Lattices and Codes

H. A. Verrill

Seminar on Finite Simple groups and Number theory

The application I'm going to talk about is from coding theory. The outline of the seminar is as follows:

1. Idea of codes
2. Example of linear codes
3. Codes and Lattices
4. Theta functions
5. Stiener systems
6. The Mathieu group M_{24}

1. Idea of codes

Information is encoded and sent (as in Figure 1), hopefully in such a way that if you make errors, they can be detected and corrected if they are not too bad.

In this diagram, we get words in a certain language, written in a certain alphabet; we might as well assume all the words are the same length, n. In the picture, we've taken an alphabet with 5 different letters. Not all words are allowed; if we get a word that is not allowed, hopefully we can correct it. We do this all the time in English, eg, you see a word like "elephint" and probably correct it to "elephant". But if you make too many errors, you might not be able to work out what was meant.

Anyway, for now we assume there are only 2 letters in our alphabet, 0 and 1. So in fact we can say that our alphabet is $\mathbb{F}_2$, and our words are elements of $\mathbb{F}_2^n$.

The collection of "allowed" words is called the collection of codes, and denoted C. The distance between two elements of C is the number of places they differ in. This is shown in Figure 2, where codes are length 2, and lines are drawn joining adjacent code words (i.e., code words distance one apart). If we choose the circled words to be our code words, then at least we will be able to recognise up to one error, but we won't be able to make a correction.

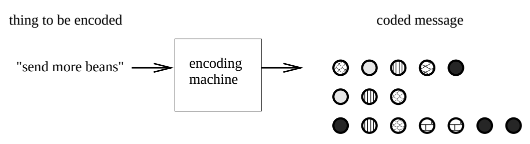


FIGURE 1. Sending a message

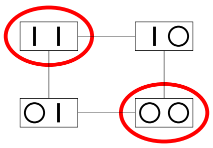


FIGURE 2. A code in $\mathbb{F}_2^2$

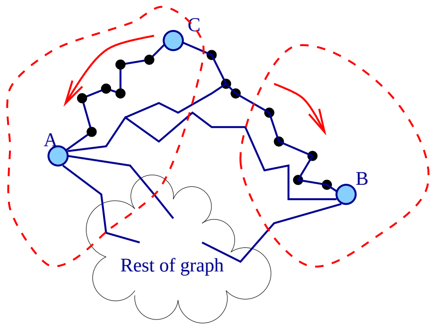


FIGURE 3. Correction of errors

We let d denote the distance between words. A code can correct $\lceil (d-2)/n \rceil$ errors, as indicated in Figure 3

If the codes are in $\mathbb{F}_2^n$, and the minimum distance is d, and there are a total of M code words, we call the code a $(n, M, d)-$code. What we want to do is get d big, and also M big, so we can send a lot without too much effort, and also have a good chance of correcting lots of errors.

For these binary codes, this is expressed as the idea of trying to maximize the information rate

$$R := \frac{log_2(M)}{n}$$

for fixed distance d.

List of nice things for codes to satifsy:

- Big d
- Big R
- Symmetry - so the automorphism group of the graph is transitive, eg, linear codes, which are linear subspaces of $\mathbb{F}_2^n$
- "covering".

We can think of these binary codes as vertices of an $n-$dimensional cube. Figure 4 shows some possible codes in $\mathbb{F}_2^4$ - which shows this is not such a good place to try and find codes.

Another way of looking at this situation is shown in figure 5

Two words is not much good for the code with distance 3; we need to make the code words longer; the Hamming code is what happens when you take the length to be 7. Before we get to the Hamming code, we can find a sequence of codes, based on the (4,2,3)-code. This is shown in Figure 6

