- *If $aH = Ha$ for all $a \in G$ then $a_1H = a_2H$ and $b_1H = b_2H$ implies that $a_1b_1H = a_2b_2H$ which means that one can define the product of the cosets aH and bH by $(aH)(bH) = abH$.*
  Why? Using the assumption several times we get

$$a_1b_1H = a_1Hb_1 = a_2Hb_1 = a_2b_1H = a_2b_2H.$$

Why is $|Gx| \cdot |G_x| = |G|$?

Assume that G is a finite group. If H is a subgroup of G then $|H| \cdot m = |G|$ where m is the number of cosets of H. Since G_x is a subgroup it thus suffices to construct a bijection ψ from the set of cosets of G_x to the orbit Gx.
Define $\psi(gG_x) = gx$. If $g_1G_x = g_2G_x$ we have $g_2^{-1}g_1 \in G_x$ so $g_2^{-1}g_1x = x$ and hence $g_1x = g_2x$ so ψ is well defined
If $g_1x = g_2x$ we have $g_2^{-1}g_1x = x$ so that $g_2^{-1}g_1 \in G_x$ and therefore $g_1G_x = g_2G_x$ so that ψ is an injection. If $y \in Gx$ there is a $g \in G$ so that $y = gx$ and hence $y = \psi(gG_x)$ and hence ψ is a surjection.

Why is the number of orbits in the action of a group $\frac{1}{|G|}\sum_{g\in G}|X_g|$?

Let $E = \{(g,x) \in G \times X : gx = x\}$. Then we get by interchanging the order of summation

$$|E| = \sum_{g\in G}|\{x \in X : gx = x\}| = \sum_{x\in X}|\{g \in G : gx = x\}|,$$

so that $\sum_{g\in G}|X_g| = \sum_{x\in X}|G_x|$.
Let X/G denote the set of orbits, which are equivalence classes under the relation $x \sim y$ if and only if $x = gy$ for some $g \in G$. Thus the orbits are disjoint and their union is X. Since $|G_x| = \frac{|G|}{|Gx|}$ and Gx is the orbit containing x we get the claim from the following calculation:

$$\sum_{g\in G}|X_g| = \sum_{x\in X}|G_x| = \sum_{A\in X/G}\sum_{x\in A}\frac{|G|}{|Gx|} = |G|\sum_{A\in X/G}\sum_{x\in A}\frac{1}{|A|}$$
$$= |G|\sum_{A\in X/G}\frac{1}{|A|}\sum_{x\in A}1 = |G|\sum_{A\in X/G}1 = |G||X/G|.$$

Proof of Pólya's theorem on the number of colorings

Assume that Ω is a set of colorings of X invariant under G. Then the number of orbits of the action of G on Ω is $\frac{1}{|G|}\sum_{g\in G}|\Omega_g|$ where $\Omega_g = \{\omega \in \Omega : g\omega = \omega\}$ is the set of colorings that are invariant under g, and these in turn are the ones that are constant on each orbit of the permutation g of X, or equivalently on each orbit in the action of the cyclic group generated by g on the set X. Thus we can consider each $g \in G$ separately and then add up.
Suppose that $A_{g,1}, A_{g,2}, \ldots A_{g,m_g}$ are the orbits of g with $s_j = |A_{g,j}|$. Now there is of course one way of using color a_j exactly s_1 times to give the points in $A_{g,1}$ the color a_j. This can be described by the generating function $a_1^{s_1} + \ldots + a_r^{s_1}$. Assume now that $p(a_1, \ldots, a_r)$ is a generating functions such that coefficient of the monomial $a_1^{i_1} \cdot a_2^{i_2} \cdot \ldots \cdot a_r^{i_r}$ is the number of ways in which the sets $A_{g,1}, \ldots A_{g,k}$ can be colored using color a_j exactly i_j times. The elements in the set $A_{g,l+1}$ can be colored using one of the colors s_k times and all different colors give rise to different cases. If we use color a_q and want to use color a_j exactly i_j times in coloring all

Proof of Pólya's theorem on the number of colorings, cont.

