READING BURNSIDE

E. KOWALSKI

In [1], W. Burnside proves what is indeed commonly known as *Burnside's Theorem* (except when that term is reserved to another of his results, most often the solvability of groups of order p^aq^b, where p, q are primes):

Theorem 1 (Burnside, 1905)**.** *Let k be an algebraically closed field, let G be a subgroup of $\mathrm{GL}_n(k)$ for some $n \geqslant 1$, and let $k[G]$ be the linear subspace of $\mathrm{End}(k^n)$ spanned by the elements of G. Then G is* irreducible*, i.e., there is no linear subspace $W \subset k^n$, $0 \neq W$, $W \neq k^n$, such that G leaves W invariant, if and only if, $k[G] = \mathrm{End}(k^n)$.*

In this note, we first simply (try to) interpret Burnside's original proof in modern terminology and notation. This is instructive for anyone interested in the evolution of mathematical writing, as Burnside's language is – at least for the author – essentially unreadable *as a rigorous mathematical proof*. One can make sense of it, with some effort, and from it construct a proof which is (or at least, seems to be) equivalent to Burnside's. (I have no doubt that his proof is correct in his terms, despite not being able to understand some important parts of it; and certainly he would be able to explain those in person...)

We then show how to extend the argument to non-irreducible representations. This leads naturally to such results as the linear independence of matrix coefficients and characters of finite-dimensional representations of a group.

1. Burnside's proof, modernized

Our reading of Burnside's proof is the following. Let $V = k^n$, and let G be a subgroup $G \subset \mathrm{GL}(V)$. We denote $E = \mathrm{End}(V)$ the ring of linear endomorphisms of V, and then we let

$$R = \{\phi \in E' = \mathrm{Hom}(E, k) \mid \phi(g) = 0 \text{ for all } g \in G\},$$

the space of linear relations satisfied by the coordinates of all elements of g. By linearity, we see that Burnside's theorem is equivalent with the statement that $R = 0$ if and only if G acts irreducibly on V. One direction is very easy: if $W \subset V$ is a non-trivial proper subspace stable under G, $\lambda \neq 0$ is an element of V' such that $W \subset \ker \lambda$ and $e \neq 0$ is a vector in W, then the linear map defined by

$$\phi(A) = \lambda(Ae)$$

(for $A \in E$) is non-zero (since one can find A mapping e to any vector, in particular one not in $\ker \lambda$) and belongs to R (since $ge \in W \subset \ker \lambda$ for $g \in G$). In concrete terms, we can think of looking at the matrices representing G in a basis starting with a basis of W, and observing that "the last coefficient of the first column" is always zero for the block-triangular matrices representing G.

To prove the converse, the idea is to exploit representation theory. Roughly speaking, we use the following steps, which show that the strategy is natural and easy to remember:

(1) Show that R is in a natural way a subrepresentation for some action of G on E';

(2) Find that all irreducible subrepresentations of E' are in fact isomorphic to (the contragredient of) the action of G on V, embedded in a specific manner;
(3) Show using this description that no non-zero irreducible subrepresentation of E' can be contained in R, if G acts irreducibly.

To implement these steps, we first introduce the representation

$$\rho : G \to \mathrm{GL}(E)$$

such that

$$\rho(g)A = gA$$

for $g \in G$ and $A \in E$ (note that this is not the usual representation of G on $\mathrm{End}(V)$, which is obtained as the tensor product of the action of G on V with its contragredient). The contragredient representation $\check{\rho}$ acts on E' by

$$\check{\rho}(g)\phi = \Big(A \mapsto \phi(\rho(g^{-1})A)\Big)$$

(or

$$\langle \check{\rho}(g)\phi, A\rangle = \langle \phi, \rho(g^{-1})A\rangle = \langle \phi, g^{-1}A\rangle$$

in duality-bracket notation).

