COSETS AND LAGRANGE'S THEOREM

KEITH CONRAD

1. Introduction

Pick an integer $m \neq 0$. For $a \in \mathbf{Z}$, the congruence class $a \bmod m$ is the set of integers $a + mk$ as k runs over $\mathbf{Z}$. We can write this set as $a + m\mathbf{Z}$. This can be thought of as a translated subgroup: start with the subgroup $m\mathbf{Z}$ and add a to it. This idea can be carried over from $\mathbf{Z}$ to any group at all, provided we distinguish between translation of a subgroup on the left and on the right.

Definition 1.1. Let G be a group and H be a subgroup. For $g \in G$, the sets

$$gH = \{gh : h \in H\}, \quad Hg = \{hg : h \in H\}$$

are called, respectively, a *left H-coset* and a *right H-coset*.

In other words, a coset is what we get when we take a subgroup and shift it (either on the left or on the right). The best way to think about cosets is that they are *shifted subgroups*, or *translated subgroups*.

Note g lies in both gH and Hg, since $g = ge = eg$. Typically $gH \neq Hg$. When G is abelian, though, left and right cosets of a subgroup by a common element are the same thing. When an abelian group operation is written additively, an H-coset should be written as $g + H$, which is the same as $H + g$.

Example 1.2. In the additive group $\mathbf{Z}$, with subgroup $m\mathbf{Z}$, the $m\mathbf{Z}$-coset of a is $a + m\mathbf{Z}$. This is just a congruence class modulo m.

Example 1.3. In the group $\mathbf{R}^\times$, with subgroup $H = \{\pm 1\}$, the H-coset of x is $xH = \{x, -x\}$. This is "x up to sign."

Example 1.4. When $G = S_3$, and $H = \{(1), (12)\}$, the table below lists the left H-cosets and right H-cosets of every element of the group. Compute a few of them for non-identity elements to satisfy yourself that you understand how they are found.

g	gH	Hg
(1)	$\{(1), (12)\}$	$\{(1), (12)\}$
(12)	$\{(1), (12)\}$	$\{(1), (12)\}$
(13)	$\{(13), (123)\}$	$\{(13), (132)\}$
(23)	$\{(23), (132)\}$	$\{(23), (123)\}$
(123)	$\{(13), (123)\}$	$\{(23), (123)\}$
(132)	$\{(23), (132)\}$	$\{(13), (132)\}$

Notice first of all that cosets are usually not subgroups (some do not even contain the identity). Also, since $(13)H \neq H(13)$, a particular element can have different left and right H-cosets. Since $(13)H = (123)H$, different elements can have the same left H-coset. (You have already seen this happen with congruences: $14 + 3\mathbf{Z} = 2 + 3\mathbf{Z}$, since $14 \equiv 2 \bmod 3$.)

In the next section we will see how cosets look in some geometric examples, where we can visualize cosets. Then we will see that cosets arise in certain decimal patterns and as "inhomogeneous solution spaces" to linear equations or differential equations. Then we will look at some general properties of cosets. The *index* of a subgroup in a group, which tells us how many cosets the subgroup has (either on the right or on the left), will lead to the most basic important theorem about finite groups: Lagrange's theorem. We will see a few applications of Lagrange's theorem and finish up with the more abstract topics of left and right coset spaces and double coset spaces.

2. Geometric examples of cosets

When a group is defined in terms of vectors and matrices, we can often get a picture of the group and its cosets.

Example 2.1. Let $G = \mathbf{R}^2$ and $H = \mathbf{R}\mathbf{e}_1$ be the x-axis. The (left) H-coset of a vector $\mathbf{v} \in \mathbf{R}^2$ is

$$\mathbf{v} + H = \mathbf{v} + \mathbf{R}\mathbf{e}_1 = \{\mathbf{v} + c\mathbf{e}_1 : c \in \mathbf{R}\}.$$

This is the line parallel to H (the x-axis) that passes through the endpoint of $\mathbf{v}$. The H-cosets in general are the lines parallel to H. Two parallel lines are either equal or disjoint, so any two H-cosets are equal or disjoint. In Figure 1, the H-cosets of $\mathbf{v}$ and $\mathbf{v}'$ are equal while those of $\mathbf{v}$ and $\mathbf{w}$ are disjoint.

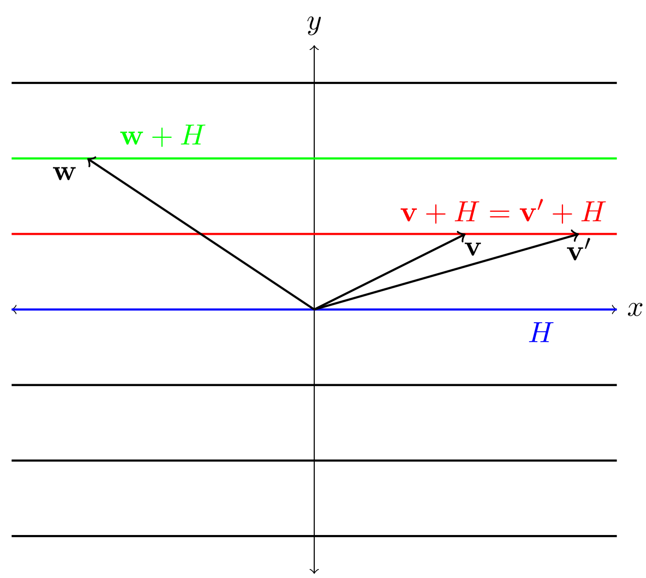

Figure 1. The cosets of $\mathbf{R}\mathbf{e}_1$ in $\mathbf{R}^2$.

Example 2.2. Let $G = \mathrm{Aff}^+(\mathbf{R})$, the 2×2 matrices $\left(\begin{smallmatrix} x & y \\ 0 & 1 \end{smallmatrix}\right)$ with $x > 0$ under matrix multiplication. Geometrically, we identify elements of G with points (x, y) in the plane where $x > 0$. Such points form a right half-plane:

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \longleftrightarrow (x, y) \in \mathbf{R}_{>0} \times \mathbf{R}.$$

The identity element becomes the point $(1,0)$. See Figure 2. This is quite similar to the idea of identifying complex numbers $x+yi$ with points (x,y) in the plane.

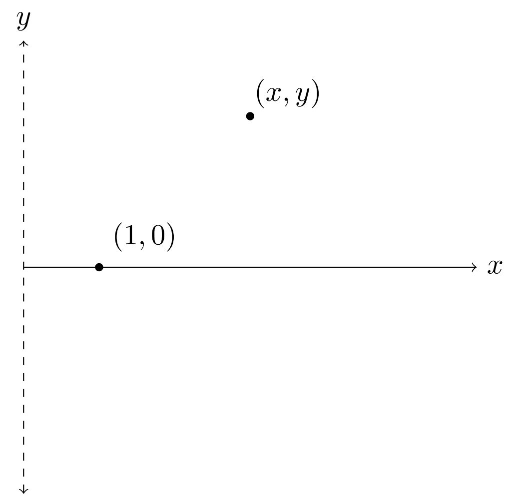

FIGURE 2. The group $\mathrm{Aff}^+(\mathbf{R})$.

Unlike the geometric interpretation of complex number addition, which is ordinary vector addition with the intuitive "parallelogram rule", the group law on $\mathrm{Aff}^+(\mathbf{R})$ when written in terms of points in the plane is unfamiliar: starting from

$$\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u & v \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} xu & xv+y \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/x & -y/x \\ 0 & 1 \end{pmatrix},$$

we get the group law on points

$$(x,y)(u,v) = (xu, xv+y), \quad (x,y)^{-1} = \left(\frac{1}{x}, -\frac{y}{x}\right).$$

Any matrix $\left(\begin{smallmatrix} x & y \\ 0 & 1 \end{smallmatrix}\right)$ in $\mathrm{Aff}^+(\mathbf{R})$ breaks up as the product $\left(\begin{smallmatrix} 1 & y \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} x & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (*not* as $\left(\begin{smallmatrix} x & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & y \\ 0 & 1 \end{smallmatrix}\right)$!). These special types of matrices each form subgroups, which we will call H and K:

$$H = \left\{ \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} : y \in \mathbf{R} \right\}, \quad K = \left\{ \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} : x > 0 \right\}. \tag{2.1}$$

They are pictured in Figure 3.

