Frobenius's theorem on Frobenius kernels

Ravi Fernando – fernando@berkeley.edu

March 12, 2016

This is a proof of Frobenius's classical theorem that the kernel of a Frobenius group is a normal subgroup. I followed Christopher Cooper's handout[1], but my exposition is intended to be more relaxed and understandable with fewer prerequisites. For this purpose, I will black-box the definition and formula for the induced character ψ^G in the proof.

Definition 1. *Let $G \subseteq S_n$ be a finite permutation group. We say G is a* Frobenius group *if G acts transitively on the points $1, \ldots, n$ and no non-identity element of G fixes more than one point. If G is a Frobenius group, then we say its* (Frobenius) kernel *is the subset consisting of all fixed-point-free elements, together with the identity.*

Example 2. *The symmetric group S_3 is a Frobenius group because if any element fixes two points, it must fix the third as well. Its kernel is the alternating group A_3.*

Example 3. *Similarly, the alternating group A_4 is a Frobenius group, since it doesn't contain single transpositions. Here, the Frobenius kernel is the Klein four-group, consisting of the double transpositions $(12)(34), (13)(24)$, and $(14)(23)$, as well as the identity.*

Example 4. *For a more complicated example, let G be the set of permutations of the finite field $\mathbb{F}_p$ of the form $x \mapsto ax + b$, where $a, b \in \mathbb{F}_p$ and $a \neq 0$. This is a group under composition, whose order is $p(p-1)$. To prove that G is a Frobenius group, suppose the map $x \mapsto ax + b$ fixes two points $c \neq d$ in $\mathbb{F}_p$. Then the equations*

$$c = ac + b \tag{1}$$

$$d = ad + b \tag{2}$$

imply that $(1-a)c = b = (1-a)d$, which forces $a = 1$ (because $c \neq d$) and then $b = 0$ (because $c = c + b$). The Frobenius kernel here consists of the translations $x \mapsto x + b$, because if $a \neq 1$, then the map $x \mapsto ax + b$ has the fixed point $\frac{b}{1-a} \in \mathbb{F}_p$.

In all of these examples, the Frobenius kernel turned out to be a (normal) subgroup of the Frobenius group. This isn't at all obvious in general, but we claim that it is true. The method of proof is remarkable: roughly speaking, we show that there is a representation whose kernel is the Frobenius kernel, but we are only really able to find the character of the representation (by adding and subtracting some other characters), and we have no idea how to construct the representation itself. Also remarkable: even a century after its discovery, this theorem still has no known purely group-theoretical proof.[2]

[1] http://web.science.mq.edu.au/∼chris/represent/CHAP08%20Frobenius%20Groups.pdf

[2] See discussion: http://mathoverflow.net/questions/63142/character-free-proof-that-frobenius-kernel-is-a-normal-subgroup

Theorem 5. *The Frobenius kernel K of any Frobenius group G is a normal subgroup of G.*

Notice that the condition of normality is not hard to check, if we somehow knew that it is a subgroup. To prove this, note that an element $g \in G$ fixes a point $i \in \{1, \dots, n\}$ if and only if its conjugate hgh^{-1} fixes $h \cdot i$. So conjugate elements of G must fix the same number of points, which implies that K is invariant under conjugation.

Let's start by thinking about what the elements of G look like. Recall that all non-identity elements of G fix either one point or no points. Let X be the subset of G consisting of the former type, and Y the subset consisting of the latter. (So we are claiming that $K = \{1\} \cup Y$ is a normal subgroup of G.) Since conjugation preserves number of fixed points, elements in X are not conjugate to those in Y; that is, each of X and Y is a union of some of the conjugacy classes of G.

Now let H be the stabilizer of the point 1. Let's say the conjugacy classes of H are $\Gamma_1, \dots, \Gamma_k$, where $\Gamma_1 = \{1\}$, and let h_i be the size of Γ_i. The following description of the conjugacy classes of G contained in X will be useful when applying the induced character formula in Lemma 7.

Lemma 6. *There are exactly $k-1$ conjugacy classes $\Omega_2, \dots, \Omega_k$ of G contained in X, where each Ω_i $(1 < i \leq k)$ consists of Γ_i along its conjugates that fix the points $2, \dots, n$ instead of 1. In particular, Ω_i has size nh_i, so X has size $\sum_{i=2}^{k} nh_i = n\sum_{i=2}^{k} h_i = n(|H|-1) = |G| - n$, and Y has size $n-1$.*

Proof. Let's start by calculating the size of X. Recall that X is the disjoint union of the stabilizers of each of the individual points, with the identity removed: $X = (\mathrm{Stab}(1) \setminus \{1\}) \cup \cdots \cup (\mathrm{Stab}(n) \setminus \{1\})$. Each stabilizer has index n by the orbit-stabilizer theorem, so it has order $\frac{|G|}{n} = h$, giving $|X| = (h-1)n = |G| - n$.