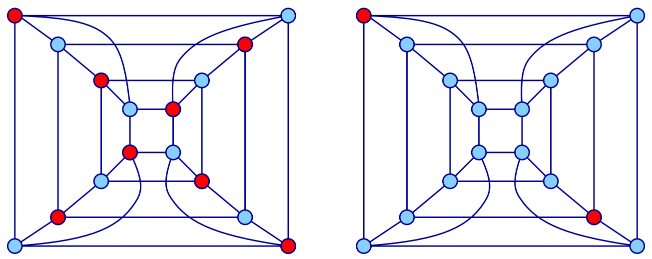


FIGURE 4. Codes in $\mathbb{F}_2^4$

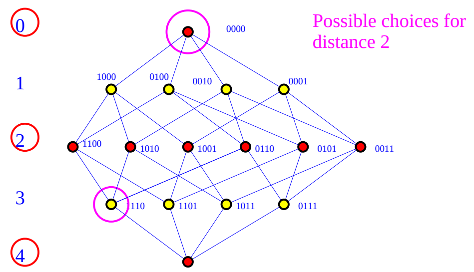


FIGURE 5. Graph for $\mathbb{F}_2^4$

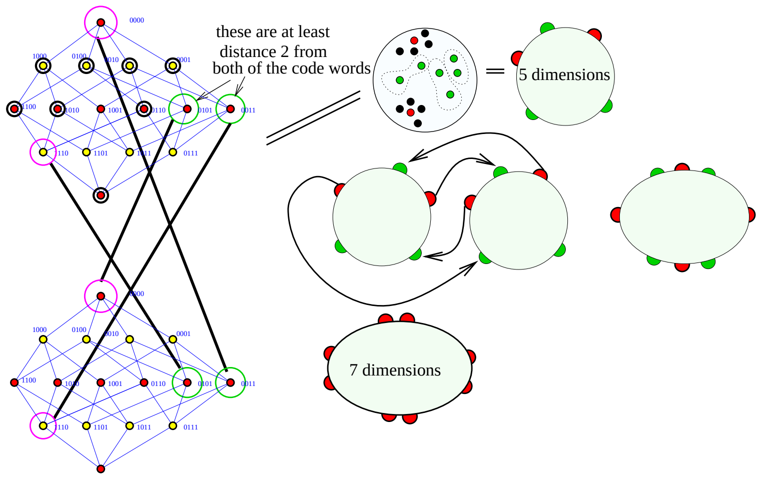


FIGURE 6. Towards the Hamming code

After we get to dimension 7, this process can't be continued. In this sequence, the information rate increase like this:

$$log_2(2)/3 = 0.3333, log_2(2)/4 = .25, log_2(4)/5 = .4, log_2(8)/6 = .5, log_2(16)/7 = .571$$

Anyway, proceeding in the above manner does give a fairly reasonable code, but obviously this is not a very intellegent way to construct it, since we can't really see how to get better codes.

2. Linear codes

Here's another way to tell you what the Hamming Code is: the vectors:

$$\begin{gathered}(0,0,0,0,0,0,0)\\(0,0,0,1,1,1,0)\\(0,0,1,0,1,0,1)\\(0,0,1,1,0,1,1)\\(0,1,0,0,0,1,1)\\(0,1,1,0,1,1,0)\\(0,1,1,1,0,0,0)\\(1,0,0,0,1,1,1)\\(1,0,0,1,0,0,1)\\(1,0,1,0,0,1,0)\\(1,1,0,0,1,0,0)\\(1,1,0,1,0,1,0)\\(1,1,1,0,0,0,1)\\(1,1,1,1,1,1,1)\end{gathered}$$

This is a linear subspace of $\mathbb{F}_2^7$. That means we can do linear algebra type stuff, eg, it must be given by the kernal of some matrix B. Eg, take

$$B := \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

Let A be the matrix given by taking it's columns to be spanning vectors of the code, eg, we can take

$$A := \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

(Notice weird thing about relation between A and B.)
then there an exact sequence:

$$0 \to \mathbb{F}_2^4 \to \mathbb{F}_2^7 \to \mathbb{F}_2^3 \to 0$$

The columns of A given the kernal of B, so the columns of B^t give the kernal of A^t, so we get another exact sequence and another code, the dual code, which is spanned by the rows of B. If the dual is the same as what you started from, the code is called **self dual**.

If a code is a linear code in $\mathbb{F}_2^n$, we say it is a $[n,k,d]-$code if the minimum weight is d, and the dimension of the code is k.