the sets $A_{g,1}, \ldots A_{g,k}, A_{g,k+1}$ then we must use color a_j exactly i_j times when $j \neq q$ and color a_q exactly $i_q - s_q$ times to color the first sets $A_{g,1}, \ldots A_{g,k}$. But this means that the generating function for coloring the sets $A_{g,1}, \ldots A_{g,k}, A_{g,k+1}$ is $p(a_1, \ldots a_r) \cdot (a_1^{s_{k+1}} + \ldots + a_r^{s_{k+1}})$. Thus the induction step works and we see that
$\zeta_{g,X}(a_1^1 + \ldots + a_r^1, a_1^2 + \ldots + a_r^2, \ldots, a_1^n + \ldots + a_r^n)$ is the generating function for coloring the orbits of g. In other words the coefficient of $a_1^{i_1} \cdot a_2^{i_2} \cdot \ldots \cdot a_r^{i_r}$ in the generating function
$\zeta_{g,X}(a_1^1 + \ldots + a_r^1, a_1^2 + \ldots + a_r^2, \ldots, a_1^n + \ldots + a_r^n)$ is the number of orbits in the action of g on the coloring of X using color a_j exactly $i - j$ times and then we can take the sum over $g \in G$ and divide by $|G|$ to get the number of orbits wwhen G acts on this set of colorings.
For the case where the set of colorings are all the ones using r colors, that is all functions from X to a set with r elements we proceed in the same way and note that if g has k orbits these can be colored in r^k different ways when all points in an orbit get the same color. But if g has k orbits then $\zeta_{g,X}(r, \ldots r) = r^k$ and the result follows.

The orbit or cycle index of rotations

The permutation $p = (1\ 2\ \ldots\ n)$ of the set $\mathbb{N}_n = \{1, 2, \ldots, n\}$ generates a cyclic group (denoted by C_n) with elements $I, p, p^2, \ldots, p^{n-1}$ where I is the identity element. This permutation p corresponds to a rotation of a regular n-gon by $\frac{2\pi}{n}$.
The order of the element p^k, $k = 0, 1, 2 \ldots, n-1$ is the smallest integer $d \geq 1$ so that $p^{d \cdot k} = I$ or, equivalently, $d \cdot k = n \cdot j$ for some integer j. Since $k \leq n - 1$ we must have $0 \leq j \leq d - 1$. If $\gcd(j, d) > 1$ we could make d smaller so we must have $\gcd(d, j) = 1$. But this implies that $d \mid n$ and we have $k = \frac{n}{d} \cdot j$ with $\gcd(j, d) = 1$ and $0 \leq j \leq d - 1$. Conversely we see that for every $d \mid n$ and $0 \leq j \leq d - 1$ with $\gcd(j, d) = 1$ we find an integer $k \in \{0, 1, \ldots, n - 1\}$ so that the order of p^k is d.
Thus there are $\varphi(d)$ elements of order d for each $d \mid n$. A permutation p^k of order d acting on the set $\{1, 2 \ldots, \}$ has $\frac{n}{d}$ orbits (since the identity leaves every point unchanged and all other elements in the cyclic group generated by p^k moves every element in the set) and thus the orbit index is

$$\zeta_{C_n, \mathbb{N}_n}(t_1, t_2, \ldots, t_n) = \frac{1}{n} \sum_{d \mid n} \varphi(d) t_d^{\frac{n}{d}}.$$

Why is $F(\alpha)$ isomorphic to $F[x]/\langle h\rangle$ when h is irreducible and $h(\alpha) = 0$?

For each equivalence class $[a] \in F[x]/\langle h\rangle$ we let $\psi(a) = a(\alpha)$. Since $[a] = [b]$ iff $a = b + c \cdot h$ and $h(\alpha) = 0$ this function ψ is well defined. It is (?) clear that $\psi([a] + [b]) = \psi([a]) + \psi([b])$, $\psi([-a]) = -\psi([a])$, and $\psi([a] \cdot [b]) = \psi([a]) \cdot \psi([b])$. Furthermore if $[a] \neq 0$ then there are polynomials b and c so that $a \cdot b = 1 + c \cdot h$ (that is, $[b] = [a]^{-1}$) and this implies that $\psi([a]) \cdot \psi([b]) = 1$ and hence $\psi([b]) = \psi([a])^{-1}$ when $\psi([a]) \neq 0$. This implies that $\{\, \psi([a]) : [a] \in F[x]/\langle h\rangle \,\}$ is a subfield of the extension field K of F. Since $[x] \in F[x]/\langle h\rangle$ this field contains α and hence $F(\alpha)$ as well. On the other hand it follows from the fact that $F(\alpha)$ is a field and $\alpha \in F(\alpha)$ that every $\psi([a]) \in F(\alpha)$. Thus ψ is an isomorphism: $F[x]/\langle h\rangle \to F(\alpha)$.

