THE INDEX OF A GROUP IN A SEMIGROUP

George M. Bergman

We shall define the right and left indices of a subgroup G in a semigroup S with unit, and show that if G has cancellation in S, and at least one of these indices is finite, then they are equal. If only right cancellation holds, and the left index is finite, the right index will be either less than the left index, or infinite. It will be shown by counterexamples that these theorems are "best results."

Let S be a semigroup with unit and G a subgroup of S with the same unit. Let $[S:G]_{\wr}$ denote the cardinality of the set of right cosets xG of G in S, and $[S:G]_{\diagup}$ the cardinality of the set of left cosets. These are equal when S is a group, because the operation $x \mapsto x^{-1}$ gives an anti-automorphism of S and of G.

For S any *finite* semigroup, we see by a counting argument that these two indices will be equal if G has left and right cancellation in S:

$$(1) \qquad\qquad gx = x \Rightarrow g = e \quad (x \in S, g \in G),$$

$$(2) \qquad\qquad xg = x \Rightarrow g = e \quad (x \in S, g \in G).$$

If only one of the cancellation conditions is assumed, say (2), then we see $[S:G]_{\wr} \leqq [S:G]_{\diagup}$. As indicated above, we shall generalize these results, replacing the assumption that S is finite by assumptions that one or both of the indices is. But let us begin with two counterexamples, one showing that one-sided cancellation is not enough to imply equal indices even for finite S; the other, that assuming cancellation, the two indices can still be any two infinite cardinals.

1. Counterexamples. (a) *Assuming only right cancellation.*

Let X be a set, G a simply transitive group of permutations of X, and S the semigroup of maps of X into itself consisting of the elements of G and the constant maps, all written on the right. Then it is easy to see that G has right cancellation in S, and that $[S:G]_{\wr} = 2$, G forming one coset, and the constant maps another; but $[S:G]_{\diagup} = 1 + |X|$: each constant map comprises a left coset. In this example, if X is finite, S will also be.

(b) *With cancellation, but infinite indices.* Let G be a group with a one-to-one endomorphism α such that $[G: \alpha(G)] > \aleph_0$. (E.g., let H be an uncountable group, $G = H \oplus H \oplus H \oplus \cdots$, and α the "right shift" endomorphism.) Let S be the semigroup of all expressions $x^n g(n \geqq 0, g \in G)$, with multiplication defined using the law $gx^n =$

$x^n\alpha^n(g)$. Then it is easy to verify that S is a semigroup with cancellation (in fact, embeddable in a group), and that elements $x^m g$ and $x^n h$, though they always belong to the same right G-coset, will belong to the same left coset if and only if $gh^{-1} \in \alpha^n(G)$. Hence $[S : G]_\ell = 1 + 1 + \cdots = \aleph_0$, but $[S : G]_\prime = 1 + [G : \alpha(G)] + [G : \alpha^2(G)] + \cdots = [G : \alpha(G)] > \aleph_0$. By using more indeterminates x, all acting in the same way on G, one can get arbitrary cardinals as left *and* right indices.

2. In the next five sections, we shall prove the two theorems mentioned in the introduction.

Let us begin with some general observations on a semigroup S with unit, having a subgroup G with $[S : G]_\prime < \infty$.

Given any $x \in S$, define $G_x \subseteq G$ as $\{g \in G \mid xg \in Gx\}$. This will be a subgroup of G: in fact, it is the isotropy subgroup of the coset Gx under the action of G, by right multiplication, on the class of left cosets of G in S. Since this class is finite, $[G : G_x] < \infty$.

If G satisfies left and right cancellation in S, the correspondence between elements g and h satisfying $xg = hx$ for given x is one-to-one and will constitute a monomorphism of groups, $\varphi_x : G_x \to G$; what we shall call, in the next section, a "partial monomorphism" of G into itself.