2.1. Self dual codes. There will be some good results about self dual codes, which means it's nicer to deal with them. We can go from the Hamming code H to the extended code, $\tilde{H}$, which is a self dual code, by adding a parity check:

$$(x_1, x_2, ...x_7) \mapsto (x_1, x_2, ..., x_7, x_1 + x_2 + x_3 + \cdots + x_7)$$

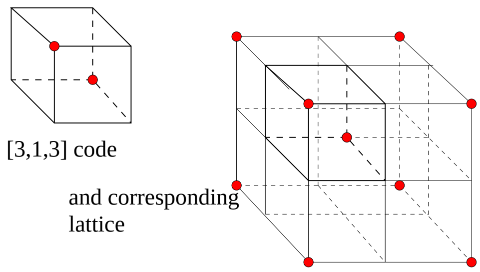


FIGURE 7. Part of lattice from a very simple code

Amazingly, it's possible to extend it and make the distance bigger at the same time, so $d = 4$; though on the other hand this has information rate $4/8 = .5$, so we've reduced our rate again.

We can get back to the Hamming code by deleting any column of the extended code.

The weight enumerator

The weight enumerator is a function that bundles up all the weights of the code into a polynomial, given by

$$\begin{aligned} W_c(x,Y) &:= \sum_{u\in C} X^{n-w(u)} Y w(u) \\ &= \sum_{u \in C} A_i X^{n-i} Y i \end{aligned}$$

where A_i is the number of codes of weight i

If we let $X = Y = 1$, this just counts all the code words, and we get 2^k.

For the Hamming code, we've got

$$W_H(X,Y) = X^7 + 7X^4Y^3 + 7X^3Y^4 + Y^7$$

for the extended Hamming code, we get:

$$W_{\tilde{H}}(X,Y) = X^8 + 14X^4Y^4 + Y^8$$

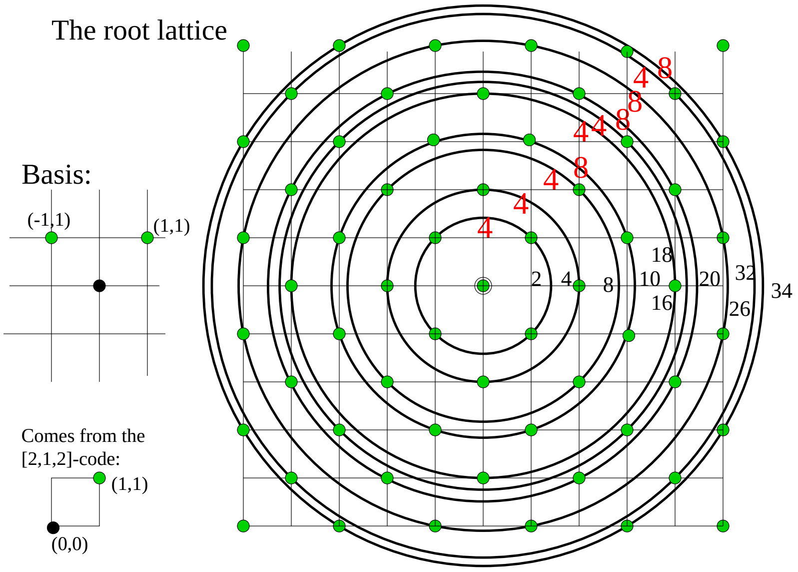
3. How to get a lattice from a code

Well, this is easy; the code, if it's a linear code, is already a lattice; but it's in $\mathbb{F}_2^n$, we'd like a lattice in $\mathbb{Z}^n$; but all you do is forget that you quotiented out by 2, and you're in $\mathbb{Z}^n$ all along. Figure 7 shows an example. The only extra thing to do is divide by $\sqrt{2}$.

Obviously we're going to get a more complicated lattice from the Hamming code. In fact what we get is the cellebrated E_8 lattice (from the extended code).

Clearly the properties we're looking for in a code - that the points are all a reasonable distance apart, but the dnesity is still as high as possible, etc, translate into properties of lattices, eg, there is a relation to the sphere packing problem.