For $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ in $\mathrm{Aff}^+(\mathbf{R})$, what are its left and right H-cosets $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)H$ and $H\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$? To compute these multiply a general element of H on the left and right by $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$:

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & ay+b \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b+y \\ 0 & 1 \end{pmatrix}.$$

Here a and b are fixed, while y varies over $\mathbf{R}$. The numbers $ay+b$ run over all of $\mathbf{R}$ and the numbers $b+y$ run over $\mathbf{R}$. This means the left and right H-cosets of $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ in $\mathrm{Aff}^+(\mathbf{R})$ are the same:

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} H = H \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \left\{ \begin{pmatrix} a & t \\ 0 & 1 \end{pmatrix} : t \in \mathbf{R} \right\}. \tag{2.2}$$

In the picture of $\mathrm{Aff}^+(\mathbf{R})$, this is simply the vertical line parallel to H passing through the point (a,b). See the first picture in Figure 4.

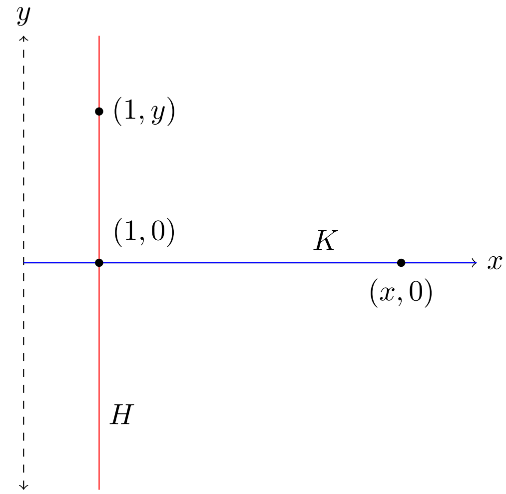

FIGURE 3. The group $\mathrm{Aff}^+(\mathbf{R})$.

To compute the left K-coset of $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$, for any $x > 0$ we have

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} ax & b \\ 0 & 1 \end{pmatrix}. \tag{2.3}$$

As $x > 0$ varies, the right side of (2.3) runs through all matrices of the form $\left(\begin{smallmatrix} t & b \\ 0 & 1 \end{smallmatrix}\right)$ with $t > 0$. Therefore

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} K = \left\{\begin{pmatrix} t & b \\ 0 & 1 \end{pmatrix} : t > 0\right\}. \tag{2.4}$$

In particular, the left K-coset of $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ is determined by b alone and is independent of the choice of $a > 0$: $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)K = \left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)K$. This is the horizontal line through (a, b). See the second picture in Figure 4.

To find out what the right K-coset of $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ is, multiply a general element of K by $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ on the right:

$$\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} ax & bx \\ 0 & 1 \end{pmatrix}.$$

Letting $x > 0$ vary, we obtain

$$K\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} = \left\{\begin{pmatrix} u & v \\ 0 & 1 \end{pmatrix} : u > 0, v = \frac{b}{a}u\right\}. \tag{2.5}$$

This is a half-line out of the origin with slope b/a (third picture in Figure 4). In particular, the left and right K-cosets of $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ in $\mathrm{Aff}^+(\mathbf{R})$ are *not* the same if $b \neq 0$. A right K-coset of $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right)$ depends on both a and b, or more precisely on their ratio b/a, while its left K-coset depends only on b.

For both H and K, its collection of all left cosets and its collection of all right cosets fill up the group $\mathrm{Aff}^+(\mathbf{R})$ in Figure 5 *without overlap*: a family of parallel half-lines, either horizontal or vertical, and a family of half-lines out of the origin.

Later in this handout, we will see that left (or right) cosets of a subgroup in any group exhibit properties that can be seen in our geometric examples: different left cosets of a

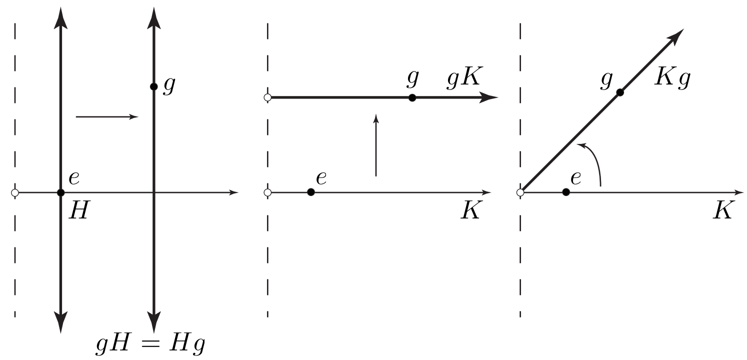

Figure 4. Left and Right Cosets of H and K in $\mathrm{Aff}^+(\mathbf{R})$.

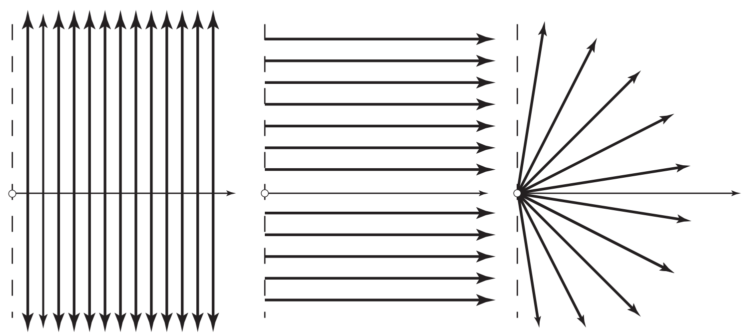

Figure 5. Left and Right Coset Decompositions of $\mathrm{Aff}^+(\mathbf{R})$ by H and K.

subgroup are disjoint, and the collection of all left cosets of a subgroup cover the group (likewise for right cosets).

3. Cosets and decimal expansions

Any rational number a/b whose denominator b is relatively prime to 10 has a decimal expansion that is purely periodic. For example,

$$\frac{1}{3} = .3333333\ldots, \quad \frac{1}{7} = .142857142857142857\ldots, \quad \frac{15}{41} = .609756097560975\ldots$$

where the repeating part has 1 term, 6 terms, and 5 terms respectively. We abbreviate the decimal expansion by writing only the periodic part and put a line over it, so

$$\frac{1}{3} = .\overline{3}, \quad \frac{1}{7} = .\overline{142857}, \quad \frac{15}{41} = .\overline{60975}.$$

In the tables below are all the *reduced* proper fractions with denominator 7, 13, and 27. (We don't include $3/27$ in the second table, for instance, since it is not reduced.)

Fraction	Decimal	Fraction	Decimal	Fraction	Decimal
1/7	$.\overline{142857}$	1/13	$.\overline{076923}$	2/13	$.\overline{153846}$
2/7	$.\overline{285714}$	3/13	$.\overline{230769}$	4/13	$.\overline{307692}$
3/7	$.\overline{428571}$	5/13	$.\overline{384615}$	6/13	$.\overline{461538}$
4/7	$.\overline{571428}$	7/13	$.\overline{538461}$	8/13	$.\overline{615384}$
5/7	$.\overline{714285}$	9/13	$.\overline{692307}$	10/13	$.\overline{769230}$
6/7	$.\overline{857142}$	11/13	$.\overline{846153}$	12/13	$.\overline{923076}$

Fraction	Decimal	Fraction	Decimal	Fraction	Decimal
1/27	$.\overline{037}$	2/27	$.\overline{074}$	4/27	$.\overline{148}$
5/27	$.\overline{185}$	7/27	$.\overline{259}$	8/27	$.\overline{296}$
10/27	$.\overline{370}$	11/27	$.\overline{407}$	13/27	$.\overline{481}$
14/27	$.\overline{518}$	16/27	$.\overline{592}$	17/27	$.\overline{629}$
19/27	$.\overline{703}$	20/27	$.\overline{740}$	22/27	$.\overline{814}$
23/27	$.\overline{851}$	25/27	$.\overline{925}$	26/27	$.\overline{962}$

There are two interesting features revealed in this table:

- the decimal expansions for all the *reduced* fractions with the same denominator have the same period length (denominator 7 has period length 6, denominator 13 has period length 6, and denominator 27 has period length 3),
- if we cyclically shift the digits in the repeating part of the decimal expansions of a reduced fraction, we get the repeating part of the decimal expansion of another reduced fraction with the same denominator.