Dropping cancellation again, we note that if x is any element of S, the family of cosets $Gx, Gx^2, Gx^3, \ldots$ will be finite, hence we can find positive integers a, b such that $Gx^a = Gx^{a+b}$. Hence $Gx^a = Gx^{a+mb}$ for all $m \geq 1$. Taking $m = a$ and multiplying on the right by $x^{a(b-1)}$, we get $Gx^{ab} = Gx^{2ab}$, in other words, there is some positive power y of x such that $Gy = Gy^2$, and we can write $y^2 = gy$. It follows that for any $n \geq 0$, $y^{n+1} = g^n y$.

Now again assume right and left cancellation of G in S. Since G_y is of finite index in G, there will exist $c > 0$ such that $g^c \in G_y$. We calculate:

$$g^{c+1}y = y^{c+2} = yg^c y = \varphi_y(g^c)yy = \varphi_y(g^c)gy .$$

Using left cancellation, we conclude that $\varphi_y(g^c) = g^c$, i.e., $yg^c = g^c y$. It follows that $y^{2c} = g^c y^c = y^c g^c$. Putting $z = y^c$, $h = g^c$, we get an equation $z^2 = zh$, where z is again a power of x.

3. **Partial monomorphisms.** By a partial monomorphism f from a group G to a group H, we shall mean a monomorphism of groups, from a subgroup G_f of finite index in G, into the group H. We shall call two partial monomorphisms $f, g : G \to H$ *equivalent* if they agree on some subgroup of finite index in G (not necessarily all of $G_f \cap G_g$).

This is an equivalence relation. (However we shall not here identify equivalent partial monomorphisms.)

By the composition of two partial monomorphisms, $F \xrightarrow{f} G \xrightarrow{g} H$, we shall mean the partial monomorphism gf defined on $F_{gf} = f^{-1}(G_g) \subseteqq F_f$. It is easy to show that this sort of composition respects equivalence.

(Thus one may form a category whose objects are groups, and whose morphisms are equivalence classes of partial monomorphisms, or more generally, of partial homomorphisms, analogously defined. This category might be worthy of some study.)

Note that the images of two equivalent partial homomorphisms $f, g: F \to G$ will be *commensurable* subgroups of G, that is, they have a common subgroup which is of finite index in each of them.

Define the *index* of a partial monomorphism $f: G \to H$ as the quotient $\nu_f = [G: G_f]/[H: f_\backslash G_f)]$, or zero if the denominator of this expression is infinite—in either case, a nonnegative rational number. It is easy to check that the indices of equivalent partial monomorphisms are equal (it is in order that this be true that we have restricted our attention to monomorphisms, rather than defining partial homomorphisms), and with the help of this fact, one checks that the index of a composition of partial monomorphisms is the product of their indices.

Note that for partial monomorphisms, $F \xrightarrow{f} G \xrightarrow{g} H$, the image of gf will have finite index in the image of g if and only if $\nu_f \neq 0$.

4. **Back to semigroups.** Once more let S be a semigroup with unit, and G a subgroup with the same unit, having two-sided cancellation in S, and such that $[S: G]_{\swarrow}$ is finite.

The maps $\varphi_x: G_x \to G$ defined in §2 ($\varphi_x(g) = h$ if $xg = hx$) will be partial monomorphisms of G into itself. Note that we are writing G_x where the terminology of §3 would have G_{φ_x}. We shall similarly put ν_x for ν_{φ_x}.

Given $x, y \in S$, we see that φ_{xy} is an extension of $\varphi_x \varphi_y$, so in particular, these are equivalent partial monomorphisms. Hence ν will be a homomorphism of semigroups from S into the multiplicative semigroup of nonnegative rationals. Note that $\nu_g = 1$ for $g \in G$.

Suppose $\nu_x = 0$ for some $x \in S$. By the argument given in §2, some power z of x will satisfy $z^2 = zh$ for some $h \in G$. Now the image of $\varphi_z \varphi_z$ will have infinite index in the image of φ_z, because $\nu_z = 0$, but it will also be commensurable with the image of the equivalent partial monomorphism $\varphi_{z^2} = \varphi_{zh}$, which is in turn commensurable with the image of $\varphi_z \varphi_h$, which has finite index in the image of φ_z—contradiction!