Why is $[F[x]/\langle h\rangle : F] = \deg(h)$

In order to be able to consider F as a subfield of $F[x]/\langle h\rangle$ we have to identify F with polynomials of degree ≤ 0 and instead of equivalence classes take the elements in $F[x]/\langle h\rangle$ to be polynomials in $F[x]$ with degree at most $m - 1$ where $\deg(h) = m$. This implies that the "vectors" $1, x, x^2, \ldots, x^{m-1}$ span the vector space $F[x]/\langle h\rangle$ over F. In addition these vectors are linearly independent because if $f_0\, 1 + f_1 x + \ldots f_{m-1} x^{m-1} = 0$ then by the definition of polynomials as functions from $\{0, 1, 2 \ldots\}$ to F we have $(f_0, f_1, \ldots, f_{m-1}, 0, \ldots) = (0, 0, \ldots)$ and hence $f_0 = f_1 = \ldots = f_{m-1} = 0$. Thus the dimension of $F[x]/\langle h\rangle$ over F is m since the dimension can be characterized as the number of linearly independet vectors that spans the space. (The fact that the field is F and not e.g. $\mathbb{R}$ does not matter.)

Why is $[F(\alpha) : F] = \deg(h)$ if h is irreducible and $h(\alpha) = 0$?

Since there is an isomorphism $\psi : F[x]/\langle h \rangle \to F(\alpha)$ one also gets an isomorphism of these sets when they are considered as vector spaces over F. Since isomorphic vector spaves have the same dimension the claim follows from the fact that $[F[x]/\langle h \rangle : F] = \deg(h)$.

Why is $[K : F] = [K : E][E : F]$ when E is an extension of F and K is an extension of E?

Let $m = [K : E]$ and $n = [E : F]$. It follows from the definition that there are $k_1, \dots, k_m \in K$ which form a basis for K as a vector space over E and hence each $k \in K$ can be written in a unique way as $k = \sum_{i=1}^m c_i k_i$ where $c_i \in E$ for $i = 1, \dots, m$. Similarly there are $e_1, \dots e_n \in E$ that form a basis for E as a vector space over F. Since each $s_i \in E$ we get $s_i = \sum_{j=1}^n s_{i,j} e_j$ which in turn implies that $k = \sum_{i=1}^m \sum_{j=1}^n s_{i,j} k_i e_j$. This shows that that the vectors $k_i e_j \in K$ span K as a vector space over F and since there are $m \cdot n$ of these vectors it remains to show that they are linearly independent. If $\sum_{i=1}^m \sum_{j=1}^n s_{i,j} k_i e_j = 0$ then $\sum_{i=1}^m \left(\sum_{j=1}^n s_{i,j} e_j \right) k_i = 0$. Now the vectors k_i are linearly independent so $\sum_{j=1}^n s_{i,j} e_j = 0$ for each $i = 1, \dots, m$. But the vectors e_j are linearly independet as well so $s_{i,j} = 0$ for each $j = 1, \dots, n$ and for each $j = 1, \dots, m$. Thus the $m \cdot n$ vectors $k_i e_j$, $i = 1, \dots, m$, $j = 1, \dots, n$ form a basis for K as a vector space over F.

How can one, in principle, find the generator of a cyclic code?

Assume first that there are at least two codewords in C.