The first feature is explained by Euler's theorem from elementary number theory, and we won't discuss it here. We want to focus on the second feature. First, let's make sure we understand what it is saying. The string of digits 142857 is the repeating part of the decimal for 1/7, and if we shift the numbers one position to the left, getting 428571, we look in the first table and see it is the repeating part of the decimal for 3/7. If we shift one position to the left again, we get 285714, and when we look at the first table we find that is the repeating part of the decimal for 2/7. Similarly, 407 is the repeating part of the decimal for 11/27, and if we shift the terms one position to the left we get 074, which the second table tells us is the repeating part of the decimal for 2/27. Are you surprised by this? It turns out this phenomenon, which even a grade school student could discover, is actually cosets in disguise. This is best explained by an example.

Example 3.1. Consider 5/13, which the first table tells us is $.\overline{384615}$. Multiplying by 10 has the effect of cyclically shifting the digits in the decimal one position to the left:

$$10 \cdot \frac{5}{13} = 10(.\overline{384615}) = 10(.384615384615\ldots) = 3.84615384615\ldots = 3.\overline{846153}.$$

The numerator of 10(5/13) is 50, and $50 \equiv 11 \bmod 13$, so $50/13 = 11/13 + \text{integer}$. Since we already saw that $50/13 = 3.\overline{846153}$, the integer part of 50/13 is 3 and therefore $11/13 = .\overline{846153}$. In other words, shifting the digits in $.\overline{384615}$ one position to the left turns 5/13 into 11/13, and the reason is that shifting digits in a decimal one position to the left corresponds to multiplication by 10 and $5 \cdot 10 \equiv 11 \bmod 13$.

If we want to shift the digits to the left one more time, the fraction we get will again be obtained from multiplying the new numerator by 10 and reducing the product mod 13: $11 \cdot 10 = 110 \equiv 6 \bmod 13$, or in terms of the original numerator, $5 \cdot 10^2 \equiv 6 \bmod 13$. In general, if $5 \cdot 10^k \equiv r \bmod 13$, where $0 \leq r < 13$, then $r/13$ will have a decimal expansion whose repeating part is the digits for $5/13$ shifted to the left k times. All these decimals have the same length, which is 6, and that means after we do this shifting 6 times we come back to the original fraction and have run through all the cycle shifts of the digits in the decimal for $5/13$. The fractions we have created in this way are all $r/13$ where $r \equiv 5 \cdot 10^k \bmod 13$ for some k. In other words, the numerators of the fractions with denominator 13 having a decimal with the same digit cycle (up to the choice of starting point) as the decimal for $5/13$ is precisely $\{5 \cdot 10^k \bmod 13 : k = 1, 2, \ldots\}$, which is a coset in $(\mathbf{Z}/(13))^\times$ of the cyclic subgroup $\langle 10 \bmod 13\rangle$.

In general, if a/b is a reduced fraction where $(10, b) = 1$, then the other reduced fractions whose decimal has a repeating block that is a shift of the repeating block of the decimal for a/b are precisely those fractions with denominator b and numerator in the coset $a\langle 10 \bmod b\rangle$ of $(\mathbf{Z}/(b))^\times$.

Example 3.2. The fractions $2/27, 11/27$, and $20/27$ all have decimals with digit cycle 074 up to a shift, and in $(\mathbf{Z}/(27))^\times$ we have $\langle 10 \bmod 27\rangle = \{1, 10, 19\}$ and modulo 27 the numerators of the fractions are $\{2, 11, 20\} = \{2, 2 \cdot 10^2, 2 \cdot 10\} = 2\langle 10 \bmod 27\rangle$, which is a coset of $\langle 10 \bmod 27\rangle$.

4. Inhomogeneous solution spaces are cosets

When solving a linear equation $A\mathbf{x} = \mathbf{b}$ (A is a matrix), where the right side is nonzero, we call the equation inhomogeneous. The general solution to this type of equation is $\mathbf{x}_0 + \mathbf{y}$, where $\mathbf{x}_0$ is a *particular* solution of $A\mathbf{x} = \mathbf{b}$ and $\mathbf{y}$ is the *general* solution to the (homogeneous) equation $A\mathbf{x} = \mathbf{0}$. Let's see how the solutions to an inhomogeneous linear equation or inhomogeneous linear different equation are an example of a coset. I do not think the two examples that follow will give *insight* into cosets; they simply serve to show that cosets appear in courses usually taken before abstract algebra.

Example 4.1. Consider a system of linear equations, written as a matrix equation $A\mathbf{x} = \mathbf{b}$, where A is an $m \times n$ real matrix, $\mathbf{b} \in \mathbf{R}^m$, and $\mathbf{x}$ is an unknown vector in $\mathbf{R}^n$. Let W be the solution set of the corresponding homogeneous equation, where the right side is $\mathbf{0}$:

$$W = \{\mathbf{x} \in \mathbf{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

This is a group under vector addition. (It contains $\mathbf{0} \in \mathbf{R}^n$, if $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x}' = \mathbf{0}$ then $A(\mathbf{x}+\mathbf{x}') = A\mathbf{x} + A\mathbf{x}' = \mathbf{0} + \mathbf{0} = \mathbf{0}$, and $A(-\mathbf{x}) = -A(\mathbf{x}) = -\mathbf{0} = \mathbf{0}$.) For any $\mathbf{b} \neq \mathbf{0}$, if the equation $A\mathbf{x} = \mathbf{b}$ has a solution, say $\mathbf{x}_0$, then the general solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x}_0 + W$. Indeed, since for any $\mathbf{x} \in \mathbf{R}^n$ we have the following equivalences:

$$\begin{aligned} A\mathbf{x} = \mathbf{b} &\iff A\mathbf{x} = A\mathbf{x}_0 \quad \text{since } A\mathbf{x}_0 = \mathbf{b} \\ &\iff A(\mathbf{x} - \mathbf{x}_0) = \mathbf{0} \\ &\iff \mathbf{x} - \mathbf{x}_0 \in W \\ &\iff \mathbf{x} \in \mathbf{x}_0 + W. \end{aligned}$$

The solution set $\mathbf{x}_0 + W$ is a coset of W in $\mathbf{R}^n$.

Remark 4.2. Consider the linear differential equation

$$y' + xy = x^3. \tag{4.1}$$

The solutions to $y' + xy = 0$, where the right side is 0 (homogeneous case), are the functions of the form $ce^{-x^2/2}$. A particular solution to the equation (4.1) is $x^2 - 2$, and anyone who had a good course in differential equations will know the general solution to (4.1) must take the form $x^2 - 2 + ce^{-x^2/2}$. Letting c vary, the solution set to (4.1) is $x^2 - 2 + \mathbf{R}e^{-x^2/2}$. This is a coset of functions, so the same phenomenon we met with linear equations also occurs for linear differential equations.

5. Properties of cosets

We will generally focus our attention on left cosets of a subgroup. Proofs of the corresponding properties of right cosets will be completely analogous, and can be worked out by the reader.

