

# GROUP CODES AND THEIR DECODING

## 1. SETUP

We consider a finite group  $\mathbf{G}$  acting on a set  $X$  (maybe faithfully).

Choose a sequence of subgroups

$$\{I\} = \mathbf{G}_0 < \mathbf{G}_1 < \mathbf{G}_2 < \cdots < \mathbf{G}_m = \mathbf{G}$$

and for each  $k$  a linearly ordered set  $S_k$  of generators for  $\mathbf{G}_k$ .

(Perhaps this might be done randomly, but what we have in mind is adding one generator to the end of the list at each step, and removing the redundant ones. The case when we add one generator from an irredundant set at each step corresponds to the “parabolic subgroup” case; putting some restriction on the generators is like the “reflection subgroup” case; and no restrictions would correspond to arbitrary subgroups.)

No matter how we choose coset leaders, the following representation holds.

**Theorem 1.** *Let  $\mathbf{G}$  be a group, and let*

$$\{I\} = \mathbf{G}_0 < \mathbf{G}_1 < \cdots < \mathbf{G}_{m-1} < \mathbf{G}_m = \mathbf{G}$$

*be a sequence of subgroups. Choose (left) coset leaders for each  $\mathbf{G}_i$  over  $\mathbf{G}_{i-1}$ , with  $I$  as the coset leader for  $\mathbf{G}_{i-1}$ . Then every element of  $\mathbf{G}$  has a unique expression as a product of coset leaders,  $g = c_m \dots c_1$ , with each  $c_i$  a coset leader for  $\mathbf{G}_i$  over  $\mathbf{G}_{i-1}$ .*

So we consider recursively the situation where we have a subgroup  $\mathbf{H} \leq \mathbf{G}$  and a linearly ordered set of generators  $A, B, C, \dots$  for  $\mathbf{G}$ .

## 2. COSET LEADERS, GRAPHS AND NORMAL FORM

We summarize a standard way to find coset leaders, graphs and spanning trees.

Define an *expression* to be a finite sequence of generators. Recall that an expression is *reduced* if it has minimum length. Define the order for expressions as follows:  $E_1 < E_2$  if the number of factors in  $E_1$  is less than the number of factors in  $E_2$ , or if both have the same number of factors and  $E_1$  precedes  $E_2$  in dictionary order.

We can choose the *coset leader* of a left coset  $g\mathbf{H}$  to be the group element represented by the smallest expression that evaluates to an element of the coset.

Next we will define an algorithm for making a list of expressions. Initially the list contains only the identity element, an expression with no elements.

For each expression  $E$  in the list, in order, multiply it on the left by each generator  $R$  in order, and then check whether the result is in the same coset as any expression on the list. If it is not, then add  $RE$  to the end of the list.

**Theorem 2.** *The list consists exactly of all the coset leader expressions for  $\mathbf{G}$  over  $\mathbf{H}$ .*

Now let us construct the graph  $\Gamma$  of (left) coset leaders for  $\mathbf{G}$  over  $\mathbf{H}$ . The vertices of  $\Gamma$  are the coset leaders of  $\mathbf{G}$  over  $\mathbf{H}$ . There is a directed edge from  $v_i$  to  $v_j$  if  $v_j = Rv_i$  for a generator  $R$  and  $\ell(v_j) = \ell(v_i) + 1$ . Label such an edge by  $R$ , so that the transformation  $v_j$  can be reconstructed by tracing a path from  $v_0$  to  $v_j$  and reading the edge labels in order. Such a path need not be unique. However, we have just seen that every coset leader has a unique minimal expression in the dictionary order, and this determines a canonical path from  $v_0$  to  $v_j$ .

The *spanning tree*  $T$  for the coset leader graph  $\Gamma$  has again as vertices all the coset leaders for  $\mathbf{G}$  over  $\mathbf{H}$ . There is a directed edge from  $v_i$  to  $v_j$  in  $T$  if the minimal expression for  $v_j$  is  $RE$ , where  $R$  is a generator and  $E$  is the minimal expression for  $v_i$ .