Hence ν will be a multiplicative homomorphism of S into the positive rationals. But since it sends G to $\{1\}$, and G has finite left

index in S, its image will be a finite subsemigroup of the positive rationals—and the only such semigroup is $\{1\}$. Hence $\nu_x = 1$ for all $x \in S$, i.e., $[G: \varphi_x(G_x)] = [G: G_x]$.

5. **Measure for measure.** Let us define two finitely additive measures μ_{ℓ} and μ_{r} on the set S, taking for a basis of measurable sets of μ_{ℓ} the left cosets Hx of subgroups H of finite index in G, to each of which we assign the measure $1/[G: H]$, and defining the second measure similarly in terms of right cosets. It is clear that these are consistent definitions, and $\mu_{\ell}(S) = [S: G]_{\ell}$, while $\mu_{r}(S)$ is defined if and only if $[S: G]_{r}$ is finite, and if this is so, equals $[S: G]_{r}$.

Given $x \in S$ let us define the semicoset of x as $Gx \cap xG = xG_x = \varphi_x(G_x)x$. Now left translation by the elements of G transitively permutes the semicosets which form a given left coset Gx, and we see that the number of them will be $[G: \varphi_x(G_x)]$. Because $\nu_x = 1$, this equals $[G: G_x]$, which we know is finite. Since there are only finitely many left cosets Gx, the number of semicosets will be finite.

The semicosets $xG_x = \varphi_x(G_x)x$ will be measurable under μ_{r} with measure $1/[G: G_x]$, and under μ_{ℓ} with measure $1/[G: \varphi_x(G_x)]$. As we just observed, these numbers are equal. Hence we have partitioned S into finitely many μ_{r}-and-μ_{ℓ} measurable sets on each of which these two measures have the same value. So $\mu_{\ell}(S) = \mu_{r}(S)$, completing the proof of:

THEOREM 1. *Let S be a semigroup with unit e, and $G \subsetneqq S$ a subgroup with the same unit, having left and right cancellation in S. Then if $[S: G]_{\ell} < \infty$, we have $[S: G]_{r} = [S: G]_{\ell}$.*

By symmetry, the same is true if $[S: G]_{r} < \infty$.

6. **One-sided cancellation.** Now let us suppose only that G has *right* cancellation in S, and $[S: G]_{\ell} < \infty$.

The subgroups G_x defined in §2 still have finite index in G. Since G_x depends only on the coset Gx, there are only finitely many distinct groups G_x, hence their intersection, which we shall call G_0, also has finite index in G.

Let us measure the failure of the left cancellation law by writing, for each $x \in S$, $H_x = \{g \in G \mid gx = x\}$, the isotropy subgroup of x. We can now prove:

THEOREM 2. *Let S be a semigroup with unit e, and $G \subsetneqq S$ a subgroup with the same unit, satisfying right cancellation: $xg = x \Rightarrow g = e$ $(x \in S, g \in G)$, and having finite left index: $[S: G]_{\ell} < \infty$. Then either of the hypotheses:*

(a) $[S:G]_{\ell} < \infty$

(b)   *None of the isotropy subgroups $H_x \subsetneqq G$ is conjugate in G to a proper subgroup of itself,*
implies:

(c) $[S:G]_{\ell} \leq [S:G]_{\nearrow}$,

(d)   *the subgroup $G_0 = \{g \in G \mid \forall x \in S, \, xg \in Gx\}$ (of finite index in G) has left (as well as right) cancellation in S,*

(e)   *Each isotropy subgroup $H_x \subsetneqq G$ has only finitely many distict conjugates, as a subgroup of G,*

(f)   *There are only finitely many distinct isotropy subgroups H_x,*

(g)   *Every isotropy subgroup $H_x \subsetneqq G$ is finite.*
In fact, conditions (a)–(g) are all equivalent.