Since $g = ge$ lies in gH, every element of G lies in some left H-coset, namely the left coset defined by the element itself. (Take a look at Example 1.4, where (13) lies in $(13)H$.) Similarly, $g \in Hg$ since $g = eg$.

A subgroup is always a left and a right coset of itself: $H = eH = He$. (This is saying nothing other than the obvious fact that if we multiply all elements of a subgroup by the identity, on either the left or the right, we get nothing new.) What is more important to recognize is that we can have $gH = H$ (or $Hg = H$) even when g is not the identity. For instance, in the additive group $\mathbf{Z}$, $10 + 5\mathbf{Z} = 5\mathbf{Z}$. All this is saying is that if we shift the multiples of 5 by 10, we get back the multiples of 5. Isn't that obvious? In fact, the only way we can have $a + 5\mathbf{Z} = 5\mathbf{Z}$ is if a is a multiple of 5, *i.e.*, if $a \in 5\mathbf{Z}$.

For a subgroup H of a group G, and $g \in G$, when does gH equals H?

Theorem 5.1. *For $g \in G$, $gH = H$ if and only if $g \in H$.*

Proof. Since $g = ge \in gH$, having $gH = H$ certainly requires $g \in H$.

Now we need to show that if $g \in H$, then $gH = H$. We prove $gH = H$ by showing each is a subset of the other. Since $g \in H$, $gh \in H$ for any $h \in H$, so $gH \subset H$. To see $H \subset gH$, note $h = g(g^{-1}h)$ and that $g^{-1}h$ is in H (since $g^{-1} \in H$). □

Example 5.2. Consider the subgroup $H = \{1, s\}$ of D_4. We have

$$sH = \{s, s^2\} = \{s, e\} = H.$$

Two of the elements of D_4 that are not in H are r and r^2. Their left H-cosets are not equal to H:

$$rH = \{r, rs\}, \quad r^2H = \{r^2, r^2s\}.$$

Notice the left H-cosets H, rH, and r^2H are not just unequal (which means each has an element not in another), but are in fact *disjoint* (which means no element of one is in another). This is similar to Example 2.1, where the cosets are a family of parallel lines: different parallel lines are disjoint. This is a completely general phenomenon, as follows.

Theorem 5.3. *If two left H-cosets share a common element, then they are equal. Equivalently, two left H-cosets that are not equal have no elements in common, i.e., they are disjoint.*

Proof. To show left H-cosets with a common element are the same, suppose x belongs to g_1H and to g_2H, say

$$x = g_1h_1 = g_2h_2 \tag{5.1}$$

where $h_1, h_2 \in H$. Then $g_1 = g_2h_2h_1^{-1}$. Any element of g_1H has the form g_1h for some $h \in H$, and

$$g_1h = g_2(h_2h_1^{-1}h) \in g_2H.$$

Since h was arbitrary in H, we see $g_1H \subset g_2H$. The reverse inclusion, $g_2H \subset g_1H$, follows by a similar argument (use the equation $g_2 = g_1h_1h_2^{-1}$ instead). $\square$

By Theorem 5.3, no element lies in more than one left H-coset. We call any element of a left coset a *representative* of that coset. A set of representatives for all the left H-cosets is called a *complete set* of left coset representatives.

Example 5.4. By the table in Example 1.4, each element of S_3 is in exactly one of H, $(13)H$, and $(23)H$, so

$$S_3 = H \cup (13)H \cup (23)H. \tag{5.2}$$

This is a union of disjoint sets. An example of a complete set of left coset representatives of H is $(1), (13)$, and (23). Another complete set of left coset representatives of H is (12), (13), and (132).

Example 5.5. Let $G = \mathbf{Z}$ and $H = m\mathbf{Z}$, for $m > 0$. The coset decomposition of H in G (left and right is the same) is just the decompsition of $\mathbf{Z}$ into congruence classes modulo m:

$$\mathbf{Z} = m\mathbf{Z} \cup (1 + m\mathbf{Z}) \cup (2 + m\mathbf{Z}) \cup \cdots \cup (m - 1 + m\mathbf{Z}). \tag{5.3}$$

It is standard to use $\{0, 1, \dots, m-1\}$ as the coset representatives.

Example 5.6. Let $G = \mathrm{Aff}^+(\mathbf{R})$ and $K = \{(\begin{smallmatrix} x & 0 \\ 0 & 1 \end{smallmatrix}) : x > 0\}$. We saw in Example 2.2 that any left K-coset of a matrix depends only on the upper-right entry of the matrix: $(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix})K = \{(\begin{smallmatrix} x & b \\ 0 & 1 \end{smallmatrix}) : x > 0\}$. Moreover, since all matrices in a given left K-coset have a common upper-right entry, we can use the upper-right entry to parametrize different left K-cosets:

$$\mathrm{Aff}^+(\mathbf{R}) = \bigcup_{b \in \mathbf{R}} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} K, \tag{5.4}$$

where the union is disjoint. Coset representatives are the matrices $(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix})$.

Example 5.7. Consider the (additive) group $\mathbf{R}$ and the subgroup $\mathbf{Z}$. Every real number, up to addition by an integer, looks like a number between 0 and 1 (more precisely, in the half-open interval $[0, 1)$). Different numbers in this range are not $\mathbf{Z}$-translates of each other, so $[0, 1)$ is a complete set of (left) coset representatives of $\mathbf{Z}$ in $\mathbf{R}$:

$$\mathbf{R} = \bigcup_{0 \le x < 1} (x + \mathbf{Z}), \tag{5.5}$$

where the sets making up this union are disjoint.

Example 5.8. We saw in Example 2.1, where $G = \mathbf{R}^2$, that the cosets of the x-axis $\mathbf{R}\mathbf{e}_1$ in the plane are the lines parallel to the x-axis. A complete set of left coset representatives of $\mathbf{R}\mathbf{e}_1$ is a choice of one point on each line parallel to the x-axis. These choices could be made at random, but a nicer method is to use the points lying on a line (any line) not parallel to the x-axis. Such a line will pass once through any line parallel to the x-axis. An example of such a line is the y-axis, *i.e.*, the y-axis is a set of coset representatives for $\mathbf{R}\mathbf{e}_1$ in $\mathbf{R}^2$:

$$\mathbf{R}^2 = \bigcup_{y \in \mathbf{R}} ((0, y) + \mathbf{R}\mathbf{e}_1), \tag{5.6}$$

where the union is disjoint. Coset representatives are $\{(0, y) : y \in \mathbf{R}\}$.

The decompositions in (5.2), (5.3), (5.4), (5.5), and (5.6) have analogues in any group G. For a subgroup H of G, every element of G lies in some left H-coset, and different left H-cosets are disjoint, so we can write G as a union of disjoint left H-cosets. Write these different cosets as g_iH, so we have a disjoint union

$$G = \bigcup_{i \in I} g_i H. \tag{5.7}$$

Here I is just an indexing set counting the number of different left H-cosets in G. Formula (5.7) is called the left H-coset decomposition of G. The particular multipliers g_i (which are nothing other than a particular choice of one element from each left H-coset) are a complete set of left coset representatives of H in G.

In a group, the passage from H to gH by left multiplication by g leads to no collapsing of terms. For instance, in Example 5.4, the cosets of $\{(1),(12)\}$ in S_3 all have size 2. Even in cases of an infinite subgroup, as in Example 2.1, the cosets all look like the original subgroup (just shifted). Let's prove such a comparison between a subgroup and any of its cosets in general.

Theorem 5.9. *Let H be a subgroup of the group G. Any left H-coset in G has a bijection with H. In particular, when H is finite, the cosets of H all have the same size as H.*

Proof. Pick a left coset, say gH. We can pass from gH to H by left multiplication by g^{-1}: $g^{-1}(gh) = h \in H$. Conversely, we can pass from H to gH by left multiplication by g. These functions from gH to H and *vice versa* are inverses to each other, showing gH and H are in bijection with each other. □

Of course there is an analogous result for right cosets, which the reader can formulate.