Within the spanning tree, the alphabetic order also determines an order on the edges emanating from a given vertex. We will refer to this as the *branch order* of  $T$ .

Spanning trees are used to navigate through the group in encoding and decoding. Now we observe that the spanning tree has a technical property that is crucial to the decoding algorithm.

**Lemma 3.** *Let  $u$  and  $w$  be vertices in  $\Gamma$ . If there is a path in  $\Gamma$  going from  $u$  to  $w$ , then the (unique) path from  $u$  to  $w$  in the spanning tree  $T$  goes through the successor of  $u$  that is least in the branch order and lies on some path from  $u$  to  $w$ .*

In fact, it is not necessary to use the alphabetical order to determine the spanning tree and the branch order. Any spanning tree and branch order with the property of Lemma 3 will work for the decoding algorithm.

### 3. A COMMENT ON LENGTHS

Suppose we have an element  $g = c_m \dots c_1$  written as a product of coset leaders. There is a length function  $\ell_{\mathbf{G}}(g)$ , which is the length of the shortest expression representing  $g$  in terms of the chosen set of generators for  $\mathbf{G}$ . A more appropriate length function in our setting is

$$\ell^*(g) = \sum_{i=1}^m \ell_{\mathbf{G}_i}(c_i)$$

using the chosen generators for  $\mathbf{G}_i$  at each stage. It isn't clear how  $\ell^*(g)$  compares with  $\ell(g)$ , though it may not matter.

If  $\mathbf{H}$  is a reflection subgroup of a reflection group  $\mathbf{G}$  and  $g \in \mathbf{H}$ , then  $\ell_{\mathbf{H}}(g) \leq \ell_{\mathbf{G}}(g)$ . Indeed, the length of  $g$  is the number of reflecting planes

separating say  $\mathbf{x}_0$  and  $g\mathbf{x}_0$ , and for a reflection subgroup, the reflecting planes of  $\mathbf{H}$  are a subset of those of  $\mathbf{G}$ . When is this true in general? and does it matter?

#### 4. A CRUCIAL PROPERTY

Here are the critical properties of reflection groups that were used.

**Theorem 4.** *If  $\mathbf{H}$  is a reflection subgroup of a reflection group  $\mathbf{G}$ , then every left coset  $x\mathbf{H}$  has a unique shortest element (which we can choose as the coset leader).*

**Theorem 5.** *If  $\mathbf{G}$  is a reflection group and  $\mathbf{H}$  is a reflection subgroup, and  $L$  is the coset leader of the coset  $L\mathbf{H}$ , and if  $S_\alpha$  is a fundamental reflection, then the coset leader of the coset  $S_\alpha L\mathbf{H}$  is either  $S_\alpha L$ , or else it is  $L$ . Note that in the former case, the length of the coset leader  $S_\alpha L$  is the length of  $L$  plus or minus one.*

**Theorem 6.** *Let  $\mathbf{G}$  be a reflection group with a sequence of reflection subgroups*

$$\{I\} = \mathbf{G}_0 < \mathbf{G}_1 < \cdots < \mathbf{G}_{m-1} < \mathbf{G}_m = \mathbf{G}.$$

*If the canonical form of  $u$  as a product of coset leaders is  $c_m \dots c_1$  and  $v = S_\alpha u$  for some fundamental reflection  $S_\alpha$ , then the canonical form of  $v$  is  $c'_m \dots c'_1$  where  $c'_i = c_i$  for all but one  $i$ , and for that index  $c'_j$  is an immediate predecessor or successor of  $c_j$  in the coset leader graph  $\Gamma_j$ .*

Now the first theorem is not true for non-reflection subgroups, but that is not a problem unless there are lots of minimal-length elements in a coset. The third theorem is a consequence of the proof of the second. What we need to hold is this formulation of the second theorem.

**Property 1.** *Let  $\mathbf{H} \leq \mathbf{G}$  with fixed generating sets for  $\mathbf{H}$  and  $\mathbf{G}$ . Let  $a$  be a generator of  $\mathbf{G}$  and  $c$  a coset leader. Then the coset leader for  $ac\mathbf{H}$  is either  $ac$ , or else it is  $c$  and  $c^{-1}ac$  is a generator for  $\mathbf{H}$ .*

Note that  $c\mathbf{H} = c(c^{-1}ac)\mathbf{H}$ .