Now the first main point is that $R \subset E'$ is a subrepresentation of $\check{\rho}$: indeed, if $g \in G$, $\phi \in R$, and $\psi = \check{\rho}(g)\phi$, then we have

$$\psi(h) = \phi(g^{-1}h) = 0$$

for all $h \in G$ (since G is a group!), and this means that $\check{\rho}(g)\phi$ is also in R.

So the principle is now that we try to write down the decomposition of R as a representation of G; Burnside's Theorem is the fact that $R = 0$ if G acts irreducibly, but if we lived in a world where Burnside's theorem does not hold, or had not yet been proved, the determination of R would be a very natural question. Indeed, in Section 3, we show how a very similar argument leads to results about different representations of a fixed group.

First, some subrepresentations of E' turn out to be easy to find. Namely, for any fixed vector $v \neq 0$ in V, we can consider the linear map defined by

$$\alpha_v : \begin{cases} V' & \to & E' \\ \lambda & \mapsto & (A \mapsto \lambda(Av)), \end{cases}$$

(or in other words

$$\langle \alpha_v(\lambda), A\rangle_E = \langle \lambda, Av\rangle_V,$$

where we use subscripts to indicate for which space the duality bracket is used.)

This linear map is injective (since $\alpha_v(\lambda) = 0$ implies that all vectors Av are in $\ker \lambda$, and this is only possible if $\lambda = 0$) and is an intertwining operator, where G acts on V' through $\check{\pi}$ and on E' through $\check{\rho}$. Indeed, for $\lambda \in V'$, both $(\alpha_v \circ \check{\pi}(g))(\lambda)$ and $(\check{\rho}(g) \circ \alpha_v)(\lambda)$ are the linear maps

$$A \mapsto \lambda(g^{-1}Av)$$

in E'; abbreviating the G-actions, this follows easily by

$$\begin{aligned} \langle \alpha_v(g \cdot \lambda), A\rangle &= \langle g \cdot \lambda, Av\rangle = \langle \lambda, g^{-1}Av\rangle \\ &= \langle \lambda, (g^{-1}A)v\rangle = \langle \alpha_v(\lambda), g^{-1}A\rangle = \langle g \cdot \alpha_v(\lambda), A\rangle. \end{aligned}$$

So we have an abundance of copies of $\check{\pi}$ inside E', namely the images $\mathrm{Im}(\alpha_v)$ as v varies. The next step is that, in fact, there are no other subrepresentations up to isomorphism, i.e., E' is isomorphic to a direct sum of copies of $\check{\pi}$. Indeed, if we fix a basis

$$(e_1, \ldots, e_n)$$

of V, we have

$$E' = \bigoplus_{1 \leqslant j \leqslant n} \mathrm{Im}(\alpha_{e_j}), \tag{1}$$

which is simply because $\mathrm{Im}(\alpha_{e_j})$ is the subspace of E' consisting of those ϕ for which $\phi(A)$ depends only on the j-th column Ae_j of the matrix representing A in the given basis.

The final step is to show that R, if it is non-zero, must contain a subrepresentation of the type $\mathrm{Im}(\alpha_v) \subset R$ for some $v \neq 0$. This follows if there are no other way to inject $\check{\pi}$ into E' than using α_v for some v, since then any non-zero irreducible subrepresentation of R is of this type. And if that is granted, then we are done, because the assumption that $R \neq 0$ leads to a contradiction: using $v \neq 0$ such that $\mathrm{Im}(\alpha_v) \subset R$, we deduce that all maps

$$A \mapsto \lambda(Av)$$

are in R, where $\lambda \in V'$ is arbitrary. Taking $A = \mathrm{Id} \in G$, we deduce that

$$0 = \langle \alpha_v(\lambda), \mathrm{Id} \rangle = \lambda(v)$$

for all λ. This, of course, contradicts $v \neq 0$.