Lattice	Code
Integral	$C \subset C^{\perp}$
Even	doubly even
Unimodular	self dual



$$\vartheta_\Gamma(q) = 1 + 4q^2 + 4q^4 + 4q^8 + 8q^{10} + 4q^{16} + 4q^{18} + 8q^{20} + 8q^{26} + 4q^{32} + 8q^{34} + \ldots$$

FIGURE 8. Example of a theta function

3.1. How to get from a lattice to a code? We can't get a code from every lattice; there is a result for irreducible root lattices:

THEOREM 3.1. *For an irreducible root lattice, the following are equivallent:*
(i) $\Gamma = \Gamma_C$ *for a binary code* $C \subset \mathbb{F}_2^n$
(ii) $A_1^n \subset \Gamma$
(iii) $2\Gamma^* \subset \Gamma$
(vi)Γ *is of type* A_1, D_n *(*$n > 4$*, even)* E_7 *or* E_8.

4. Theta functions

Theta functions are very useful functions that come up all over the place; one important use is the theta function ϑ_Γ of a lattice Γ. This is defined as follows:

$$\vartheta_\Gamma(\tau) = \sum_{x \in \Gamma} q^{\frac{1}{2}x.x}$$

Where $q = exp(2\pi i \tau)$, $\tau \in \mathbb{H}$, the upper half plane. By the way, this is well defined.

This function is analogous to the weight enumeration polynomial, so this means we can apply results about theta functions to get results about codes.

First note that for an even integral lattice,

$$\vartheta_\Gamma(\tau) = \sum_{x \in \Gamma} a_r q^r$$

where a_r is the number of vectors of length $2r$, so we're counting lattice points on a sphere radius $\sqrt{2r}$. So this tells us about the growth of the a_r, and that the sequence converges. See Figure 8. (Unfortunately this is not a unimodular lattice).

THEOREM 4.1. *If Γ is an even unimodular lattice in $\mathbb{R}^n$, then $n \equiv 0$ mod(8), and the theta function $\vartheta_\Gamma(\tau)$ is a modular form of weight $\frac{n}{2}$.*

The Weight ennumerator tells us about the minimum distance of vectors in the code.

THEOREM 4.2. *if our code is in $\mathbb{F}^{24}$, then $A_i = 0$ if $i \not\equiv 0 (mod\ 4)$*

THEOREM 4.3. *Also $A_8 = 759 - 4A_4$.*

THEOREM 4.4. *(Mac Williams Identity) For a binary $[n,k,d]-$code $C \subset \mathbb{F}_2^n$,*

$$W_{C^\perp}(X,Y) = \frac{1}{2^k} W_C(X+Y, X-Y).$$

COROLLARY 4.5. *For a self dual code,*

$$W_C(X,Y) = W_C(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}).$$

THEOREM 4.6. *(Gleason) If $C \subset \mathbb{F}_2^n$ is a doubly even self dual code, then $W_C(X,Y)$ is a polynomial in*

$$\phi := W_{\tilde{H}}(X,Y) = X^8 + 14X^4Y^4 + Y^8$$

and

$$\xi := X^4Y^4(X^4+Y^4)^4$$

This result comes from modular forms. Now, if we want to get a big minimum distance, we would like to make $A_4 = 0$.

4.1. sketch proofs.

THEOREM 4.7. *The algebra of modular forms is isomorphic to the polynomial algebra*

$$\mathbb{C}[E_4, E_6]$$

Here E_4 is a modular form of weight 4 given by:

$$E_4(\tau) := \frac{1}{2\zeta(k)} \sum_{(m,n)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m\tau+n)^4}$$

and E_6 is a modular form of weight 6 given by:

$$E_6(\tau) := \frac{1}{2\zeta(k)} \sum_{(m,n)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{1}{(m\tau+n)^6}$$

Another important modular form is the weight 12 modular form Δ, which is given by

$$\Delta(\tau) = 1728q \prod_{r=0}^{\infty} (1-q^r)^{24}.$$

Now we get going on the theorems about codes....

THEOREM 4.8. *Theorem 4.3*

Proof. Assume that n is not divisible by 8. wlog $n \equiv 4 \bmod 8$.
The idea is to show that

$$\vartheta(TS\tau) = -\tau^{n/2}\vartheta(\tau),$$

which gives

$$\vartheta((TS)^3\tau) = -\vartheta(\tau),$$

but $(TS)^3 = 1$, which gives a contradiction. □

Theorem 4.9. *ϑ is a modular form of weight 12; the space of modular forms of weight 12 is spanned by E_4^3 and E_6^2.*

Corollary 4.10. *for a weight 12 modular form $\sum a_i q^i$, we have*

$$a_2 = -24a_1 + 196560a_0$$

Since the theta function for a unimodular lattice is a modular form of weight $n/2$, which is 12 in our case, and the coefficient of q^0 is 1, we have

$$a_2 = 196560 - 24a_1$$

for the theta function.

