These are some notes on error correcting codes. Two good sources for this material are

- *From Error Correcting Codes through Sphere Packings to Simple Groups*, by Thomas Thompson.
- *Sphere Packings, Lattices, and Simple Groups* by J. H. Conway and N. Sloane

Planet Math (on the internet) also some information.

1 Main Definitions

Let $\boldsymbol{F}$ be a finite set. A *code of length* n is a subset $S \subset \boldsymbol{F}^n$. The elements of S are called *codewords*. When $\boldsymbol{F} = \boldsymbol{Z}/2$, as is usually the case, a code is called a *binary code*.

The *Hamming distance* between two elements is the number of coordinates in which they differ. For instance, the Hamming distance between 101000 and 111001 is 2 because the two words differ in spots 2 and 6. We will write $H()$ for the Hamming distance.

A code is called *k-error correcting* if the Hamming distance between any two distinct codewords is at least $2k+1$. The rationale for this definition is as follows. Suppose that one intends to send the message s but ends up sending s' instead, and s' contains at most k errors. We know that $H(s, s') \leq k$. If there was some other codeword s'' such that $H(s', s'') \leq k$, then we would have $H(s, s'') \leq 2k$. This forces $s = s''$. So, if we have received s' in place of s, we can recover s. That is, we can correct up to k errors.

The code S is called a *linear* (n, k) *code* if $\boldsymbol{F}$ is a finite field and S is a k-dimensional vector subspace of $\boldsymbol{F}^n$. The intuitive idea here is that we have some isomorphism $\phi : S \to \boldsymbol{F}^k$ (often given by some coordinate projections, but not always.) You want to send some element $t \subset \boldsymbol{F}^k$. So, you send $s = \phi^{-1}(t)$. The receiver then performs $t = \phi(s)$ to recover the message. If S has some error correcting properties, then s can be sent with some errors, and one can still recover t.

The *weight* of a codeword is the number of nonzero entries. The *minimum weight* of a linear code is the minimum weight, taken over all nontrivial codewords. Thanks to the equality $H(s, t) = H(0, s - t)$, the minimum Hamming distance between any two distinct words in a linear code is the same as the minimum weight of the code.

2 Hamming's Rectangle Codes

We say *BL code* in place of *binary linear code* for the sake of brevity. Here is an example of a BL $(9, 4)$ code which is 1-error correcting.

Say that a *grid* is a 3×3 matrix with entries in $\boldsymbol{Z}/2$ such that the sum of every row and every column is congruent to 0 mod 2. Let S be the set of grids. There is a linear map $f : \boldsymbol{F}^9 \to \boldsymbol{F}^6$, which just strings out all the mod 2 sums of the rows and columns. Clearly S is the kernel of f. Hence S is a linear code.

It is clear that any nontrivial grid must have weight at least 4. Hence the minimum Hamming distance between any two words in S is 4. In particular, S is 1-error correcting.

There is a natural isomorphism from S to $\boldsymbol{F}^4$. We just write out the entries of the upper 2×2 block. knowing these entries, it is easy to fill in the rest of the grid. Here is an example.

$$\begin{bmatrix} 1 & 0 & * \\ 1 & 1 & * \\ * & * & * \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}. \tag{1}$$

This code can be improved somewhat. Suppose we consider the $(8, 4)$ BL code obtained from the one above simply by deleting the bottom right entry. The resulting code has minimum weight 3 and hence is still 1-error correcting.

You could make a similar code based on any $(m + 1) \times (n + 1)$ shape, where $\min(m, n) \geq 3$. The result would be a BL $((m + 1)(n + 1), mn)$-code that is 1-error correcting. Leaving off the bottom right entry, you would get a BL $((m + 1)(n + 1) - 1, mn)$-code that was 1-error correcting.

Generalizing to higher dimension, one could consider (for example) a $3 \times 3 \times 3$ cube of numbers over $\boldsymbol{Z}/2$. It is not hard to check that this gives a BL $(27, 8)$ code that is 3-error correcting. The minimum weight in this case is 8. The map to $\boldsymbol{F}^8$ is again obtained by stringing out the entries of the upper $2 \times 2 \times 2$ block. This code can be improved by leaving off certain entries. I'll leave this to you to work out.

One of the problems with these rectangle codes is that they are not as efficient as possible. According to Thompson, these are the first codes that Hamming discovered, but later he found better ones.

3 Hamming's $(7,4)$ code

A k-error correcting code S is called *perfect* if every element of $\boldsymbol{F}^n$ is within k of a unique element of S. Put another way, $\boldsymbol{F}^n$ is covered by disjoint Hamming k-balls centered at the points of S. (A Hamming k-ball is the set of points having Hamming distance at most k from a given point.) The codes above are not perfect.