To study the conjugacy classes contained in X, fix some $x \in X$ and let $\mathrm{Cl}_G(x)$ denote the conjugacy class of x in G. First, notice that x has exactly one fixed point; call it i. If $g \in G$ is any permutation that sends i to j (and such an element exists for all i, j, by transitivity), then $g^{-1}xg$ has the fixed point j instead. In fact, $x \mapsto g^{-1}xg$ gives a bijection from elements of G fixing i to elements fixing j. Since all we are doing is conjugating, it's even a bijection from elements of $\mathrm{Cl}_G(x)$ fixing i to those fixing j. So we can now restrict our attention to conjugates of x fixing, say, $j = 1$, and we know that there will be equally many fixing any other point. So let's assume without loss of generality that x itself fixes 1 (i.e. belongs to H), and study the conjugates of x that also lie in H.

We know some conjugates of x that belong to H already: x belongs to one of the conjugacy classes $\Gamma_2, \dots, \Gamma_k$ in H, say Γ_i. Everything in Γ_i is conjugate to x in H (i.e. equals gxg^{-1} for some $g \in H$), so it is still conjugate to x in G. But the reverse isn't true a priori: if two elements are conjugate in G, in principle it could be the case that one can be conjugated to the other only using an element that lies outside of H. But this doesn't happen here; in fact, gxg^{-1} belongs to H if and only if g does. (The previous sentence will be useful later.) The reason is simple: since x fixes only the point 1, gxg^{-1} fixes only the point $g \cdot 1$, and this equals 1 if and only if $g \in H$. So the conjugacy class of x in G is exactly its conjugacy class in H, along with its bijective copies that fix elements other than 1. So the G-conjugacy classes $\Omega_2, \dots, \Omega_k$

in X correspond to the non-identity conjugacy classes $\Gamma_2, \ldots, \Gamma_k$ of H, and each Ω_i is exactly n times larger than Γ_i. □

Now let's think about characters. Let θ be the permutation character of G, coming from the representation ρ_θ on the vector space $\mathbb{C}^n$ given by permuting coordinates. Its values $\theta(g)$ are given by the number of fixed points of the permutation g. Recall that the permutation representation ρ_θ decomposes into two pieces: a trivial representation consisting of vectors with all their coordinates equal, and a complement consisting of vectors whose coordinates sum to 0. The character of the former is the trivial character χ_1, and the character of the latter is $\theta - \chi_1$.

Our main tool will be the following.

Lemma 7. *Fix any irreducible character ψ of H, with degree m. Let $\psi^G = \mathrm{Ind}_H^G \psi$ be the corresponding induced character of G. Then the function $\psi^* = \psi^G - m(\theta - \chi_1)$ is an irreducible character of G.*

Proof. We will take a few facts for granted about the induced character ψ^G. Given a character (or representation) of a subgroup $H \leq G$, induction is a natural way to produce a character (or representation) of the full group G. (We are not generally guaranteed anything about whether an induced representation is irreducible.) Inducing from H to G always multiplies the degree of a character by the index $[G : H]$, which in this case is n. Induced representations are best defined using tensor products of modules over group algebras, so we'll black-box this, but there is actually a very explicit formula for an induced character: $\mathrm{Ind}_H^G(\chi)(s) = \frac{1}{|H|} \sum_{t \in G, t^{-1}st \in H} \chi(t^{-1}st)$, where the sum ranges over whichever choices of $t \in G$ satisfy the condition $t^{-1}st \in H$.

The induced character formula tells us the following: $\psi^G(1) = mn$ (this is the earlier-mentioned fact about the degree of an induced character), ψ^G of anything in Y is 0 (there are no $t \in G$ with $t^{-1}st \in H$), and finally $\psi^G(\Omega_i) = \psi(\Gamma_i)$ for $1 < i \leq k$. Explanation for the last one: choose s to be a representative of the class Ω_i, and as in Lemma 6 choose it to be in H. Then, by an observation from the proof of that lemma (with $s = x, t = g^{-1}$), the only t's that count are $t \in H$, and all of the $\chi(t^{-1}st)$ are equal to $\chi(s)$. So we get $\psi^G(s) = \frac{1}{|H|} \sum_{t \in H} \chi(s) = \frac{|H|}{|H|} \psi(s) = \psi(s)$.