- *Let* $m = \min\{\deg(\mathbf{a}) : \mathbf{a} \neq 0,\ \mathbf{a} \in C_p\}$ *and let* $\mathbf{g} \in C_p$ *be such that* $\deg(\mathbf{g}) = m$. *If* $\tilde{\mathbf{g}}$ *is another such polynomial, then* $\tilde{\mathbf{g}} + \mathbf{g} \in C_p$ *since the code is linear and* $\deg(\tilde{\mathbf{g}} + \mathbf{g}) < m$ *because* $1 + 1 = 0$ *in* F_2 *and the coefficient of* x^m *is* 1 *in both* $\tilde{\mathbf{g}}$ *and* $\mathbf{g}$. *Thus we must have* $\tilde{\mathbf{g}} = \mathbf{g}$, *that is* $\mathbf{g}$ *is unique.*
- *Since the code is cyclic it follows that* $x^j \cdot \mathbf{g} \in C_p$ *for all* $0 \leq j \leq n - m - 1$, *and since cyclicity includes linearity we have* $\mathbf{a} \cdot \mathbf{g}$ *for all* $\mathbf{a} \in F[x]$ *with* $\deg(\mathbf{a}) \leq n - m - 1$.
- *If* $\mathbf{c} \in C_p$ *then* $\mathbf{c} = \mathbf{a} \cdot \mathbf{g} + \mathbf{r}$ *with* $\deg(\mathbf{a}) \leq n - m - 1$ *and* $\deg(\mathbf{r}) < m$. *By the previous result and the linearity of the code we have* $\mathbf{c} - \mathbf{a} \cdot \mathbf{g} \in C_p$ *so we must have* $\mathbf{r} = 0$ *by the definition of* m *and this shows that* C *is generated by* $\mathbf{g}$.

If the only element in C_p *is* 0, *then one can can choose* $\mathbf{g} = x^n - 1$

Why is $\mathbf{g}$ the generator of a cyclic code if and only if $\mathbf{g}$ divides $x^n - 1$?

- *If* $\deg(\mathbf{g}) = n$ *then* $\mathbf{g}$ *divides* $x^n - 1$ *if and only if* $\mathbf{g} = x^n - 1$ *which happens if and only if* $C = \{00\cdots0\}$. *Thus we may assume that* $0 \leq \deg(\mathbf{g}) < n$.
- *By definition the code* C *is cyclic if and only if it is linear and the remainder when* $x^j \cdot \mathbf{c}$ *is divided by* $x^n - 1$ *is in* C_p *for every* $\mathbf{c} \in C_p$, *or equivalently if and only if it is linear and the remainder when* $\mathbf{a} \cdot \mathbf{c}$ *divided by* $x^n - 1$ *is in* C_p *for every* $\mathbf{a} \in F_2[x]$ *and* $\mathbf{c} \in C_p$.
- *Assume that* g *is the generator of a cyclic code. Then* $x^n - 1 = \mathbf{h} \cdot \mathbf{g} + \mathbf{r}$ *where* $\deg(\mathbf{r}) < \deg(\mathbf{g})$. *Thus the remainder when* $\mathbf{h} \cdot \mathbf{g}$ *is divided by* $x^n - 1$ *is* $\mathbf{r}$ *which thus belongs to* C_p *and this contradicts the construction of* $\mathbf{g}$ *unless* $\mathbf{r} = 0$ *in which case* $\mathbf{g}$ *divides* $x^n - 1$.

Why is $\mathbf{g}$ the generator of a cyclic code if and only if $\mathbf{g}$ divides $x^n - 1$? Cont.

- *Assume that* $\mathbf{g}$ *divides* $x^n - 1$ *and* $\deg(\mathbf{g}) = n - k$. *Then there is a polynomial* $\mathbf{h}$ *with* $\deg(\mathbf{h}) = k$ *so that* $\mathbf{h} \cdot \mathbf{g} = x^n - 1$. *Furthermore, we note that the set* $C_p = \{ \mathbf{a} \cdot \mathbf{g} : \mathbf{a} \in F_2[x],\ \deg(\mathbf{a}) \leq k - 1 \}$ *is such that if* $\mathbf{c}_1$ *and* $\mathbf{c}_2 \in C_p$ *then* $c_1 + c_2 \in C_p$, *that is the code* C *generated by* $\mathbf{g}$ *is linear. If now* $\mathbf{c} \in C_p$ *then* $\mathbf{c} = \mathbf{a} \cdot \mathbf{g}$ *and the remainder when* $x^j \cdot \mathbf{c}$ *is divided by* $x^n - 1$ *is* $\mathbf{r} = x^j \cdot \mathbf{a} \cdot \mathbf{g} - \mathbf{q} \cdot (x^n - 1) = (x^j \cdot \mathbf{a} - \mathbf{q} \cdot \mathbf{h}) \cdot \mathbf{g}$ *and since* $\deg(\mathbf{r}) \leq n - 1$ *we have* $\deg(x^j \cdot \mathbf{a} - \mathbf{q} \cdot \mathbf{h}) \leq k - 1$ *which implies that* $\mathbf{r} \in C_p$ *and hence the code* C *is cyclic.*