People usually credit Richard Hamming with discovering the first perfect 1-error correcting code around 1948, but some people give priority to Marcel Golay because, technically, he published it first.

Let $\boldsymbol{F} = \boldsymbol{Z}/2$. Let $P = P^2(\boldsymbol{F})$ denote the projective plane over $\boldsymbol{F}$. We can think of $\boldsymbol{F}^7$ as the set of labellings of P by elements of $\boldsymbol{F}$. There is a canonical linear map $f : \boldsymbol{F}^7 \to \boldsymbol{F}^3$. One simply takes the mod 2 weighted sum of the coordinates. An example will clarify this. After we choose some ordering on the points of P (as indicated below), the element 1101010 gets mapped to 011, as follows.

$$\begin{array}{cc} 1 & 1:0:0 \\ 1 & 0:1:0 \\ 0 & 0:0:1 \\ 1 & 1:1:0 \\ 0 & 1:0:1 \\ 1 & 0:1:1 \\ 0 & 1:1:1 \\ - & - \\ - & \mathbf{0:1:1} \end{array} \tag{2}$$

We are simply taking the mod 2 dot product with 1101010 with each of the columns in the matrix above.

The kernel $S = \ker(f)$ is the $(7,4)$ code. The map $\phi : S \to \boldsymbol{F}^4$ is obtained by simply ignoring the labels of points with only one nonzero coordinate, and recording the rest. This map is clearly an injection, and both the domain and range have 16 elements. Hence, it is an isomorphism.

Clearly, no nontrivial element of S has weight 1 or 2. Hence, S is 1-error correcting. Given an arbitrary $s' \in \boldsymbol{F}^n$, there is a unique point $p \in P$ which has the same coordinates as $f(s')$. We just change the bit of s' corresponding to this point. The resulting word s lies in S. This shows that S is perfect. One can also see this by counting: Each hamming 1-ball has size $1 + 7 = 2^3$,

and we are placing disjoint 1-balls about 2^4 points, thereby covering all 2^7 points in $\boldsymbol{F}^7$.

4 Other Perfect 1-error Correcting codes

Let $\boldsymbol{F}$ be set with m elements. Suppose that we have a perfect 1-correcting (n, k) code based on $\boldsymbol{F}$. Then the following holds.

- The number of codewords is m^k.
- The size of the Hamming 1-balls are $(m-1)n+1$.
- The total number of elements in $\boldsymbol{F}^n$ is m^n.

Since our code partitions $\boldsymbol{F}^n$ into 1-balls, we have

$$m^k((m-1)n+1) = m^n. \tag{3}$$

Rearranging this gives

$$n = \frac{m^a - 1}{m-1}; \qquad a = n - k. \tag{4}$$

Golay discovered that one can make this work when $\boldsymbol{F} = \boldsymbol{Z}/p$, p prime, and Cocke later showed how to extend the method to all finite fields. We'll describe the general case. It is the natural generalization of the Hamming $(7, 4)$ code. To avoid trivialities, we take $a \geq 3$.

Let $\boldsymbol{F}$ be a finite field and let $P = P^a(\boldsymbol{F})$ denote the projective space of dimension $a-1$ over $\boldsymbol{F}$. When $a = 3$ we get the projective plane. In general, Equation 4 gives the number of elements of P. Recall that each point in P is an equivalence class of points in $\boldsymbol{F}^a$. We choose one representative for each point in P. We call these points the *special points*. We consider elements of $\boldsymbol{F}^n$ to be labellings of P by elements of $\boldsymbol{F}$. We define $f : \boldsymbol{F}^n \to \boldsymbol{F}^a$ to be the weighted sum of the special points, according to the labels. The summation is done in $\boldsymbol{F}^a$. The code S is defined to be the kernel of f.

The same argument as for the $(7, 4)$ code shows that S is a perfect 1-error correcting code. For instance, if $s' \in \boldsymbol{F}^n - S$ then $f(S)$ represents some point $[p]$ of P. We choose the label λ so that $f(S) = \lambda p$, where $p \in \boldsymbol{F}^a$ is the special point representing $[p]$. We then let s be the result of subtracting off λ from the label of $[p]$. Then $s \in S$ and $H(s, s') = 1$. We get the isomorphism to $\boldsymbol{F}^k$ by ignoring the labels of points which have only one nonzero coordinate. A counting argument shows that ϕ is an isomorphism.

5 Golay's perfect 3-error correcting code

Probably the most famous of all codes is Golay's $(23, 12)$-code. This is a perfect 3-error correcting BL code. To see that such a beast is possible, note that the Hamming 3-balls in $\boldsymbol{F}^{23}$ have size

$$\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} = 2^{11} = \frac{2^{23}}{2^{12}}. \tag{5}$$

Golay found this amazing identity first and then worked out the corresponding code in an *ad hoc* way. Since then, this code has been studied extensively and related to many things – e.g. the densest possible packing of balls in 24-dimensions. Looking carefully at Golay's code is like staring into the sun.