6. The index and Lagrange's theorem

For any integer $m \neq 0$, the number of cosets of $m\mathbf{Z}$ in $\mathbf{Z}$ is $|m|$. This gives us an interesting way to think about the meaning of $|m|$, other than its definition as "m made positive." Passing from $\mathbf{Z}$ to other groups, counting the number of cosets of a subgroup gives a useful numerical invariant.

Definition 6.1. Let H be a subgroup of a group G. The *index* of H in G is the number of left cosets of H in G. This number, which is a positive integer or ∞, is denoted $[G : H]$.

Concretely, the index of a subgroup tells us how many times we have to translate the subgroup around (on the left) to cover the whole group.

Example 6.2. Since $H = \{(1),(12)\}$ has three left cosets in S_3, by Example 5.4, $[S_3 : H] = 3$.

Example 6.3. The subgroup $H = \{1, s\}$ of D_4 has four left cosets:

$$H, \quad rH = \{r, rs\}, \quad r^2H = \{r^2, r^2s\}, r^3H = \{r^3, r^3s\}.$$

The index of H in D_4 is 4.

Example 6.4. For a positive integer m, $[\mathbf{Z} : m\mathbf{Z}] = m$, since $0, 1, \ldots, m-1$ are a complete set of coset representatives of $m\mathbf{Z}$ in $\mathbf{Z}$.

Example 6.5. What is the index of $15\mathbf{Z}$ inside $3\mathbf{Z}$? (Not inside $\mathbf{Z}$, but $3\mathbf{Z}$.) Modulo 15, a multiple of 3 is congruent to $0, 3, 6, 9$, or 12. That is, we have the disjoint union

$$3\mathbf{Z} = 15\mathbf{Z} \cup (3 + 15\mathbf{Z}) \cup (6 + 15\mathbf{Z}) \cup (9 + 15\mathbf{Z}) \cup (12 + 15\mathbf{Z}).$$

Thus $[3\mathbf{Z} : 15\mathbf{Z}] = 5$.

Example 6.6. The index $[\mathbf{R} : \mathbf{Z}]$ is infinite, by Example 5.7: there are infinitely many cosets of $\mathbf{Z}$ in $\mathbf{R}$.

Remark 6.7. If we allow ourselves to use the language of cardinal numbers, which permits a distinction between different orders of infinity, then we could define $[G : H]$ as the cardinality of the set of left H-cosets. This would have no effect on the meaning of a finite index, but would allow a more refined meaning in the case of an infinite index. Since we will not have much use for the index concept when it is infinite, we stick with the more concrete approach in Definition 6.1.

In the case of finite groups, there is a simple formula for the index of a subgroup.

Theorem 6.8. *When G is a finite group, and H is a subgroup, $[G : H] = |G|/|H|$.*

For example, this formula says the index of $\{(1), (12)\}$ in S_3 is $6/2 = 3$. Compare with the computation of the same index in Example 6.2. Theorem 6.8 lets us read off the index of the subgroup just from knowing the size of the subgroup, without actual coset constructions.

Proof. Since G is finite, H has finitely many left cosets in G. Let $t = [G : H]$, and write the different left cosets of H as $g_1H, \dots, g_tH$. We know that any two left cosets of H are the same or are disjoint. Therefore we have

$$G = g_1H \cup \cdots \cup g_tH, \tag{6.1}$$

where the union is disjoint. By Theorem 5.9, each left H-coset has the same size as H, so computing the size of both sides of (6.1) tells us

$$|G| = t|H|, \tag{6.2}$$

Thus $[G : H] = t = |G|/|H|$. □

Using the formula for the index as a ratio, we get the next "transitivity" result quite easily.

Theorem 6.9. *In a finite group G, indices are multiplicative in towers: for subgroups $K \subset H \subset G$,*

$$[G : K] = [G : H][H : K].$$

Proof. Theorem 6.8 implies

$$[G : H][H : K] = \frac{|G|}{|H|} \cdot \frac{|H|}{|K|} = \frac{|G|}{|K|} = [G : K].$$

□

While Theorem 6.9 is only stated for finite G, the formula does hold for infinite groups containing subgroups with finite index (but our proof of Theorem 6.9 does not make sense for infinite groups). Consider $G = \mathbf{Z}$, $H = 3\mathbf{Z}$, and $K = 15\mathbf{Z}$: $[G : K] = 15$, $[G : H] = 3$ and $[H : K] = 5$.

In the proof of Theorem 6.8, we found at the end that $|H|$ divides $|G|$. This is called Lagrange's theorem.

Theorem 6.10 (Lagrange, 1771). *In any finite group, the size of any subgroup divides the size of the group.*

Proof. Let G be a finite group, and H a subgroup. By (6.2), $|H|$ divides $|G|$. In fact, we found the ratio $|G|/|H|$ counts the number of left H-cosets in G. $\square$

We will see some applications of Lagrange's theorem in the next section.

The converse of Lagrange's theorem is true for some groups (*e.g.*, all cyclic groups), but it is false in general: given a divisor of the size of the group, there need not be a subgroup with that size.

Example 6.11. The smallest counterexample is A_4, which has size 12. While A_4 has subgroups of size 1,2,3,4, and 12, it has no subgroup of size 6.

To prove there is no subgroup of size 6, we argue by contradiction. Suppose there is a subgroup of A_4 with size 6, say H. Then $[A_4 : H] = 12/6 = 2$. We will show for each $g \in A_4$ that $g^2 \in H$.

If $g \in H$ then clearly $g^2 \in H$. If $g \notin H$ then gH is a left coset of H different from H (since $g \in gH$ and $g \notin H$), so from $[G : H] = 2$ the only left cosets of H are H and gH. Which one is g^2H? If $g^2H = gH$ then $g^2 \in gH$, so $g^2 = gh$ for some $h \in H$, and that implies $g = h$, so $g \in H$, but that's a contradiction. Therefore $g^2H = H$, so $g^2 \in H$.

Every 3-cycle (abc) in A_4 is a square: (abc) has order 3, so $(abc) = (abc)^4 = ((abc)^2)^2$. Thus H contains all 3-cycles in A_4. The 3-cycles are

$$(123),\ (132),\ (124),\ (142),\ (134),\ (143),\ (234),\ (243)$$

and that is too much since there are 8 of them while $|H| = 6$. Hence H does not exist.

Example 6.12. The next counterexample to the full converse of Lagrange's theorem, in terms of size, is $\mathrm{SL}_2(\mathbf{Z}/(3))$. This group has size 24. It has subgroups of size 1, 2, 3, 4, 6, 8, and 24, but no subgroup of size 12.

To see why there is no subgroup of size 12, assume otherwise. Any subgroup H of size 12 in $\mathrm{SL}_2(\mathbf{Z}/(3))$ has index 2, so just as in the previous example one can show any square in $\mathrm{SL}_2(\mathbf{Z}/(3))$ must lie in H. A tedious count shows there are 10 squares in $\mathrm{SL}_2(\mathbf{Z}/(3))$. Since $|H| = 12$, we don't yet have a contradiction. To reach a contradiction, we use the two squares $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)^2$ and $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)^2$. They must be in H, so their products $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ lie in H. These products turn out not to be squares, so they fill up the available space in H. The inverses of these products also have to lie in H, but an explicit calculation shows the inverses of these two products are matrices not yet taken into account. Thus H is too small, so H does not exist.

Example 6.13. Quite generally, A_n does not have a subgroup of size $n!/4$ (that is, of index 2) for $n \geq 4$.

Remark 6.14. There is no simple dividing line between those finite groups that satisfy the converse to Lagrange's theorem (for each divisor, there is a subgroup of that size) and those finite groups that do not. Three classes of finite groups that satisfy the converse include abelian groups, dihedral groups, and groups of prime-power size. (These are all special cases of "supersolvable" groups, and all supersolvable groups satisfy the converse of Lagrange's theorem, but some non-supersolvable groups also satisfy the converse of Lagrange's theorem.)