Proof. The scheme of proof will be:
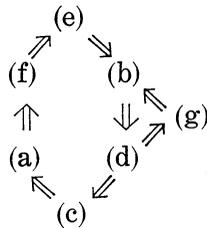
Aside from an application of Theorem 1, the one key step, with which we shall begin, is:

(b) $\Rightarrow$ (d):   We wish to prove that for all $x \in S$, $H_x \cap G_0 = \{e\}$. Suppose we have an x for which this is not true. By the argument given in §2 we can find a positive power y of x such that $y^2 = gy$ for some $y \in G$. Note that $H_{y^2} \supseteq H_y \supseteq H_x$. Let h be an element of $H_y \cap G_0 - \{e\}$. Since $h \in G_0 \subsetneqq G_y$, we can write $yh = iy$ ($i \in G$). By right cancellation we have $y \neq yh = iy$, so $i \notin H_y$. But $iy^2 = yhy = y^2$, so $i \in H_{y^2} = H_{gy} = g(H_y)g^{-1}$. Hence H_{y^2} is conjugate to a proper subgroup, H_y, of itself, contradicting (b).

(d) $\Rightarrow$ (g):   By (d), no H_x meets the subgroup G_0, hence each H_x must have order $\leq [G:G_0]$.

(g) $\Rightarrow$ (b):   Immediate.

(d) $\Rightarrow$ (c): G_0 will have finite left index in S, so we can apply Theorem 1 to conclude $[S:G_0]_{\ell} = [S:G_0]_{\nearrow}$. Comparing G-cosets and G_0-cosets, we get $[S:G]_{\ell} = [S:G_0]_{\ell}/[G:G_0] = [S:G_0]_{\nearrow}/[G:G_0] \leq [S:G]_{\nearrow}$.

(c) $\Rightarrow$ (a):   Immediate.

(a) $\Rightarrow$ (f):   It is easy to see that two elements of the same right coset have the same left isotropy subgroup. This makes the implication clear.

(f) $\Rightarrow$ (e):   For, any conjugate of an isotropy subgroup is an isotropy

subgroup: $gH_xg^{-1} = H_{gx}$.

(e) $\Rightarrow$ (b): If $gH_xg^{-1} \subset H_x$, then clearly $H_x \subset gH_xg^{-1} \subset g^2H_xg^{-2} \subset \cdots$ are all distinct, contradicting (e).

One can state many special cases and corollaries to this result. For instance, if G has right cancellation and finite right index in S, then $[S: G]_\angle \geqq [S: G]_y$. If G is a group with ascending chain condition on subgroups, or an abelian group then condition (b) is automatic, hence given $S \supseteqq G$ as in Theorem 2, all of (a)–(g) hold.

A question we have not been able to answer is: if G in Theorem 2 is finitely generated, or more generally, if S is finitely generated, must conditions (a)–(g) hold? (A finitely generated group *can* have a subgroup properly containing a conjugate of itself. E.g., the group defined by two generators a and b, and one relation $aba^{-1} = b^2$.)

7. **Another counterexample.** We shall now give a rather complicated example (perhaps necessarily so) showing that conditions (a)–(g) of Theorem 2 can, in fact, fail to hold.

Let G be a free group on countably many generators $g_0, g_1, \cdots$.

Let S be obtained by adjoining to G, as a semigroup with unit, one generator y, and the relations

$$(3) \qquad\qquad yy = g_0y ,$$

$$(4) \qquad\qquad yg_i^{\pm 1} = g_{i+1}^{\pm 1}y, \qquad (i = 0, 1, \cdots) .$$

(Actually, the conditions (4) with exponent -1 are implied by those with $+1$.)

Applying (3) and (4) to the expression $yyg_{i-1}^{\pm 1}(i = 1, 2, \cdots)$ in the two possible orders, we get the equations $g_{i+1}^{\pm 1}g_0y = g_0g_i^{\pm 1}y$. Applying (3) and (4) to yyy in the two possible ways, we get $g_1g_0y = g_0g_0y$, i.e., the $i = 0$ case of the preceding equation with exponent $+1$ only. Multiplying on both sides by g_1^{-1} and transposing, we get $g_1^{-1}g_0^2y = g_0y$.