In this section, I'll give one construction of it, without a (conceptual) proof that it works. We'll revisit this code from a deeper perspective in later sections. It is known that there is only one perfect 3-error correcting $(23, 12)$ code up to isomorphism, so all the many constructions give the same result.

Given any graph Γ, whose vertices are labelled by integers $1, ..., k$, the *reverse incidence matrix* M_Γ is the $k \times k$ matrix which has a 1 in the ijth entry if and only if either $i = j$ or i and j label non-incident vertices. The next page of these notes shows the pertinent example.

We let G be the 24×12 binary matrix having the form $[I][M]$ where I is the 12×12 identity matrix and M is the reverse incidence matrix for the graph made from the edges of the icosahedron. See Figure 1 below. These two matrices are just stuck side-by-side. We define $S \subset \boldsymbol{F}^{24}$ to be the row space of G. Clearly S is a 12-dimensional vector subspace of $\boldsymbol{F}^{24}$. The code S has a total of 2^{12} elements. A brute force enumeration checks that all words have weight congruent to 0 mod 4, and that the minimum weight is 8. Hence S is 3-error correcting. Taking the first 12 coordinates gives an obvious map $S \to \boldsymbol{F}^{12}$. Hence S is a 3-error correcting $(24, 12)$ code.

To obtain the $(23, 12)$ code, we simply delete the last coordinate of S. The new code S' is a 12-dimensional subspace of $\boldsymbol{F}^{23}$ and still has minimum Hamming distance 7. If we place Hamming 3-balls about the points of S', we cover exactly 2^{23} points, by Equation 5. Hence, S' is a perfect 3-error correcting $(23, 12)$ code.

It is worth pointing out that one can go the other way around. Had we started with the $(23, 12)$ code, we could produce the $(24, 12)$ code just by adding a single coordinate so that all the words have even weight.

Here is the matrix M_Γ.

$$M_\Gamma = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Here is the isosahedron graph.

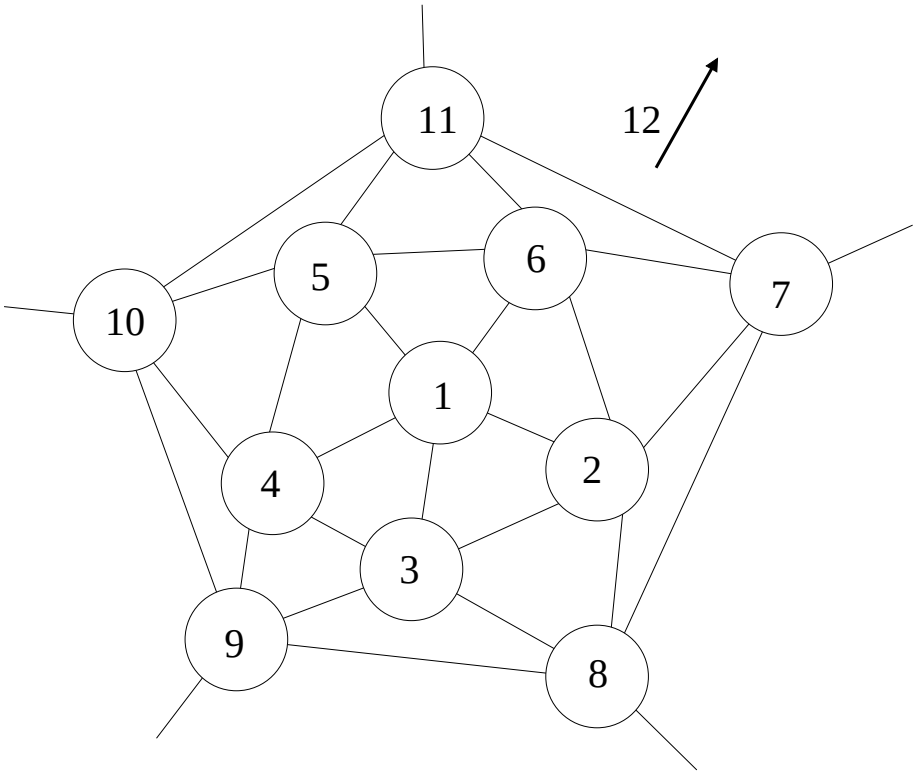

Figure 1.

6 The Hexacode

Our main goal is to understand the Golay $(24, 12)$ code in a more conceptual way. For these purposes, it is useful to define what is known as the *hexacode*. The hexacode is a 3 dimensional subspace of $\boldsymbol{F}_4^6$, where $\boldsymbol{F}_4$ is the field with 4 elements.