Since we defined the index of a subgroup as the number of its left cosets, presumably we need to introduce a corresponding index for the count of right cosets. However, an extra "right index" concept is not necessary, as we now explain.

Theorem 6.15. *For any subgroup H of G, there are as many left H-cosets as right H-cosets.*

Proof. We will give two proofs, one in the case of finite G and the other in the case of general G.

The "right" analogue of Theorem 5.9 says that any right H-coset is in bijection with H. Therefore, supposing G is finite, running through the proof of Theorem 6.8 with right cosets in place of left cosets shows $|G|/|H|$ is a formula for the number of right H-cosets in G. We already saw this is a formula for the number of left H-cosets, so the number of left and right H-cosets is the same.

Now suppose G is an arbitrary group, possibly infinite. We want to give a bijection between the collections of left and right H-cosets. One's first guess, to send gH to Hg, is *not well-defined*. For instance, taking $G = S_3$ and $H = \{(1), (12)\}$, we have

$$(13)H = (123)H = \{(13), (123)\},$$

but

$$H(13) = \{(13), (132)\}, \quad H(123) = \{(23), (123)\}.$$

So passing from gH to Hg depends on the coset representative g, which means it makes no sense as a function from cosets to cosets. (If you want to remember an example of an attempt to define a function that is in fact *not* well-defined, this is it.)

The correct way to turn left cosets into right cosets is to use inversion. For any subset $S \subset G$, let $S^{-1} = \{s^{-1} : s \in S\}$. For instance, since subgroups are closed under inversion, check that $H^{-1} = H$ and $(H^{-1})^{-1} = H$. If we invert a left coset gH, we obtain

$$(gH)^{-1} = H^{-1}g^{-1} = Hg^{-1}.$$

Similarly, $(Hg)^{-1} = g^{-1}H$. The function $f(gH) = (gH)^{-1}$ gives a bijection between left and right H-cosets. $\square$

Theorem 6.15 does not say every right coset is a left coset, but only that the number of each kind of coset is the same. For instance, we saw in Example 5.4 that the different left cosets of $H = \{(1), (12)\}$ in S_3 are

$$\{(1), (12)\}, \quad \{(13), (123)\}, \quad \{(23), (132)\}.$$

The right coset $H(13) = \{(13), (132)\}$ is not the same as any of these. However, the collection of right H-cosets has 3 members:

$$\{(1), (12)\}, \quad \{(13), (132)\}, \quad \{(23), (123)\}.$$

7. Applications of Lagrange's theorem

Lagrange's theorem leads to group-theoretic explanations of some divisibility properties of integers. The idea is this: to prove $a \mid b$, find a group of size b containing a subgroup of size a. After illustrating this in a few cases, we will use Lagrange's theorem to extract information about subgroups of a group.

Theorem 7.1. *Binomial coefficients are integers: for $n \geq 1$ and $0 \leq m \leq n$, the ratio*

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

is an integer.

Proof. We are going to write down a subgroup of S_n with size $m!(n-m)!$.

Consider the permutations in S_n that separately permute the subsets $\{1, \dots, m\}$ and $\{m+1, \dots, n\}$:

$$H = \{\sigma \in S_n : \sigma \text{ permutes } \{1, 2, \dots, m\} \text{ and } \{m+1, \dots, n\}\}.$$

The reader can check H is a subgroup, and $|H| = m!(n-m)!$. If the reader is concerned about the degenerate cases $m = 0$ and $m = n$, where one of the conditions defining H is an empty condition, note these cases of the theorem are easily checked directly. □

Theorem 7.2. *For positive integers a and b, the ratios*

$$\frac{(ab)!}{(a!)^b}, \quad \frac{(ab)!}{(a!)^b b!}$$

are integers.

Proof. Write out the integers from 1 to ab in groups of a consecutive integers:

$$1, 2, \dots, a; a+1, \dots, 2a; 2a+1, \dots, 3a; \dots; (b-1)a+1, \dots, ba.$$

There are b of these sets, separated by semi-colons. The permutations of $1, \dots, ba$ that permute each set within itself is a subgroup of S_{ab} and has size $(a!)^b$. Therefore $(a!)^b \mid (ab)!$.

The stronger divisibility relation $(a!)^b b! \mid (ab)!$ can be explained by finding a subgroup of S_{ab} with size $(a!)^b b!$. This is left as an exercise for the reader. Hint: The new subgroup of S_{ab} should contain the one constructed in the previous paragraph. □

Theorem 7.3. *For $m \geq 1$, let $\varphi(m) = |(\mathbf{Z}/(m))^\times|$ be the number of invertible numbers modulo m. For $m \geq 3$, $\varphi(m)$ is even.*

Here is a table illustrating the evenness of $\varphi(m)$ once $m \geq 3$. Can you find an explanation yourself?

m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\varphi(m)$	1	1	2	2	4	2	6	4	6	4	10	4	12	6	8

Proof. For $m \geq 3$, $\{\pm 1\}$ is a subgroup of $(\mathbf{Z}/(m))^\times$ with size 2, so $2 \mid \varphi(m)$ by Lagrange. □

Now we turn to applications of Lagrange's theorem in group theory itself. (The previous application to $\varphi(m)$ is more properly considered an application to number theory than to group theory.) In each application below, pay attention to the way Lagrange's theorem is used.

Theorem 7.4. *Let G be a finite group and let H and K be subgroups with relatively prime size. Then $H \cap K = \{e\}$.*

Proof. Since $H \cap K$ is a subgroup of both H and K, its size divides both $|H|$ and $|K|$ by Lagrange. Therefore, by relative primality of these sizes, $|H \cap K| = 1$, so $H \cap K = \{e\}$. □

In the handout on orders of elements in a group, we proved the order of every element in a finite abelian group divides the size of the group. Now we can show this for all finite groups.

Theorem 7.5. *Let G be a finite group. If g has order n, then $n \mid |G|$. In particular, $g^{|G|} = e$ for every $g \in G$.*

Proof. Let $H = \langle g \rangle$ be the subgroup generated by g. The size of H is exactly the order of g, so $|H| = n$. By Lagrange, $n \mid |G|$. Since $g^n = e$ and $n \mid |G|$, we get $g^{|G|} = e$. $\square$

It is interesting to compare this method of showing the order of g divides $|G|$ with the proof from the abelian case. In the abelian case, we first proved $g^{|G|} = e$ and then deduced the order of each element is a factor of $|G|$. In Theorem 7.5, we see that the proof in the general case goes the other way: first show each element has order dividing $|G|$ and then conclude $g^{|G|} = e$ for each g.

Corollary 7.6. *If G is a finite group and k is an integer relatively prime to $|G|$, then the k-th power function on G is invertible.*

This is not saying the k-th power function is multiplicative on G, but only that it is a bijection.

Proof. By Bezout, when $(k, |G|) = 1$ we can write $k\ell + |G|m = 1$ for some integers ℓ and m. Therefore

$$g = g^1 = g^{k\ell}(g^{|G|})^m = g^{k\ell}$$

since $g^{|G|} = 1$. So we see how to invert the k-th power function: raise to the power ℓ, where $k\ell \equiv 1 \bmod |G|$. $\square$

Example 7.7. When $G = D_5$, which has order 10, let's raise all the elements to the third power. The result is a rearrangement of the elements of the group. See the table below. (Reflections don't change since an odd power of a reflection is itself.)

g	1	r	r^2	r^3	r^4	s	rs	r^2s	r^3s	r^4s
g^3	1	r^3	r	r^4	r^2	s	rs	r^2s	r^3s	r^4s

Since $3 \cdot 7 \equiv 1 \bmod 10$, the seventh power is the inverse of the third power: $(g^3)^7 = g^{21} = g \cdot (g^{10})^2 = g$. For example, $(r^4)^3 = r^{12} = r^2$ and $(r^2)^7 = r^{14} = r^4$.