So this is a matter of choosing subgroups and their generators properly for our purposes. That includes the following stipulation (see the description of the algorithm below): *The geometry of  $X$  should be such that the closest neighbors of a point  $u^{-1}\mathbf{x}_0$  are the points  $v^{-1}\mathbf{x}_0$  with  $v = au$  for some generator  $a$  of  $\mathbf{G}$ .* Intuitively, for groups of isometries, that just says that the generators move elements by the minimum distance.

#### 5. FUNDAMENTAL REGIONS

In this section we will show that, from an *algebraic* viewpoint, choosing fundamental regions with the necessary properties for decoding is almost trivial. Of course there are also topological or geometric considerations that we are ignoring temporarily, e.g., fundamental regions should be convex.

Given  $\mathbf{G}$  acting on  $X$  and  $\mathbf{H} \leq \mathbf{G}$ , we want to define the following terms:

- *fundamental regions*  $\text{FR}(\mathbf{H})$  and  $\text{FR}(\mathbf{G})$ ,
- other *regions* that are translates of the fundamental region,
- the *guts* of a region

so that the following properties hold:

- (1)  $\text{FR}(\mathbf{H}) \supseteq \text{FR}(\mathbf{G})$ ,
- (2) more generally each region of  $\mathbf{G}$  is contained in a region of  $\mathbf{H}$ ,
- (3)  $X$  is the union of all its  $\mathbf{H}$ -regions (or  $\mathbf{G}$ -regions),
- (4) the guts of distinct  $\mathbf{H}$ -regions (or  $\mathbf{G}$ -regions) are disjoint,
- (5) if  $\mathbf{x} \in \text{FR}(\mathbf{H})$ , then there is a (left) coset leader  $c$  such that  $c\mathbf{x} \in \text{FR}(\mathbf{G})$ ,
- (6) if  $\mathbf{x}$  is in the guts of  $\text{FR}(\mathbf{H})$ , then  $c$  is unique.

It is straightforward to see that all this works, so long as we abide by this scheme.

- $\text{FR}(\mathbf{G})$  contains at least one point from each orbit of  $\mathbf{G}$ , and exactly one for each faithful orbit.
- $\text{FR}(\mathbf{H}) = \{c^{-1}\mathbf{x} : c \text{ is a coset leader and } \mathbf{x} \in \text{FR}(\mathbf{G})\}$
- A region for  $\mathbf{H}$  is a set of the form  $h\text{FR}(\mathbf{H})$  for some  $h \in \mathbf{H}$ , and likewise for regions of  $\mathbf{G}$ .
- The guts of a region consists of the points that are in a faithful orbit of  $\mathbf{G}$ .

So any topology we have that is consistent with the above scheme should yield the critical properties (1) and (5) above.

## 6. DESCRIPTION OF THE ALGORITHM

The outline of our group coding algorithm is straightforward. The setup involves selecting a particular group  $\mathbf{G}$  and a sequence of its subgroups. The coset leader graph and its spanning tree are determined, and a binary correspondence is established. Choose an *initial vector*  $\mathbf{x}_0$  in the guts of the fundamental region of  $\mathbf{G}$ . The *code* consists of  $\mathbf{G}\mathbf{x}_0 = \{g\mathbf{x}_0 : g \in \mathbf{G}\}$ . All these things appear as subroutines or parameters in the implementation.

The message  $\mathbf{m}$  to be sent corresponds to a group element  $g = \gamma(\mathbf{m})$ . The element  $g$  has a canonical expression  $g = c_m \dots c_1$  as a product of coset leaders. We transmit the vector

$$\mathbf{x} = g^{-1}\mathbf{x}_0 = c_1^{-1} \dots c_m^{-1}\mathbf{x}_0,$$

The expression for each  $c_i$  as a product of generators is obtained from the coset leader tree  $\Gamma_i$ , and these combine to give the expression for  $g$ . The expression for  $c_i^{-1}$  is obtained by going backwards through the coset leader tree, and these are combined in reverse order to give  $g^{-1}$ .