So we need only check the claim. For this, we must first show that the map $\Xi : v \mapsto \alpha_v$ gives a linear isomorphism $V \to H$, where

$$H = \mathrm{Hom}_G(\check{\pi}, E') = \mathrm{Hom}_G(V', E')$$

is the space of G-equivariant linear maps from $\check{\pi}$ to E'. Indeed, we first note that both spaces have dimension n; in the case of H, this is because of the known decomposition (1) which gives

$$H \simeq \bigoplus_j \mathrm{Hom}_G(\check{\pi}, \mathrm{Im}(\alpha_{e_j})),$$

and Schur's Lemma ensures that $\mathrm{Hom}_G(\check{\pi}, \mathrm{Im}(\alpha_{e_j})) \simeq k$ (because $\check{\pi}$ is also irreducible: if $W \subset V'$ is G-stable, then $W^\perp = \{x \mid \lambda(x) = 0 \text{ for all } \lambda \in W\} \subset V$ is G-stable also). Then we see that the map

$$\Psi \begin{cases} H \to V \\ \alpha \mapsto ({}^t\alpha)(\mathrm{Id}) \in V, \end{cases}$$

is the inverse of Ξ: we have $\Psi \circ \Xi = \mathrm{Id}_V$, since for $v \in V$ and $w = \Psi \circ \Xi(v)$, we find (denoting $\alpha = \Xi(v)$) using the various definitions that

$$\lambda(w) = \langle \lambda, \Psi(\Xi(v)) \rangle_V = \langle \lambda, {}^t\alpha(\mathrm{Id}) \rangle_V = \langle \alpha(\lambda), \mathrm{Id} \rangle_E = \langle \lambda, \mathrm{Id}(v) \rangle_V = \lambda(v)$$

for all $\lambda \in V'$, and hence $v = w$.

As already indicated, if $R \neq 0$, it must contain a subrepresentation of E', which must be isomorphic to $\check{\pi}$ (see the remark below), and hence we must have $R \supset \mathrm{Im}(\alpha)$ for some non-zero $\alpha \in H$. This means $R \supset \mathrm{Im}(\alpha_v)$ for some non-zero vector in V, and we are done.

Remark 2. We have used some general facts concerning finite-dimensional linear representations of a group G (not necessarily finite). First, Schur's Lemma is valid: for any two finite-dimensional irreducible representations of G on k-vector spaces, say π_1 and π_2, we have

$$\mathrm{Hom}_G(\pi_1, \pi_2) \simeq \begin{cases} 0 & \text{if } \pi_1 \text{ and } \pi_2 \text{ are not isomorphic,} \\ k & \text{if they are.} \end{cases}$$

Second, if a finite-dimensional representation ρ of G on F can be written as a direct sum

$$F = \bigoplus_j F_j$$

where each F_j is isomorphic to a single irreducible representation π, then any non-zero subrepresentation R contains the image of a G-equivariant map $\pi \to R$. We recall how to prove this: first, R contains (by dimension consideration) some non-zero irreducible subrepresentation, say $R_1 \subset R$. Since $R_1 \neq 0$, one composite map

$$R_1 \hookrightarrow F \to F_j$$

at least must have non-zero image. By Schur's Lemma, this means that $R_1 \to F_j$ is an isomorphism, and since $F_j \simeq \pi$ by assumption, we obtain a map $\pi \to R_1$ which must also be an isomorphism.

2. Comments on Burnside's writing

It is of some interest to compare more precisely the wording of the proof in the previous section with what Burnside actually wrote. His paper is roughly four pages long, but the main part – the proof that irreducibility implies that there is no linear relation between the entries of the matrices in G – takes only two pages. Here are some remarks that come to mind.

(1) It is quite surprising that Burnside should require seventeen lines (from line -8, page 430 to line 9, page 431) to prove the trivial direction (existence of relations when G is reducible). This suggests that part (at least) of the complexity of Burnside's proof is more than simply a question of distance in terms of time and language.

(2) The language of Burnside uses the following type of expressions: "successive sets of symbols", "sets of linear functions", "symbols with the same second suffix are transformed among themselves", "undergo the same transformation for every substitution of the group", "group in a set of symbols with the same second suffix", etc. These are very hard to decipher nowadays.