Finally, we can take each of the above equations, multiply on the right by an arbitrary sequence F_0 of $g_j^{\pm 1}$'s and y's, and apply (3) and (4) to get:

$$(5) \quad \begin{cases} g_{i+1}^{\pm 1}g_0Fy = g_0g_i^{\pm 1}Fy \\ \text{where } i = 0, 1, \cdots, \text{ but only the exponent} \\ +1 \text{ is allowed for } i = 0, \text{ and} \\ g_1^{-1}g_0^2Fy = g_0Fy . \end{cases} \begin{array}{l} (F \text{ any expression} \\ \text{in } g_0, g_1^{\pm 1}, g_2^{\pm 1}, \cdots) . \end{array}$$

To have a set of relations defining S as a semigroup, we also want to write down the semigroup-relations defining the free group G:

$$(6) \qquad \begin{aligned} g_i g_i^{-1} &= e, \\ g_i^{-1} g_i &= e \end{aligned} \qquad (i = 0, 1, \cdots) .$$

We now claim that the elements of S can be uniquely represented by the set of words w in the $g_i^{\pm 1}$ and y not involving any of the expressions on the left-hand side of equations (3)–(6). This means, in fact, words $w = A$, where A is a reduced expression for an element of the free group G, and words $w = Ay$, where A is as before, and w does not *end* in one of the left-hand expressions of (5).

It is clear on the one hand that repeated arbitrary applications of the reductions (3)–(6) will eventually bring any word to such a form. To show the uniqueness of this expression, it suffices to verify that when we have an overlap of two reducible expressions of the sort given by (3)–(6), the results of the two possible reductions can be further reduced to some common expression. (Cf. [1, Th. III. 9.3].)

All but *one* of these calculations are either trivial, or reduce to calculations made in the preceding paragraphs, and we leave the verification to the reader. The one case requiring special consideration is that of an expression yAy, where Ay is the left-hand side of one of the equations of (5). Let us denote by By the right-hand side of the same equation.

By repeated application of (4), yAy can be reduced to $A'yy$, where A' is the result of raising the subscript of each term in A by 1. Equation (3) then gives $A'g_0 y$, and repeated applications of (5) yield $g_0 A y$, since g_0 lowers subscripts as it moves to the left, so long as it encounters no g_i^{-1}'s (which it won't by the conditions in (5)) and so long as it has no g_0^{-1} to the right of it (ditto). We can *then* reduce $g_0 A y$ to $g_0 B y$. Alternatively, we could have begun by reducing yAy to yBy, and this then to $B'yy$, $B'g_0 y$, and $g_0 B y$. The resulting expressions are the same, completing the desired step in the verification of our normal-form assertion.

To determine $[S:G]_{\ell}$ and $[S:G]_{\ell'}$, let us consider an element of S whose expression in our reduced form is Ay. Write A as BC', where $B, C \in G$, and C is chosen so that C' (as in the preceding paragraph) is the largest right-hand segment of A not involving g_0 or g_0^{-1}. Then B must either be empty or end in $g_0^{\pm 1}$. We can thus, by (4) rewrite our element as ByC, and we observe that By will again be irreducible. It is easily deduced that every element of S can be written uniquely either as a reduced expression for an element of G, or as ByC, where B and C are such reduced expressions, and B either is empty, or ends in $g_0^{\pm 1}$. It follows that $[S:G]_{\ell} = \aleph_0$, the cosets other than G corresponding to these elements B, and that G has right cancellation in S.

On the other hand, since every element of S can be written

as g or (nonuniquely) gy, we have $[S:G]_{\swarrow} = 2$.

Larger values of $[S:G]_z$ could have been obtained by adjoining to G still more free generators, $h_{\alpha,i}$ (where i takes on *all* integer values, and α ranges over a large index set) and the relations $yh_{\alpha,i}^{\pm 1} = h_{\alpha,i+1}^{\pm 1}y$, for all i and α.

REFERENCE

1. P. M. Cohn, *Universal Algebra*, Harper and Row, 1965.

Permanent affiliation:
HARVARD UNIVERSITY
BEDFORD COLLEGE, LONDON
UNIVERSITY OF CALIFORNIA, BERKELEY