$\boldsymbol{F}_4$ is the splitting field for the polynomial $x^3 - 1$. The roots of this polynomial are denoted 1, ω, and $\overline{\omega} = \omega^2$. It is natural to think of ω as a 3rd root of unity. The field $\boldsymbol{F}^4$ is thus $\{0, 1, \omega, \overline{\omega}\}$. It is worth pointing out that $1 + \omega + \overline{\omega} = 0$, just as what happens for the number field $\boldsymbol{Q}(\omega)$. Indeed, reduction mod 2 gives a beautiful surjective ring homomorphism from the Eisenstein integers $\boldsymbol{Z}[\omega]$ to $\boldsymbol{F}_4$.

The hexacode consists of elements $abcdef \in \boldsymbol{F}_4^6$ such that

$$a + b = c + d = e + f = s,$$

$$\begin{aligned} a + c + e = a + d + f = b + c + f = b + d + e = \omega s, \\ b + d + f = b + c + e = a + c + f = a + d + e = \overline{\omega} s \end{aligned} \tag{7}$$

for some $s \in \boldsymbol{F}_4$.

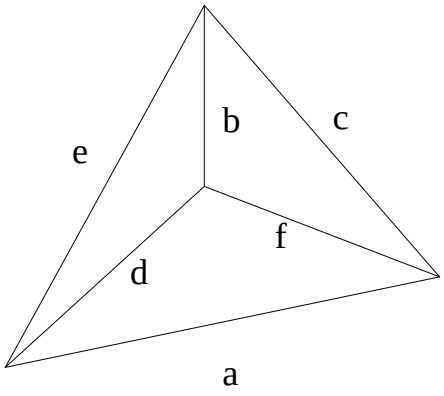

Figure 2.

There is a nice geometric interpretation of these rules. We think of $\boldsymbol{F}_4^6$ as a labelling of the edges of a regular tetrahedron T by elements of $\boldsymbol{F}_4$. Given any vertex v of T, we define the *vertex sum* of v to be the sum of the labels of the edges incident to that vertex. Similarly, given any face of T, we define the *face sum* to be the sum of the edges bounding the face. Finally, given any *axis* of T, namely, a line that goes through the centers of two opposite

edges, we define the *axis sum* to be the sum of the two edges corresponding to the axis. The hexacode consists of labellings with the following property. There exists an element $s \in \boldsymbol{F}_4$ such that

- All axis sums are s.
- All face sums are ωs.
- All vertex sums are $\overline{\omega} s$.

Given a face f of T, we can create a word by labelling the 3 edges of T by ω and the other 3 edges by $\overline{\omega}$. This gives us a word $w(f)$. For instance, if f is the back face in Figure 2, then $a = c = e = \omega$ and $b = d = f = \overline{\omega}$. In this case $s = 1$. One can check easily that $w(f) \in S$ for each face $f = f_1, f_2, f_3, f_4$.

Lemma 6.1 *The code S is the linear span of $W(f_j)$ for $j = 1, 2, 3, 4$.*

Proof: Let S' be the code which is the linear span of $W(f_j)$. We have already seen that $S' \subset S$. The matrix

$$\begin{bmatrix} \overline{\omega} & \omega & \overline{\omega} & \omega & \omega & \overline{\omega} \\ \omega & \overline{\omega} & \overline{\omega} & \omega & \overline{\omega} & \omega \\ \overline{\omega} & \omega & \omega & \overline{\omega} & \overline{\omega} & \omega \end{bmatrix} \tag{8}$$

is obtained by writing down 3 of the words $W(f_j)$. One checks easily that this matrix has rank 3 over $\boldsymbol{F}_4$, meaning that any 3 rows are linearly independent. (Just take determinants.) Hence $\dim(S') \geq 3$. On the other hand, one checks easily that the rules in Equation 7 allow us to recover a word in S if any 3 entries are supplied. Hence $\dim(S) \leq 3$. But $\dim(S') \leq \dim(S')$. All this is possible only if $\dim(S) = \dim(S') = 3$ and $S = S'$. ♠

Being a 3-dimensional vector space of $\boldsymbol{F}_4$, the set S has $64 = 4^3$ elements. One checks easily from the rules in Equation 7 that all nontrivial words in S have even weight at least 4. In particular, there is no word in S with exactly one 0. (This would be a word of weight 5.) By construction, S is a 1-error correcting code of type $(6, 3)$.

7 The Miracle Octad Generator

R.T. Curtis invented the Miracle Octad Generator as a way of efficiently working with the Golay $(24, 12)$-code. The definition here is (I think) due to J. H. Conway. In any case, the explanation is given in the book *Sphere packings* by Conway and Sloane. In a later section, I'll discuss some of Curtis's ideas on the MOG.