Lemma 4.11. *For a self dual doubly even code,*

$$a_1 = 24.2 + 16A_4$$

and

$$a_2 = 2^8 + 16A_4.20.2 + \begin{pmatrix} 24 \\ 2 \end{pmatrix}.4$$

Proof.

$$a_1 = \text{number of elements } x \in \Gamma_C \text{with} x.x = 2$$
$$a_2 = \text{number of elements } x \in \Gamma_C \text{with} x.x = 4$$

Figure 9 gives a first idea of how the code gets repeated throughout $\mathbb{Z}^{24}/\sqrt{2}$ to give a lattice. Figure 10 shows how we get the above numbers □

From this we deduce Theorem 4.3
Next for Theorem 4.6

Theorem 4.12. *The algebra of modular forms of weight divisible by 4 is generated by E_4 and Δ.*

The idea of the proof of Theorem 4.6 is to take some functions $A(\tau)$ and $B(\tau)$, and look at $W_C(A(\tau), B(\tau))$, and show that this function must be a modular form; but A and B have been chosen so that

$$E_4 = A^8 + 14A^4B^4 + B^8$$

and

$$\Delta = 172816A^4B^4(A^4 - B^4)^4.$$

From this we have the result.
A and B are given by

$$A(\tau) := \sum_{x \in \mathbb{Z}} q^{(x.x)}$$
$$B(\tau) := \sum_{x \in 2\mathbb{Z}+1} q^{\frac{1}{4}(x.x)}$$