What about the converse of Corollary 7.6? That is, is the k-th power function on G invertible *only* when k is relatively prime to $|G|$? This can be answered using Cauchy's theorem, but we don't do that here.

Corollary 7.8. *Any group of prime size is cyclic, and in fact any non-identity element is a generator.*

Proof. Let G be the group, with $p = |G|$. Pick any non-identity element g from G. By Theorem 7.5, the order of g divides p and is greater than 1, so g has order p. Therefore $|\langle g \rangle| = p$, so $\langle g \rangle = G$. $\square$

Remark 7.9. Be careful not to mis-apply Corollary 7.8. While it tells us that the additive group $\mathbf{Z}/(p)$ is cyclic (any non-zero number modulo p is an additive generator), it does not tell us anything about the nature of the multiplicative group $(\mathbf{Z}/(p))^\times$, which has size $p-1$. The group $(\mathbf{Z}/(p))^\times$ *is* cyclic for any prime p, but the proof of that is not as simple as for $\mathbf{Z}/(p)$. It certainly does not come from Corollary 7.8, since $(\mathbf{Z}/(p))^\times$ does not have prime size (when $p > 3$).

Corollary 7.10. *Let p and q be primes. Any non-trivial proper subgroup of a group of size pq is cyclic.*

Proof. A proper subgroup will have size equal to a non-trivial proper factor of pq, which is either p or q. In either case, the subgroup has prime size and therefore is cyclic by Corollary 7.8. $\square$

For instance, if we want to find all the subgroups of a group of size 6 (or 15 or 21...), we can compute the cyclic group generated by each element. Every non-trivial proper subgroup will arise in this way, and we then throw in the trivial group and the whole group to complete the list.

Here is an application of Lagrange's theorem to a characterization of cyclic groups from combinatorial information about their subgroups.

Theorem 7.11. *A finite group that has at most one subgroup of any size is cyclic.*

Proof. I learned this proof from F. Ladisch.

Let G be a group with at most one subgroup of any size, and let n be its order. We will count the elements of G by *collecting together all elements that generate the same subgroup.* That is, for each cyclic subgroup H we will collect together all of its generators and then add up the number of these generators as H varies. This accounts for each element of G exactly once (each element generates some cyclic subgroup), so

$$n = \sum_{\text{cyc. subgp. } H} |\{\text{generators of } H\}|.$$

Every subgroup of G has order dividing n, by Lagrange's theorem, so we can sum over cyclic subgroups of G by summing over cyclic subgroups of order d for each d dividing n:

$$n = \sum_{d|n} \sum_{\text{cyc.} H, |H|=d} |\{\text{generators of } H\}|.$$

All cyclic groups of order d have the same number of generators: if $\langle h \rangle$ has order d then each element of $\langle h \rangle$ is h^a for a unique a between 1 and d, and h^a generates $\langle h \rangle$ if and only if $(a, d) = 1$. Thus the number of generators in a cyclic group of order d is $|\{1 \le a \le d : (a, d) = 1\}| = \varphi(d)$, so

$$n = \sum_{d|n} \sum_{\text{cyc.} H, |H|=d} \varphi(d) = \sum_{d|n} \varphi(d) |\{H \subset G : |H| = d, H \text{ cyclic}\}|. \tag{7.1}$$

This equation is valid for *any* finite group G of order n.

In the particular group $\mathbf{Z}/(n)$ of order n, which is cyclic, there is a unique cyclic subgroup of every order dividing n, so (7.1) for $G = \mathbf{Z}/(n)$ becomes

$$n = \sum_{d|n} \varphi(d). \tag{7.2}$$

If G is a group of order n and it has at most one subgroup of order d for each d dividing n, then $|\{H \subset G : |H| = d, H \text{ cyclic}\}| = 0$ or 1 for all d dividing n, so

$$n = \sum_{d|n} \varphi(d) |\{H \subset G : |H| = d, H \text{ cyclic}\}| \le \sum_{d|n} \varphi(d) \overset{(7.2)}{=} n. \tag{7.3}$$

Therefore $|\{H \subset G : |H| = d, H \text{ cyclic}\}| = 1$ for all d dividing n: if the count were ever 0 then the inequality in (7.3) would be strict and hence $n < n$, a contradiction. Since G has a cyclic subgroup of order d for each d dividing n, taking $d = n$ shows G is cyclic. $\square$

The only subgroups used in this proof are cyclic, so we proved something slightly stronger than what was stated in the theorem: a finite group that has at most one *cyclic* subgroup of any size is cyclic.

Corollary 7.12. *Let G be a finite group such that, for each d dividing $|G|$, the equation $x^d = 1$ in G has at most d solutions. Then G is cyclic.*

Proof. We will show G has at most one subgroup of any size. If H is a subgroup of G with order d, then every element of H satisfies the equation $x^d = 1$ and thus, by hypothesis, H is the complete set of solutions in G to the equation $x^d = 1$. This shows there is at most one subgroup of order d: if H and H' are both subgroups of order d then H and H' both equal $\{g \in G : g^d = 1\}$, so $H = H'$. □

We conclude our list of applications of Lagrange's theorem with a result that stands in some sense "dual" to Theorem 7.4: instead of the subgroups having relatively prime size, and getting information about the size of the intersection, we look at what happens when the subgroups have relatively prime index. What can be said about the index of their intersection?

Theorem 7.13. *Let G be a finite group, with subgroups H and K. Set $m = [G : H]$ and $n = [G : K]$. Then*

$$[m, n] \leq [G : H \cap K] \leq mn.$$

In particular, if m and m are relatively prime, then $[G : H \cap K] = mn = [G : H][G : K]$.

Proof. Since $H \cap K \subset H \subset G$ and $H \cap K \subset K \subset G$, Theorem 6.9 tells us

$$[G : H \cap K] = [G : H][H : H \cap K] = [G : K][K : H \cap K].$$

Thus m and n each divide $[G : H \cap K]$, so their least common multiple divides $[G : H \cap K]$ as well.

Now we want to show $[G : H \cap K] \leq mn$. Writing this as

$$[G : H][H : H \cap K] \leq [G : H][G : K],$$

our desired inequality is the same as

$$[H : H \cap K] \leq [G : K]. \tag{7.4}$$

The number $[G : K]$ counts how many left K-cosets are in G. Every left K-coset has the form gK for some $g \in G$. Among these, some cosets contain elements of H. How many? We will show there are $[H : H \cap K]$ such cosets, and thus we obtain the inequality (7.4).

The point is, for h_1 and h_2 in H, that

$$h_1 K = h_2 K \Longleftrightarrow h_1(H \cap K) = h_2(H \cap K). \tag{7.5}$$

(If this is true, then it tells us the number of left K-cosets of G represented by an element of H is the same as the number of left $(H \cap K)$-cosets in H, since $H \cap K$ is a subgroup of H.) Well, on the left side of (7.5), there is such equality exactly when $h_1 = h_2 k$ for some $k \in K$. Then $k = h_2^{-1} h_1$ lies in H as well, so $k \in H \cap K$. But then the equation $h_1 = h_2 k$ tells us $h_1(H \cap K) = h_2(H \cap K)$. The reverse implication in (7.5) is left as an exercise for the interested reader. □

8. Left and right coset spaces

The usefulness of modular arithemetic (in number theory, cryptography, *etc.*) suggests the possibility of carrying it over from $\mathbf{Z}$ to other groups. After all, since congruence classes modulo $m\mathbf{Z}$ are just a special instance of a coset decomposition, why not think about (left) cosets for subgroups of groups other than $\mathbf{Z}$ as a generalization of $\mathbf{Z}/(m)$?

Definition 8.1. Let G be a group and H be a subgroup. The *left coset space* $G/H := \{gH : g \in G\}$ is the set consisting of the left H-cosets in G.