The check matrix and the code generator matrix for a cyclic code

(a) *Do the codewords* $c = aG$ *satisfy* $Hc^{\mathsf{T}} = 0$?

(b) *Are there other solutions to* $Hc^{\mathsf{T}} = 0$ *than the vectors* $c = aG$?

Answers:

(a) *Let* $h_j = 0$ *when* $j = k + 1 \ldots n - 1$ *and* $h_{j+n} = h_j$ *for all* $j \in \mathbb{Z}$. *Similarly let* $g_j = 0$ *for* $j = n - k + 1, \ldots, n - 1$ *and* $g_{j+n} = g_j$ *for all* $j \in \mathbb{Z}$. *Thus* $H(i,j) = h_{j-i}$ *and* $G(i,j) = g_{n-k+i-j}$. *Now* $HG^{\mathsf{T}} = 0$ *because* $(HG^{\mathsf{T}})(i,j) = \sum_{m=1}^{n} h_{m-i} g_{n-k+j-m}$ *which is the coefficient for* $x^{n-k+j-i}$ *in* $\mathbf{h} \cdot \mathbf{g}$ *and hence zero because* $1 \leq n - k + j - i \leq n - 1$ *when* $i = 1, \ldots, n - k$ *and* $j = 1, \ldots, k$. *Thus the codewords* $c = aG$ *satisfy* $Hc^{\mathsf{T}} = 0$.

(b) *Since* $\mathbf{h} \cdot \mathbf{g} = x^n - 1$ *we must have* $h_0 = g_0 = 1$ *so* H *is in row echelon form with pivot elements in the first* $n - k$ *columns, hence the dimension of tke kernel of* H *is* k *and hence there are* 2^k *solutions to the equation* $Hc^{\mathsf{T}} = 0$. *But* G *is also in row echelon form so the mapping* $a \mapsto aG$ *is an injection and there are* 2^k *vectors of the form* aG. *Thus there are no other solutions than the vectors* $c = aG$.

An auxilary result

Assume that F is a field, $n > 1$, $u \in F \setminus \{0\}$ is such that $u^j \neq 1$ when $1 \leq j < n$, $\mathbf{c} \in F[x] \setminus \{0\}$ with $\deg(\mathbf{c}) < n$, and $\mathbf{c}(u^m) = 0$ when $m = m_0 + 1, \ldots, m_0 + s$. Then $\mathbf{c}$ has at least $s + 1$ nonzero coefficients.

Proof

Suppose that $\mathbf{c}$ has at most s nonzero coefficients, so that $\mathbf{c} = \sum_{j=1}^{s} a_j x^{b_j}$ with $0 \leq b_1 < \cdots < b_s < n$. Thus we have the system of equations $\sum_{j=1}^{s} M_{i,j} a_j = 0$, $i = 1, 2 \ldots, s$ where $M_{i,j} = u^{(k_o+i)b_j} = u^{(m_0+1)b_j}(u^{b_j})^{i-1}$. Now the matrix $V_{i,j} = (u^{b_j})^{i-1}$ is a Vandermonde matrix and $\det(M) = u^{(m_0+1)\sum_{j=1}^{s} b_j} \prod_{1 \leq j < k \leq s}(u^{b_j} - u^{b_k})$. It follows from $0 \leq b_j < b_k < n$ and the assumption $u^j \neq 1$ when $1 \leq j < n$ that $\det(M) \neq 0$ and thus $a_j = 0$ for $j = 0, 1, \ldots, s$ which is a contradiction since $\mathbf{c} \neq 0$.