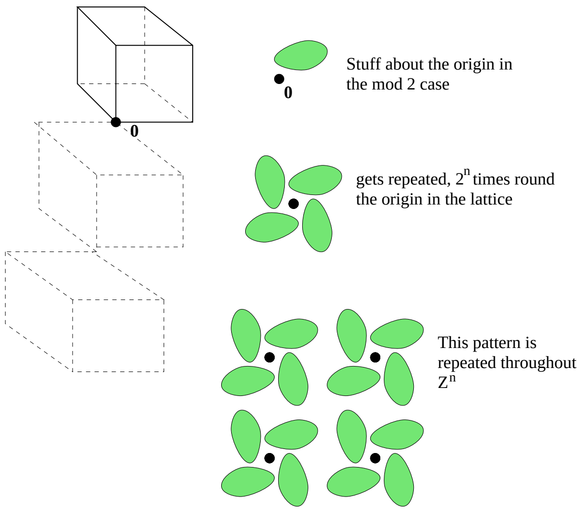


FIGURE 9. Idea of code repeating in the lattice

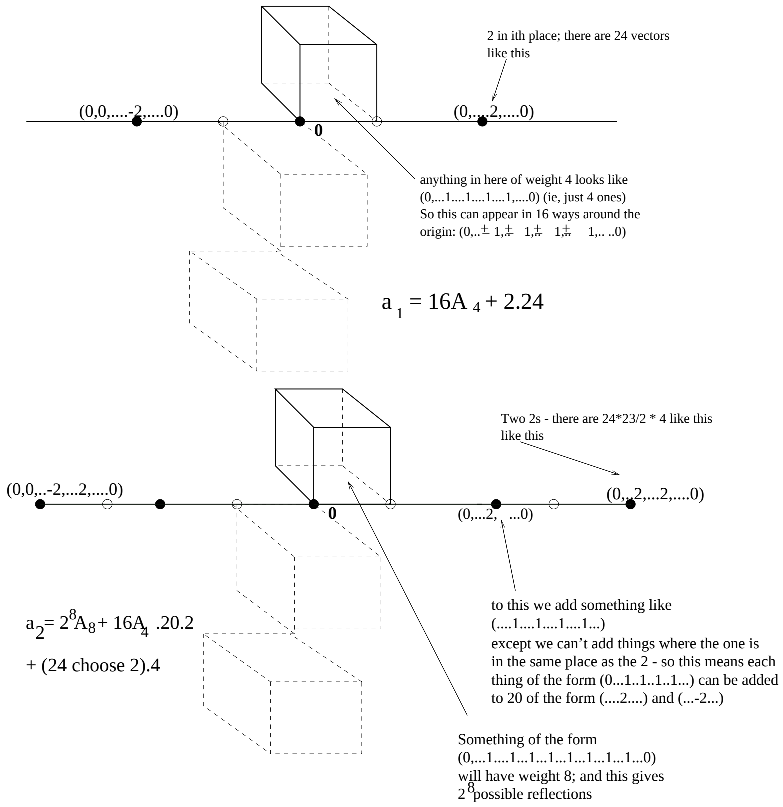


FIGURE 10. Where relations between a_i and A_i come from

I won't go through the proof that $W_C(A,B)$ is a modular form; I will go over a useful lemma that's important in the proof, and gives you an idea of how things might go:

LEMMA 4.13. *We have*

$$\vartheta_{\Gamma_C} = W_C(A,B)$$

PROOF. We have a map

$$\rho : \mathbb{Z}^n \to \mathbb{F}_2^n$$

remember that the lattice Γ_C is defined to be

$$\Gamma_C := \frac{1}{\sqrt{2}}\rho^{-1}(C)$$

and

$$\vartheta_{\Gamma_C}(\tau) = \sum_{x\in\Gamma_C} q^{\frac{1}{2}x.x}$$

$$= \sum_{c\in C} \sum_{x\in\frac{1}{\sqrt{2}}\rho^{-1}(c)} q^{\frac{1}{2}x.x}$$

Now we also have

$$\sum_{x\in\frac{1}{\sqrt{2}}\rho^{-1}(c)} q^{\frac{1}{2}x.x}$$

$$= \sum_{x\in\rho^{-1}(c)} q^{\frac{1}{4}x.x}$$

$$= A^{n-w(c)}B^{w(c)}$$

x □

Now, if we want to get a big minimum distance, we would like to make $A_4 = 0$.

So, now we might set about looking for such a code... It turns out there is a unique doubly even linear code with $A_4 = 0$. This is called the Golay code. It's automorphism group is the finite simple group M_{24}. Wow! This group is a Mattieu group.

5. Stiener Systems

After the above results about coding theory and number theory, we will go on to look at the automorphism groups of some of the very nicest codes; these are the automorphism groups of certain Stiener systems, and turn out to be examples of finite simple sporadic groups.

A Steiner system $S(l,m,n)$ is a collection of subets of size m in a set of size n such that any subset of size l is in exactly one of them.

The vectors of weight 4 in the extended Hamming code form a Steiner $S(2,3,7)$ system. See Figure 11.

THEOREM 5.1. *For a self dual doubly even code $\tilde{G}$ in $\mathbb{F}_2^{24}$, with $A_4 = 0$, the code words of weight* 8 *form a Steiner* $S(5,8,24)$ *system.*

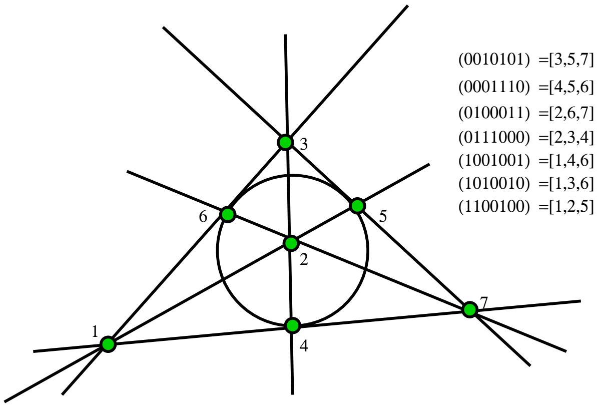


FIGURE 11. Steiner $S(2,3,7)$ system

PROOF. By the above theorem about th A_i, we know there are 759 vectors of weight 8; these can be veiwed as subsets of size 8 from a set of size 24.