The received vector has the form  $\mathbf{r} = \mathbf{x} + \mathbf{n}$  where  $\mathbf{n}$  represents channel noise. Hopefully,  $\mathbf{r}$  will be in the same region as the transmitted vector  $\mathbf{x}$ , or if not, in a neighboring region. We decode by finding the sequence of coset leaders  $d_1, \dots, d_m$  such that  $d_i \dots d_1 \mathbf{r}$  is in the fundamental region of the subgroup  $\mathbf{G}_i$ . This in turn is done by going through the coset leader trees

and applying generators, so that at each step the vector obtained is closer to  $\mathbf{x}_0$  in some appropriate geometry on  $X$ , and hence to the fundamental region of  $\mathbf{G}$ , than the preceding vector. Thus the final vector  $d_m \dots d_1 \mathbf{x}$  is in the fundamental region of  $\mathbf{G}$ . We decode by taking  $g' = d_m \dots d_1$  and the received message as  $\mathbf{m}' = \gamma^{-1}(g')$ .

The geometry of  $X$  is left intentionally vague in the preceding description, but it is crucial for navigating the coset leader tree. You must be able to determine whether a vector  $a\mathbf{x}$  is closer to  $\mathbf{x}_0$  than  $\mathbf{x}$  is.

## 7. A CONCRETE EXAMPLE

In this section, we apply the above program to the complex reflection groups  $\mathbf{G}(r, 1, n)$ . These groups are wreath products, extensions of  $(\mathbb{Z}/r\mathbb{Z})^n$  by the symmetric group  $\mathbf{S}_n$ . Let  $\rho = e^{\frac{2\pi i}{r}}$ . The matrix representation of  $\mathbf{G}(r, 1, n)$  acting on the complex space  $\mathbb{C}^n$  consists of all matrices with one non-zero entry in each row and column, that entry being a power  $\rho^k$ . In particular,  $\mathbf{G}(2, 1, n)$  is the real Coxeter group  $B_n$ . The group coding scheme for  $B_n$  extends very naturally to  $\mathbf{G}(r, 1, n)$  with minimal changes and a straightforward encoding/decoding scheme.

For a (redundant) set of generators, take the transformations  $s_i$  ( $1 \leq i \leq n$ ) that multiply the  $i$ -th entry of a vector by  $\rho$ , and the transpositions  $t_j$  ( $2 \leq j \leq n$ ) that switch the  $j$ -th and  $(j-1)$ -st components.

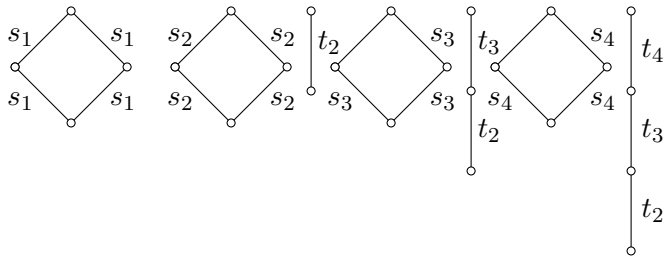
Consider the nested sequence of subgroups

$$\begin{aligned} \mathbf{G}_0 &= \{I\} \\ \mathbf{G}_1 &= \langle s_1 \rangle \\ \mathbf{G}_2 &= \langle s_1, s_2 \rangle \\ \mathbf{G}_3 &= \langle s_1, t_2 \rangle \\ \mathbf{G}_4 &= \langle s_1, t_2, s_3 \rangle \\ \mathbf{G}_5 &= \langle s_1, t_2, t_3 \rangle \\ &\dots \\ \mathbf{G}_{2n-1} &= \mathbf{G}(r, 1, n) \end{aligned}$$

which has the coset leader graphs given in Figure 1. At the even steps,  $\mathbf{G}_{2k}$  is obtained by adding adding a generator  $s_j$  that commutes with the elements of  $\mathbf{G}_{2k-1}$ . At the odd steps,  $\mathbf{G}_{2k+1}$  is obtained by adding adding a generator  $t_j$ , and the relations  $t_j s_j t_j = s_{j-1}$  make all but say  $s_1$  redundant.