The MOG works as follows. Let $M = M_{4,6}(\boldsymbol{Z}/2)$ denote the set of 4×6 binary matrices. (We mean to have 4 rows and 6 columns.) We define a map $\psi : M \to \boldsymbol{F}_4^6$ by the equation

$$\psi(M) = (0, 1, \omega, \overline{\omega}) \cdot M. \tag{9}$$

That is, we simply take the dot product of each column of M with the vector $(0, 1, \omega, \overline{\omega})$. Equation 10 gives an example. The $(-)$ signs are just placeholders.

$$\begin{bmatrix} 0 & - & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & - & 0 & 1 & 1 & 0 & 1 & 0 \\ \omega & - & 1 & 1 & 0 & 0 & 0 & 0 \\ \overline{\omega} & - & 0 & 0 & 1 & 1 & 1 & 1 \\ & - & - & - & - & - & - & - \\ & & \omega & \overline{\omega} & \omega & \overline{\omega} & \omega & \overline{\omega} \end{bmatrix} \tag{10}$$

The element $\psi(M) \in \boldsymbol{F}_4^6$ is called the *score* of M. We call M *balanced* if the parity of the top row of M coincides with the parity of each column of M. So, these 7 parity calculations must be the same. In our example, the parity is odd in all cases, so our example is balanced.

The Golay $(24, 12)$ code has an alternate description as the set S of balanced elements whose scores are in the hexacode. Thus, our example above belongs to S. The rest of this section is devoted to proving that S is a $(24, 12)$ code of minimum Hamming distance 8. In particular S is 3-error correcting.

S is certainly a binary code of length 24. From the rules defining S, it is clear that S is a vector subspace of $M_{4,6}(\boldsymbol{Z}/2) = \boldsymbol{F}^{24}$.

Lemma 7.1 $\dim(S) = 12$.

Proof: We will show that ψ maps S surjectively onto the set of 2^6 hexacode words, and that the kernel of ψ has 2^6 elements. Our example above shows that $\psi(S)$ contains one of the special words $W(f)$ generating the hexacode.

Just by permuting the columns appropriately, we see that $\psi(S)$ contains each of the words $W(f)$. But then $\psi(S)$ contains the span of these words. Hence $\psi(S)$ is exactly the hexacode.

Let K be the kernel of ψ. One can see by inspection that the following elements

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{11}$$

all belong to K. Moreover, so do the complements of these elements. (The complement of an element is obtained by switching each entry by its opposite.) Finally, any matrix obtained by permuting the columns of the above matrices belongs to K. All in all, we count that there are

$$\sum_{i=1}^{6} \binom{6}{i} = 2^6$$

such matrices. Hence K has at least 2^6 entries.

Any column of any element of K must have one of the two forms shown in Equation 11. The parity rules then force such a matrix to be a permutation of one of the matrices shown in Equation 11, or the complement of such a matrix. Hence K has exactly 2^6 elements. ♠

Lemma 7.2 *The weight of any element of S is a multiple of 4.*

Proof: The parity condition guarantees that every element of S has even weight. We just have to show that no element of S has weight congruent to 2 mod 4. If some element M of S has weight congruent to 2 mod 4, then each column of M must have odd weight. If some column of M has weight 3, then at least 2 columns have weight 3, because $3+1+1+1+1+1 = 8$. But then we can replace M by $M + M'$, where M' is some permutation of the middle matrix in Equation 11, so as to reduce the total weight by 4. Hence, it suffices to consider the case when M has total weight 2 or 6. No element of the kernel of ψ as weight 2 or 6, so $\psi(M)$ is a nontrivial hexacode word.

M cannot have total weight 2 because then $\psi(M)$ would have weight less than 4. Suppose that M has weight 6. If M has more than one 1 in the top

row, then $\psi(M)$ has weight at most 3, which is impossible. Hence M has just one 1 in the top row. If some column of M has weight 3 then $\psi(M)$ again has weight less than 4. Hence all columns of M have weight 1. But then $\psi(M)$ has weight 5. This is impossible. ♠

The following lemma completes our proof, because it combines combines with the previous lemma to show that the minimum Hamming distance between elements of S is 8, as claimed.

Lemma 7.3 *No word of S has weight 4.*

Proof: Suppose M is a word of weight 4. No element of the kernel of ψ has weight 4. So, $\psi(M)$ is a nontrivial hexacode word.