The following table of values summarizes the characters we've discussed:

	1	Ω_2	$\cdots$	Ω_k	Y		
χ_1	1	1	$\cdots$	1	1	$\cdots$	1
θ	n	1	$\cdots$	1	0	$\cdots$	0
$\theta - \chi_1$	$n-1$	0	$\cdots$	0	-1	$\cdots$	-1
ψ^G	mn	$\psi(\Gamma_2)$	$\cdots$	$\psi(\Gamma_k)$	0	$\cdots$	0
ψ^*	m	$\psi(\Gamma_2)$	$\cdots$	$\psi(\Gamma_k)$	m	$\cdots$	m

So far, all we know about $\psi^* = \psi^G - m(\theta - \chi_1)$ is that it is a function from G to $\mathbb{C}$ that happens to equal a linear combination of irreducible characters with integer coefficients, $\sum_i a_i \chi_i$. But since the inner product of any irreducible character with itself is 1, and the inner product of two different irreducibles is 0, we can calculate $\sum_i a_i^2$ by taking $\langle \psi^*, \psi^* \rangle$:

$$\begin{aligned}
\langle \psi^*, \psi^* \rangle &= \frac{1}{|G|} \sum_{g \in G} |\psi^*(g)|^2 && (3) \\
&= \frac{1}{|G|} \left(m^2 + \sum_{j=2}^{k} \left(nh_j \cdot |\psi(\Gamma_j)|^2 \right) + (n-1)m^2 \right) && (4) \\
&= \frac{1}{|G|} \left(nm^2 + \sum_{j=2}^{k} nh_j \cdot |\psi(\Gamma_j)|^2 \right) && (5) \\
&= \frac{n}{|G|} \left(m^2 + \sum_{j=1}^{k} h_j \cdot |\psi(\Gamma_j)|^2 - 1 \cdot m^2 \right) && (6) \\
&= \frac{1}{|H|} \sum_{j=1}^{k} h_j \cdot |\psi(\Gamma_j)|^2 && (7) \\
&= \langle \psi, \psi \rangle = 1, && (8)
\end{aligned}$$

because we chose ψ to be an irreducible character of H. So if ψ^* is a linear combination of irreducible characters of G with coefficients $a_i \in \mathbb{Z}$, then $\sum a_i^2 = 1$, which means it must be a single irreducible character with coefficient ±1. But ψ^* cannot be the negative of a character, because $\psi^*(1) = m$ is positive. So ψ^* must actually be the character of an irreducible representation, proving the lemma. □

We are now ready to finish proving the theorem. Notice that the character ψ^* in the lemma is a degree-m character of G that attains the value m on every element of Y. In particular, if $\rho : G \to \mathrm{GL}_m(\mathbb{C})$ is the representation whose trace equals ψ^*, and $y \in Y$, then $\rho(y)$ is an $m \times m$ matrix, which is diagonalizable, and whose eigenvalues are m roots of unity adding up to m. It follows that all the eigenvalues are 1, so $\rho(y) = I$; that is, Y is contained in the kernel of ρ.

We would like $K = \{1\} \cup Y$ to be exactly the kernel of ρ, since kernels must be normal subgroups. This is equivalent to saying that ψ^* does not attain the value m outside of K. Unfortunately, this isn't necessarily true. What is true, however, is that if ψ ranges over all irreducible characters of H, the intersection of the kernels of the representations $\rho = \rho_{\psi^*}$ is exactly K. To prove this, suppose that some $x \in X$ lies in this intersection, so that every possible $\psi^*(x) = m$. (Note that the degree m may change from one ψ to another.) Since every $x \in X$ is conjugate to an element of H, we may assume $x \in X \cap H$ without loss of generality. But ψ^* agrees with ψ itself on elements of $H \leq G$, so it follows that $\psi(x) = m = \psi(1)$ for all irreducible characters ψ of H. But this implies that the 1 and x columns of the character table of H are equal to each other, and in particular the columns of the character table are linearly dependent. This is impossible, because the character table is a square matrix with linearly independent rows, so it must have linearly independent columns as well. So by contradiction, we see that K is the intersection of the kernels of the representations ρ, which is a normal subgroup of G, as claimed.