The same type of calculations show that the crucial Property 1 holds for this subgroup sequence.

The fact that, in the complex case, some of the generators are not transpositions introduces a minor change in the algorithm. The generators  $s_i$  satisfy  $s_i^r = 1$ , and in terms of the algorithm one should regard  $s_i^{r-1} = s_i^{-1}$  as a coset leader of length one, as it will move the initial vector by the same

FIGURE 1. Coset leader graphs for  $\mathbf{G}(4, 1, 4)$ 

distance that  $s_i$  does. This means that in decoding a vector  $\mathbf{x}$ , at the appropriate stage one should consider  $\|\mathbf{x} - \mathbf{x}_0\|$ ,  $\|s_i\mathbf{x} - \mathbf{x}_0\|$  and  $\|s_i^{-1}\mathbf{x} - \mathbf{x}_0\|$ . If the first distance is a minimum, you consider the next subgroup. If the minimum is one of the latter two, then proceed around the cycle in that direction until you reach a minimum, at most halfway around. The halfway point, however, could be approached from either direction.

Next comes the task of choosing an appropriate initial vector. The only strict requirement is that  $\mathbf{x}_0$  not be fixed by any element of  $\mathbf{G}$ . It is not quite clear that the following chooses it optimally, but perhaps so. The initial vector should be chosen in the middle of the fundamental region for  $\mathbf{G}$ ; the fundamental regions for proper subgroups will be larger, and hence not as crucial. Thus one should require that the generators for the whole group each move the initial vector by the same amount. Mimicking the case of  $B_n$ , we assume the form

$$\mathbf{x}_0 = \langle a, a + b, a + 2b, \dots, a + (n - 1)b \rangle$$

and require that

$$\|s_1\mathbf{x}_0 - \mathbf{x}_0\| = \|t_2\mathbf{x}_0 - \mathbf{x}_0\| = \dots = \|t_n\mathbf{x}_0 - \mathbf{x}_0\|.$$

This is a straightforward computation that leads to the requirement

$$\frac{b}{a} = \sqrt{1 - \cos \frac{2\pi}{r}}.$$

Initially we set  $a = 1$ , and then normalize so that  $\|\mathbf{x}_0\| = 1$ . This works, but has not been properly justified. Note that  $\|s_i\mathbf{x}_0 - \mathbf{x}_0\|$  will be greater for  $i > 1$ .

Now we can define the fundamental regions geometrically, in a way that fits into the overall scheme. For  $\mathbf{H} = \mathbf{G}_i$  with  $0 \leq i \leq 2n - 1$ , let

$$\text{FR}(\mathbf{H}) = \{\mathbf{x} \in \mathbb{C}^n : \|h\mathbf{x} - \mathbf{x}_0\| > \|\mathbf{x} - \mathbf{x}_0\| \text{ for all } h \in \mathbf{H} - \{1\}\}.$$

The closure of the fundamental region is obtained by replacing  $>$  with  $\geq$ .

Notice that the fundamental regions depend on the choice of the initial vector  $\mathbf{x}_0$ . This is different from the real case. Consider for example the first case with  $\mathbf{H} = \langle s_1 \rangle$ . Writing  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{x}_0 = (u_1, \dots, u_n)$ ,

we have

$$\begin{aligned} \text{FR}(\langle s_1 \rangle) &= \{\mathbf{x} \in \mathbb{C}^n : \|s_1^k \mathbf{x} - \mathbf{x}_0\| > \|\mathbf{x} - \mathbf{x}_0\| \text{ for } 1 \leq k < r\} \\ &= \{\mathbf{x} \in \mathbb{C}^n : \text{Re}(x_1 \overline{u_1}) > \text{Re}(\rho^k x_1 \overline{u_1}) \text{ for } 1 \leq k < r\}. \end{aligned}$$

To simplify matters, we will take the standard choice of  $\mathbf{x}_0$  as indicated above, with  $a = 1$ .