If the common parity of M is even, then at most 2 columns of M have nonzero weight. But then $\psi(M)$ has weight at most 2. This is a contradiction. Hence the common parity of M is odd. But then the top row of M has at least 1 and so $\psi(M)$ has weight at most 3. This is a contradiction. ♠

Lemma 7.4 *Aside from the words $0...,0$ and $1,...,$, all Golay codewords have weight 8, 12, or 16.*

Proof: One codeword is $0,...,0$ and another one is $1...1$. Ignoring these two words, the minimum weight of a codeword is 8. The maximum weight is 16 because the Golay code is closed under complementation. That is $s \in S$ if and only if $s + 1...1 \in S$. So, all Golay codewords have weight either 8, 12, or 16. ♠

Lemma 7.5 *The Golay $(24, 12)$ code has 759 codewords of weight 8 and 2576 codewords of weight 12.*

Proof: This can be established by a direct enumeration. ♠

8 The Steiner $(759, 8, 5)$ design

It is worth mentioning an amazing property of the Golay $(24, 12)$ code S. As we mentioned, there are exactly 759 codewords of weight 8. These are called *octads*. Say that a *pentad* is a length 24 vector of weight 5. There are 24 choose 5 pentads – many more than there are octads. Conveniently,

$$759 = \frac{\binom{24}{5}}{\binom{8}{5}} \tag{12}$$

This suggests the following result.

Lemma 8.1 *Any length* 24 *vector of weight* 5 *is contained in a unique octad.*

Proof: The mod 2 sum of two octads is again an element of S. In particular, the sum of two octads is either trivial or has weight at least 8. From this we see that two octads can have at most four 1's in common. Hence, any pentad is contained in at most 1 octad. On the other hand, each octad contains 8 choose 5 pentads, and (one can count that) there are 759 octads. Hence, by Equation 12, every pentad arises as a subset of some octad. ♠

One can give a more constructive proof, in which one starts with a pentad and uses the rules above to complete it uniquely to an octad. I'll leave this to you. There are even java applets on the web which do this for you. The set of 759 octads of the Golay code are known as the $(759, 8, 5)$ Steiner design. This is one of the great objects of combinatorics.

9 The MOG and Projective Space

I mentioned above that the given description of the MOG is really due to Conway. Here I'll explain some of the ideas of R.T. Curtis, the inventor of the MOG. It turns out that the MOG encodes one of the beautiful bijections in combinatorics.

Let $P = P^3(\boldsymbol{Z}/2)$ be the projective space over $\boldsymbol{Z}/2$. There are 35 lines in P. To see this, note that a line is determined by a choice of 2 points. There are 15 points in P, so we can choose 2 points in 15 choose 2 ways. Each line

has 3 points on it, so there is a redundancy of 3 choose 2 in our description of the lines. Hence the total number of lines is

$$35 = \frac{\binom{15}{2}}{\binom{3}{2}}.$$

At the same time, there are

$$35 = \binom{7}{3}$$

ways to pick 3 elements from a 7 element set. Say that a 3 element subset of a 7 element set is a *triad*.

Obviously, there is a bijection between the set of triads and the set of lines in P. What is less obvious is that there is a bijection which preserves some of the combinatorial structure of the two sets. Say that two triads are *incident* if they share exactly 1 member in common. For instance 1110000 and 1101000 are incident. The beautiful fact is that there is a bijection between the set of triads and the set of lines such that two triads are incident if and only if the corresponding lines intersect. Thus, the pattern of intersections of lines in P is encoded by the structure of the triads. The MOG "implements" such a bijection.

We think of a 7-element set as a 2×4 binary matrix having four 1's, one of which is in the upper left corner. Here is one way to encode a triad with such a matrix

$$abcdefg \to \begin{pmatrix} 1 & d \\ b & e \\ c & f \\ d & g \end{pmatrix} \tag{13}$$

Incident triads are mapped to matrices which have exactly two 1's in common.

At the same time, we can represent a line in P as a 4×4 matrix with exactly four 1's, one of which is in the upper left corner. The remaining 3 entries encode the line. Not any matrix is allowed. To describe the allowable matrices, we let Ω be the 4×4 matrix whose entries are certain binary strings, namely

$$\Omega = \begin{bmatrix} 0000 & 0100 & 1000 & 1100 \\ 0001 & 0101 & 1001 & 1101 \\ 0010 & 0110 & 1010 & 1110 \\ 0011 & 0111 & 1011 & 1111 \end{bmatrix} \tag{14}$$

A matrix M is allowable if and only if

$$\sum_{ij} M_{ij}\Omega_{ij} = 0. \tag{15}$$

We are not doing matrix multiplication here. Rather, we are think of both M as vectors of length 16 and taking their dot product.

Ω is really a planar representation of the 16 points in the 4 dimensional vector space $V = (\boldsymbol{Z}/2)^4$ and the nonzero entries of M are selecting an affine subspace of V that contains the origin. Thus, the allowable matrices are in bijection with the lines in P. We can read off the coordinates of the line by looking at the entries of Ω which pair with nonzero entries. For instance, the matrix

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

encodes the line containing the points

$$[0:1:0:1], [1:1:1:0], [1:0:1:1]$$

Our allowability condition just says that the sum of these vectors is 0.