(3) On page 432, Burnside considers, roughly, the projection maps $E' \to W_j$ (in terms of a decomposition (1)) and first considers the case where the image has positive dimension $< n$, arguing that this implies that G is reducible, without using Schur's Lemma however as seems most natural.

(4) The argument starting from the middle of page 432 is – to the author – incomprehensible. Burnside concocts, by manipulating the coefficients of linear equations between the matrix coefficients of the group (after they have been subjected to some transformations and eliminations which must have been obvious to his readers), some linear maps between representations isomorphic to π and invokes Schur's Lemma (without using the name, see [1, p. 432, l. 6]) to say that (provided G is irreducible) they are scalars; his final step is to use

these to prove that $\det(g) = 0$ for $g \in G$ (the scalars being used to create an element in the kernel of g). This endpoint is similar to the one we used, but whether is it really equivalent is unclear to the author...

(5) There is no mention of the underlying field k in Burnside's paper, and the only implicit property of k that is used is Schur's Lemma.

(6) It can be argued that one can condense the non-trivial direction of Burnside's proof, in our writing, in barely more than ten or fifteen lines, leaving only for the reader to check straightforward algebraic calculations, and to remember elementary background facts about linear group actions. In that respect, this proof is not very much harder than simple ones found in the literature, and it seems more intrinsic.

3. Generalizing

In this section, we show how arguments which are very similar to those of Section 1 also lead to the following results, which are incarnations of the fact that non-isomorphic irreducible representations of a group are "independent":

Theorem 3. *Let G be a group and let k be an algebraically closed field. Let $\pi_1, \ldots, \pi_n$ be irreducible k-representations*

$$\pi_i : G \to \mathrm{GL}(V_i)$$

which are pairwise non-isomorphic, with V_i of dimension d_i. Let π be the direct sum of the π_i acting on the direct sum V of the V_i.

(1) *The k-linear subspace of $\mathrm{End}(V)$ spanned by the elements $\pi(g) \in \mathrm{End}(V)$ is equal to the space of those $A \in \mathrm{End}(V)$ which stabilize each subspace $V_i \subset V$.*

(2) *For $v \in V_i$ and $\lambda \in V_i'$, let*

$$f_{v,\lambda} : g \mapsto \lambda(\pi_i(g)v)$$

be the corresponding matrix coefficient. Then the functions $f_{v,\lambda}$, where v runs over the union of bases of V_i, and λ runs over the corresponding dual bases, are k-linearly independent. Moreover, the span of all $f_{v,\lambda}$, in the space

$$C = \{f : G \to k\}$$

of k-valued functions on G, is a subrepresentation of the left-regular representation isomorphic to

$$\bigoplus_i \pi_i^{d_i}.$$

Corollary 4. *Let G be a group, k an algebraically closed field. Then the characters*

$$g \mapsto \mathrm{Tr}(\pi(g))$$

of all irreducible representations

$$\pi : G \to \mathrm{GL}_n(k)$$

are linearly independent over k.

Proof of Theorem 3. We start with (1). The statement can be interpreted as identifying which relations are satisfied by the subspace $k(\pi(G))$ generated by the $\pi(g)$ in $\mathrm{End}(V)$. Hence it is very similar to the proof of Burnside's Theorem.

We therefore proceed similarly. Let $E = \mathrm{End}(V)$, with the action ρ of G given by

$$\rho(g)A = \pi(g)A,$$

and let E' denote the contragredient representation, and

$$R = \{\phi \in E' \mid \phi(\pi(g)) = 0 \text{ for all } g \in G\}$$

the space of relations satisfied by the image of π. As before, R is a subrepresentation of E'.