The coset leaders were chosen algebraically, but they should have the geometric property that *g is a coset leader for G over H if and only if*

$$\|g\mathbf{x}_0 - \mathbf{x}_0\| \leq \|gh\mathbf{x}_0 - \mathbf{x}_0\|$$

for all  $h \in \mathbf{H}$ . For our particular choice of  $\mathbf{G}$  and  $\mathbf{x}_0$ , at least, this is true.

A more serious matter is to prove that the algorithm works, which can be put as follows. Assume that we are given  $\mathbf{v}$  in the closure of the fundamental region of  $\mathbf{H}$ , so that

$$\|h\mathbf{v} - \mathbf{x}_0\| \geq \|\mathbf{v} - \mathbf{x}_0\|$$

for all  $h \in \mathbf{H}$ . Find a coset leader  $c$  for  $\mathbf{G}$  over  $\mathbf{H}$  such that

$$\|d\mathbf{v} - \mathbf{x}_0\| \geq \|c\mathbf{v} - \mathbf{x}_0\|$$

for all coset leaders  $d$ . Then we need that

$$\|g\mathbf{v} - \mathbf{x}_0\| \geq \|c\mathbf{v} - \mathbf{x}_0\|$$

for all  $g \in \mathbf{G}$ . It is not clear under what circumstances this works. However, for the groups  $\mathbf{G}(r, 1, n)$ , with the given subgroups sequence and choice of initial vector, it is true.

Again let  $\mathbf{x}_0 = (u_1, \dots, u_n)$  with  $0 < u_1 < \dots < u_n$  real, and consider  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{C}^n$ . An easy calculation shows that, for  $x \in \mathbb{C}$  and  $u \in \mathbb{R}$ , the value of  $k$  that minimizes  $|\rho^k x - u|$  is the one that maximizes the real part of  $\rho^k x$ , independently of the size of  $u$ .

N.B. The preceding observation can be used to speed up the algorithm considerably. Writing  $x = |x|e^{i\theta}$ , then  $\rho^k x = |x|e^{(\frac{2\pi k}{r} + \theta)i}$ , the real part of which is maximized by making  $\frac{2\pi k}{r} + \theta$  as closed to  $2\pi$  as possible. Thus  $k$  should be chosen as the nearest integer to  $r - \frac{r\theta}{2\pi}$ .

Now consider the sequence

$$\begin{aligned} &\|\mathbf{x} - \mathbf{x}_0\| \\ &\|s_1^k \mathbf{x} - \mathbf{x}_0\| \\ &\|s_2^\ell s_1^k \mathbf{x} - \mathbf{x}_0\| \\ &\|t_2^\delta s_2^\ell s_1^k \mathbf{x} - \mathbf{x}_0\| \\ &\|s_3^m t_2^\delta s_2^\ell s_1^k \mathbf{x} - \mathbf{x}_0\| \\ &\|cs_3^m t_2^\delta s_2^\ell s_1^k \mathbf{x} - \mathbf{x}_0\| \\ &\dots \end{aligned}$$

where  $c$  is a coset leader for  $\mathbf{G}_5$  over  $\mathbf{G}_4$ , thus one of  $\{I, t_3, t_2 t_3\}$ . First  $k$  is chosen to maximize  $\text{Re}(\rho^k x_1)$ , then  $\ell$  to maximize  $\text{Re}(\rho^\ell x_2)$ . Now since

$u_1 < u_2$ , an easy calculation shows that if  $\operatorname{Re}(\rho^k x_1) > \operatorname{Re}(\rho^\ell x_2)$ , then we should apply  $s_2$ , switching the values, to minimize the distance; otherwise not. Next  $m$  is chosen to maximize  $\operatorname{Re}(\rho^m x_3)$ . Then, since  $u_1 < u_2 < u_3$ , we apply the correct coset leader  $c$  to put  $\operatorname{Re}(\rho^k x_1)$ ,  $\operatorname{Re}(\rho^\ell x_2)$ ,  $\operatorname{Re}(\rho^m x_3)$  into increasing order (insertion sort). Continue until pau.