Now, we define the bijection from the set of triads to the set of lines. Let T be one of our 2×4 matrices representing a triad. Then there is a unique allowable matrix M such that the 6×4 allowable matrix TM belongs to the MOG – i.e. is a balanced matrix whose score is in the hexacode.

To see why the map is well defined, suppose we let M_0 denote the matrix having a 1 in the upper left corner and having all other entries 0. Then TM_0 is a pentad. There is a unique way to add 3 more 1's to TM_0 to make an octad, and it turns out that these 1's must all be added to M_0, and that the resulting matrix, M, is allowable. A few examples will convince you that this works.

Suppose T_1 and T_2 are incident. Consider the MOG elements T_1M_1 and T_2M_2. These elements both have weight 8 and can have at most 4 elements in common. On the other hand, their sum also has weight 8. Hence, they have exactly 4 elements in common. This means that M_1 and M_2 have exactly 2 elements in common. But M_1 and M_2 automatically share the upper left 1. Hence, the lines representing M_1 and M_2 intersect in exactly one point. A similar analysis shows that if T_1 and T_2 are not incident, then M_1 and M_2 share only one element in common. Hence the corresponding lines do not intersect.

10 The E_8 lattice

As preparation for the definition of the Leech lattice, we consider a simpler sphere packing in 8 dimensions. The Hamming $(7,4)$ code can be used to create a very dense packing of spheres in 8 dimensions. Let $\boldsymbol{F} = \boldsymbol{Z}/2$, and let $S' \subset \boldsymbol{F}^7$ denote the Hamming $(7,4)$-code. We form a code $S \subset \boldsymbol{F}^8$ by attaching a single digit so that the weight of every word is even. The minimum weight of S is 4. Note that 1111111 belongs to S', so 11111111 belongs to S. Every word in S has weight 4 except 00000000 and 11111111.

We define $\Lambda_8 \subset \boldsymbol{R}^8$ to be the set of vectors V such that V mod 2 belongs to S. That is, coordinatewise V is congruent mod 2 to some element of S. Since S is a linear subspace of $\boldsymbol{F}^8$, the set Λ_8 is a subgroup of $\boldsymbol{R}^8$.

Lemma 10.1 *The minimum distance between points in* Λ_8 *is* 2.

Proof: It suffices to show that the shortest nontrivial vector in Λ_8 has length 2. The vector $(2,0,0,0,0,0,0,0)$ certainly belongs to Λ and has length 2. Any vector congruent to $(0,0,0,0,0,0,0,0)$ mod 2 has at least one 2 and hence has length at least 2. Any vector congruent to a nontrivial element of S has at least 4 odd coordinates. Hence, again, such a vector has length at least $2 = \sqrt{1+1+1+1}$. ♠

The Λ_8 sphere packing is obtained by placing balls of radius 1 about each of the points of Λ_8. Thanks to the previous result, these balls are all disjoint from each other.

Lemma 10.2 *Each ball in the* Λ_8 *sphere packing is tangent to* 240 *other balls.*

Proof: It suffices to prove that there are exactly 240 vectors in Λ_8 which have length 2. The vector $(\pm 2,0,0,0,0,0,0)$ and its permutations provides 16 such vectors. If we add signs arbitrarily to each of the 14 weight 4 words of S we produce another $2^4 \times 16 = 224$ such vectors. Adding together everything, we get 240 vectors of length 2. It is easy to see that any other vector in Λ_8 has longer length. ♠

What worked well here is that the vectors

$$(2,0,0,0,0,0,0,0)$$

and

$$(\pm 1, ..., \pm 1)$$

all have the same length. It is tempting just to define Λ_8 to be the lattice generated by these vectors. However, it then turns out that some of these vectors are too close together. The beauty of the $(8, 4)$ code is that it picks out a nice subset of these vectors, guaranteeing that no two come within 2 of each other.

11 The Leech Lattice

Just as the $(8, 4)$ extended Hamming code is used to produce the E_8 lattice packing, the $(24, 12)$ Golay code is used to construct the Leech lattice. The Leech lattice was (re)discovered by Leech in 1964, but according to wikipedia it was known to Witt in 1940 but not published. (I don't know the history of this.)

What works well in 24 dimensions is that the vectors

$$(2, 2, 0, ..., 0); \qquad (2, 2, 2, 2, 2, 2, 2, 2, 0, ..., 0); \qquad (3, 1, ..., 1)$$

and their permutations and negations all have the same length. The idea behind the Leech lattice is to use the Golay $(24, 12)$ code as a guide for selecting which of these vectors should generate a lattice. This is analogous to what we did for the E_8 lattice.

Let S denote the $(24, 12)$ Golay code. Given any $s \in S$ and any integer m, let $L(s, m) \subset \boldsymbol{R}^{24}$ denote those vectors $v = (v_1, ..., v_{24})$ such that

- $v_1 + ... + v_{24} = 4m$.
- $v_k - m \equiv 2s_k \bmod 4$.