A left coset space is a set whose elements are the left cosets of a subgroup. It is a "set of sets," as the following examples illustrate.

Example 8.2. In the group S_3, let $H = \{(1),(12)\}$. By (5.2), there are 3 left H-cosets:

$$\begin{aligned} S_3/H &= \{H, (13)H, (23)H\} \\ &= \{\{(1),(12)\}, \{(13),(123)\}, \{(23),(132)\}\}. \end{aligned}$$

Example 8.3. In $\mathrm{Aff}^+(\mathbf{R})$, with $K = \{\left(\begin{smallmatrix} x & 0 \\ 0 & 1 \end{smallmatrix}\right) : x > 0\}$, each left K-coset contains one matrix of the form $\left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$, so

$$\mathrm{Aff}^+(\mathbf{R})/K = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} K : b \in \mathbf{R} \right\}.$$

Example 8.4. In $\mathbf{R}$, every $\mathbf{Z}$-coset is represented by a real number in $[0,1)$, so (5.5) tells us

$$\mathbf{R}/\mathbf{Z} = \{x + \mathbf{Z} : 0 \le x < 1\}.$$

Can we make G/H into a group using the group operation from G, in the same way that $\mathbf{Z}/(m)$ inherits an additive group structure from that of $\mathbf{Z}$? The natural idea would be to multiply two left cosets by multiplying two representatives:

$$g_1H \cdot g_2H := g_1g_2H.$$

Unfortunately, such an operation does not always make sense. That is, the coset on the right can change if we change the coset representatives on the left without changing the cosets themselves.

Example 8.5. Take $G = S_3$ and $H = \{(1),(12)\}$. We have

$$(13)H = (123)H = \{(13),(123)\}, \quad (23)H = (132)H = \{(23),(132)\},$$

but

$$(13)(23)H = (132)H, \quad (123)(132)H = H,$$

and $(132)H \neq H$, so changing the coset representatives changed the 'product' coset.

In Example 8.5, multiplication of left cosets is *not well-defined*: it depends on the choice of coset representative, and thus is not an honest operation at the level of cosets.

When G is abelian, this kind of difficulty does not arise: if $g_1H = g_1'H$ and $g_2H = g_2'H$, then $g_1g_2H = g_1'g_2'H$. One can proceed and turn G/H into a group, just as $\mathbf{Z}/(m)$ is a group for addition. When G is non-abelian (the far more typical case), there are special kinds of subgroups, called *normal* subgroups, for which multiplication of left cosets is a well-defined operation in terms of representatives, and one can make G/H into a group by multiplying coset representatives.

Why should we care about trying to make G/H into a group? There are at least three reasons:

- The usefulness of modular arithmetic in $\mathbf{Z}/(m)$ suggests having an analogous construction for other groups has got to be worthwhile, even if it is not yet clear what we might be able to do with the construction. (Note: While we can both add and multiply in $\mathbf{Z}/(m)$, we only have in mind that G/H could be a group, inheriting the one operation on G, analogous just to addition in $\mathbf{Z}/(m)$.)
- Having groups of the form G/H is a method of constructing new groups out of old ones (by replacing G with G/H). This could allow us to find alternate models for certain kinds of groups, which might be more convenient to use than other models.
- When G is finite and H is a non-trivial proper subgroup of G, both H and G/H have size less than the size of G. If G/H is a group, then we have two groups related to G that have size less than $|G|$. This turns out to be very useful in proving theorems about finite groups by induction on the size of the group.

Remark 8.6. In addition to the left coset space G/H, one can also consider the right H-coset space

$$H\backslash G := \{Hg : g \in G\}.$$

9. Double cosets

A coset can be viewed as the result of multiplying an element on one side by a subgroup. Allowing multiplying on both sides (by possibly different subgroups) leads to the idea of a double coset.

Definition 9.1. Let G be a group and H and K be subgroups. Any set of the form

$$HgK = \{hgk : h \in H, k \in K\}$$

is called an (H, K) *double coset*, or simply a double coset if H and K are understood.

While special instances of double cosets appeared in the work of Cauchy, their systematic consideration in group theory is due to Frobenius.

In this section we will look at some examples of double cosets and see how they are similar to and different from left and right cosets.

Example 9.2. Let's look at double cosets in S_3, with $H = \{(1), (12)\}$ and $K = \{(1), (13)\}$. The double coset of (1) is

$$H(1)K = HK = \{hk : h \in H, k \in K\} = \{(1), (12), (13), (132)\}$$

and the double coset of (23) is

$$\begin{aligned} H(23)K &= \{(1)(23)(1), (1)(23)(13), (12)(23)(1), (12)(23)(13)\} \\ &= \{(23), (123)\}. \end{aligned}$$

Unlike a coset, we now see that a double coset does not have to have size dividing the size of the group (4 does not divide 6). Moreover, different double cosets for the same pair of subgroups can have different sizes (such as 4 and 2).

The reader can compute the (H, K)-double cosets of the other four elements of S_3 and find only the two examples above repeating. (Or appeal to Theorem 9.7 below.)

Example 9.3. Consider the group S_3 again, but now use $H = K = \{(1), (12)\}$. There are only two different double cosets:

$$H(1)H = HH = \{(1), (12)\}, \quad H(13)H = \{(13), (23), (123), (132)\}.$$

Example 9.4. Take $G = D_4 = \langle r, s\rangle$ and $H = K = \{1, s\}$. The different (H, H) double cosets are

$$HH = H = \{1, s\}, \quad HrH = \{r, rs, r^3s, r^3\}, \quad Hr^2H = \{r^2, r^2s\}.$$

This example shows the number of double cosets for a pair of subgroups need not divide the size of the group (3 does not divide 8). This is a contrast to what happens with left and right cosets of a subgroup.

Example 9.5. An ordinary coset is a special kind of double coset, where one of the subgroups is trivial: $Hg\{e\} = Hg$ and $\{e\}gK = gK$.

Example 9.6. If G is abelian, then the product set HK is a subgroup of G and an (H, K) double coset is just an ordinary coset of the subgroup HK.

Theorem 9.7. *Fix two subgroups H and K of the group G. Every element of G lies in some (H, K) double coset, and any two (H, K) double cosets that overlap are equal. Equivalently, different (H, K) double cosets are disjoint and they collectively cover the whole group.*

Proof. Clearly $g \in HgK$: use $h = e$ and $k = e$. Thus, every element of G is in some (H, K) double coset.

Now assume HgK and $Hg'K$ overlap:

$$hgk = h'g'k', \tag{9.1}$$

where h and h' belong to H and k and k' belong to K. We want to prove $HgK = Hg'K$. By (9.1) $g = h^{-1}h'g'k'k^{-1}$, so g lies in $Hg'K$. For any h'' in H and k'' in K,

$$h''gk'' = h''h^{-1}h'gk'k^{-1}k'' \in Hg'K.$$

Letting h'' and k'' vary, we obtain $HgK \subset Hg'K$. The reverse inclusion is proved in the same way, so $HgK = Hg'K$. □

As a generalization of left or right coset spaces from Appendix 8, we can contemplate a double coset space

$$H\backslash G/K = \{HgK : g \in G\}.$$

This is a left or right coset space when H or K is trivial. Inversion of elements turns any left coset space G/H into a right coset space $H\backslash G$. Similarly, inversion on $H\backslash G/K$ turns it into $K\backslash G/H$ and *vice versa*.

We noted in Example 9.4 that the number of (H, K) double cosets does not have to divide the size of the group. However, Frobenius did find a formula involving double coset spaces that did exactly generalize the "index formulas" $|G| = |H||G/H| = |H||H\backslash G|$, as follows.

Theorem 9.8. *Let G be a finite group and H and K be subgroups. Then*

$$|\{(h, g, k) \in H \times G \times K : hgk = g\}| = |H||K||H\backslash G/K|.$$

When one of the subgroups is trivial, this reduces to the index formula for the other subgroup.