We now express E' itself as a semisimple representation. Namely, for $1 \leqslant i, j \leqslant n$, let

$$E_{i,j} = \mathrm{Hom}_k(V_j, V_i),$$

given with the G-action

$$g \cdot T = \pi_i(g)T$$

for $T \in E_{i,j}$. If we identify $E_{i.j}$ with the elements of E that map V_j to V_i and all other direct summands V_ℓ to 0, this action corresponds of course to the representation ρ on E, and we have a direct sum decomposition

$$E = \bigoplus_{i,j} E_{i,j} \tag{2}$$

of G-representations, hence also

$$E' = \bigoplus_{i,j} E'_{i,j}.$$

Now we decompose further each $E'_{i,j}$ into irreducibles. For this purpose, we use embeddings

$$\alpha_v \;:\; E'_i \to E'_{i,j}$$

defined for $v \in V_j$ as in the proof of Theorem 1, namely

$$\langle \alpha_v(\lambda), T\rangle = \langle \lambda, Tv\rangle$$

for $\lambda \in E'_i$ and $T \in E_{i,j}$ (observe that this is well-defined since T maps V_j to V_i). Each such embedding is an intertwining operator between E'_i and $E'_{i,j}$ (i.e., between the contragredient of π_i and ρ). By a choice of basis (v_ℓ) of V_j, we then also see that

$$E'_{i,j} = \bigoplus_{\ell} \mathrm{Im}(\alpha_{v_\ell}) \simeq n_j E'_i,$$

and hence that

$$E' = \bigoplus_{i,j} \bigoplus_{\ell} \mathrm{Im}(\alpha_{v_\ell}) \simeq \bigoplus_{i} (\dim V) E'_i, \tag{3}$$

a direct sum of irreducible representations.

Therefore E' is a semsimple representation of G. This implies, as is well-known, that $R \subset E'$ is also semisimple, and that its irreducible components are among the E'_i. However, and again as in the previous proof, the direct sum of the maps

$$\begin{cases} V_j \to \mathrm{Hom}_G(V'_i, E') \\ v \mapsto \alpha_v \end{cases}$$

gives an isomorphism

$$V \to \mathrm{Hom}_G(V'_i, E')$$

(the dimensions agree, using (3), and one has injectivity).

If $i \neq j$ and $v \in V_j$ is non-zero, the image of the corresponding

$$\alpha_v : V_i' \to E'$$

always lies in R: since all $\pi(g)$ stabilize the subspace V_j, they have no component in (2) with $i \neq j$. On the other hand, if $i = j$ and $v \in V_j$ is non-zero, it is not possible that the image of any α_v lie in R: this is the same as end of the proof of Theorem 1.

We can now easily prove (2). Indeed, if we build a basis of V by taking the union of bases of the V_j, and consider the associated dual basis, the "matrix coefficients" $f_{v,\lambda}$ are simply the functions which give the coefficients of the matrix representing $\pi(g)$ in this basis (from and to the given basis of V). Thus any non-trivial linear-dependency relation would correspond to an element of the relation representation R which is *not* of the type permitted by the analysis of R above. It follows that such a relation does not exist.

Now we form the span in C of the functions $(f_{v,\lambda})$, say M. If $v \in V_j$, we note that for g, $x \in G$, denoting by $\cdot$ the left-regular representation on C, we have

$$g \cdot f_{v,\lambda}(x) = f_{v,\lambda}(xg) = \langle \lambda, \pi_j(x)\pi_j(g)v \rangle = f_{\pi_j(g)v,\lambda}(x),$$

or

$$g \cdot f_{v,\lambda} = f_{\pi_j(g),\lambda}.$$

This means that the span of $f_{v,\lambda}$, for fixed λ and $v \in V_j$, is a subrepresentation of C isomorphic to π_j. Varying λ in a basis of V_j', we get a subrepresentation of C isomorphic to the sum of d_j copies of π_j. Then the direct sum over j gives the desired statement $M \subset C$. □

Proof of the corallary. This is a consequence of (2) in Theorem 3: since every character of an irreducible representation π is a linear combination of some matrix coefficients of π, it follows that for any finite set of distinct irreducible representations, the characters are linearly independent. □

References

[1] W. Burnside: *On the condition of reducibility of any group of linear substitutions*, Proc. L.M.S (1905), 430–434; doi:10.1112/plms/s2-3.1.430.

ETH Zürich – DMATH, Rämistrasse 101, 8092 Zürich, Switzerland
E-mail address: kowalski@math.ethz.ch