Then

$$\Lambda_{24} = \bigcup_{s \in S} \bigcup_{m \in \boldsymbol{Z}} L(s, m). \tag{16}$$

Lemma 11.1 *Λ_{24} is a lattice.*

Proof: Suppose that $v_1, v_2 \in \Lambda_{24}$. Then there are elements $s_1, s_2 \in S$ and integers $m_1, m_2 \in \boldsymbol{Z}$ such that $v_j \in L(s_j, m_j)$. Evidently, the conditions

imply that $v_1 + v_2 \in L(s_1 + s_2, m_1 + m_2)$. Here $s_1 + s_2$ is taken mod 2. This works because S is a linear code. Similarly $-v_1 \in L(-s_1, -m_1) = L(s_1, m_1)$. ♠

Lemma 11.2 *Every two points in Λ_{24} are at least $4\sqrt{2}$ units apart.*

Proof: It suffices to prove that the shortest nontrivial vectors in Λ_{24} have length $4\sqrt{2}$. Let $v \in L(s, m)$. There are 4 cases to consider.

- Suppose s is trivial and m is even. In this case, each coordinate of v is a multiple of 4, and least two of them are nonzero. Hence $\|v\| \geq \sqrt{16 + 16} = 4\sqrt{2}$.
- Suppose s is trivial and m is odd. In this case, all coordinates of v are even and nonzero. Hence $\|v\| \geq \sqrt{96} > 4\sqrt{2}$.
- Suppose that s is nontrivial and m is even. In this case, all the nonzero coordinates of v are even and at least 8 of them are nontrivial. Hence $\|v\| \geq \sqrt{4 + ... + 4} = 4\sqrt{2}$.
- Suppose that s is nontrivial and m is odd. Then all coordinates of v are odd. The vector $(1, ..., 1)$ does not belong to Λ_{24} because it would correspond to $m = 6$. Hence, at least one coordinate is 3, leading to $\|v\| \geq \sqrt{23 + 9} = 4\sqrt{2}$.

This takes care of all the cases. ♠

The Leech packing is obtained by placing balls of radius $\sqrt{2} = \sqrt{8}/2$ about each point of Λ_{24}.

Lemma 11.3 *Each ball in the Leech packing is tangent to* 196560 *other balls.*

Proof: It suffices to prove that there are 196560 vectors of length $2\sqrt{2}$ in the Leech lattice. We'll divide the count into 5 cases. The first three cases are the same as in the previous result, and the last two cases come from subdividing the fourth case above. Let $v \in L(m, s)$ where s is trivial.

- Suppose m is even and s is trivial. The vectors $(\pm 2, \pm 2, 0, ..., 0)$ and all their permutations are the only possibilities. This leads to

$$2 \times 23 \times 24 = 1104$$

vectors.

- Suppose m is odd and s is trivial. As we saw above, this gives no vectors of length $4\sqrt{2}$.

- Suppose that s is nontrivial and m is even. In this case v must have 8 nontrivial entries, all ± 2, and there must be an even number of $+$ signs. Since there are 759 octads, we get

$$759 \times \sum_{k=0^4} \binom{8}{2k} = 97152$$

possibilities. Here we are choosing the locations of the $(-)$ signs.

- Suppose that s is nontrivial and m is odd and v has a 3 in one of its coordinates. In this case $m \equiv 3 \bmod 4$ and v must have be a permutation of the vector $(3, \pm 1, ..., \pm 1)$, where all the entries 3 and -1 occur in the same spots as an element of S. When s has weight k, this leads to $k \times N(k)$ vectors, where $N(k)$ is the number of words of weight k in the Golay code. The possibilities here are $k = 24, 16, 12, 8$. Hence, there are

$$24 + 759 \times 16 + 2576 \times 12 + 759 \times 8 = 49152$$

possibilities.

- Suppose that s is nontrivial and m is odd and v has a -3 in one of its coordinates. In this case, $m \equiv 1 \bmod 4$ and v must have be a permutation of the vector $(-3, \pm 1, ..., \pm 1)$, where all the -1 entries occur in the same spots as an element of S, and the remaining entries do not. When s has weight k there are $(24 - k) \times N(k)$ possibilities. The possibilities here are $k = 0, 8, 12, 16$. Since $N(k) = N(24 - k)$. we get the same sum as in the previous case, namely

$$49152$$

possibilities.

Adding everything up, we get 196560 vectors of length $4\sqrt{2}$. ♠

The Leech lattice has been shown to be the densist possible lattice packing. Moreover, the configuration of 196560 balls has been shown to be unique: Up to rotations, there is no other way to place 196560 balls around another one so that they are all tangent to it.