Take a subset of size 5, eg, represented by

$$(1,1,1,1,1,\overbrace{0,0,\cdots,0,0}^{19zeros}).$$

Suppose there are two codewords of weight 8 containing this subsequence:

$$c_1 = (1,1,1,1,1,\ \overbrace{0,0,\cdots,0,0}^{16zerosand3ones}\ ).$$

$$c_2 = (1,1,1,1,1,\ \overbrace{0,0,\cdots,0,0}^{16zerosand3ones}\ ).$$

Add these together to get

$$(0,0,0,0,0,*)$$

Here $*$ either could be something like

$$\begin{array}{ccccccc} & 0 & 0 & 0 & 1 & 1 & 1 \\ + & 1 & 1 & 1 & 0 & 0 & 0 \\ \hline & 1 & 1 & 1 & 1 & 1 & 1 \end{array}$$

$$\begin{array}{cccccc} & 0 & 0 & 1 & 1 & 1 \\ + & 1 & 1 & 1 & 0 & 0 \\ \hline & 1 & 1 & 0 & 1 & 1 \end{array}$$

$$\begin{array}{ccccc} & 0 & 1 & 1 & 1 \\ + & 1 & 1 & 1 & 0 \\ \hline & 1 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccc} & 1 & 1 & 1 \\ + & 1 & 1 & 1 \\ \hline & 0 & 0 & 0 \end{array}$$

but none of these is weight divisible by 4, which is a constradiction.

So every subset of size 5 is in at most one of the code words.

But there are $\begin{pmatrix} 8 \\ 5 \end{pmatrix} = 56$ subsets of size 5 in a code of weight 8. This means that altogether $759 * 56$ different subsets of size 5 are contained in these weight 8 codes; but there are only $\begin{pmatrix} 24 \\ 5 \end{pmatrix} = 56 * 759 = 8 * 7 * 11 * 3 * 23$ subsets of size 5 altogether, so this is all of them. □

The Mathieu group M_{24} is the automorphism group of $S(5, 8, 24)$. It has order $244, 823, 040$.

I'll just prove one result about this group:

PROPOSITION 5.2. *The automorphism group of* $S(5, 8, 24)$ *acts transitively on the* 24 *co-ordinate positions.*

6. Sugar and Spice and everything nice

Finite simple group: a group G with no normal subgroup
Normal subgroup: $H \leq G$, s.t $\forall g \in G$, $g^{-1}Hg \leq H$.
Lattice, Γ: $\Gamma = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2 \oplus \cdots \oplus \mathbb{Z}e_n \subset \mathbb{R}^n$, $\{e_i\}$ a basis of $\mathbb{R}^n$.
Integral lattice, Γ: $\forall x, y \in \Gamma$, $x.y \in \mathbb{Z}$.
Dual lattice, Γ: $\Gamma^* = \{x \in \mathbb{R}^n | x.y \in \mathbb{Z} \forall y \in \Gamma\}$
Unimodular lattice Γ: $\Gamma = \Gamma^*$.
even lattice Γ: $\forall x \in \Gamma, 2 | x.x$.
Root vector $x \in \Gamma$: $x.x = 2$.
Root Lattice Γ: Γ is generated by it's roots.
Code of lenght n: $C \subset \mathbb{F}_q^n$.
Binary code of length n: $C \subset \mathbb{F}_2^n$. (We just cover this case today.)
distance between code words: number of places they differ.
Weight of a code word: Number of non zero entries.
Linear code: Linear subspace of ${\mathbb{F}_q}^n$
(n,M,d)-code: $C \subset \mathbb{F}_2^n$, $|C| = M$, min distance between code words is d.
[n,k,d]-code: $C \cong \mathbb{F}_2^k \subset \mathbb{F}_2^n$, minimum distance between code words is d.
Information rate: $R := \frac{log_2(M)}{n}$
modular form of weight k: A function $f(\tau)$, $\tau \in \mathbb{H}$, the upper half plane, with $f(\gamma\tau) = (c(\gamma, \tau))^k f(\tau)$, and $f(\tau)$ has an expansion in terms of $q = exp(2\pi i\tau)$. Here $c(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tau) = (c\tau + d)$.

Theta function of a lattice Γ, $\vartheta_\Gamma(\tau)$: $\vartheta_\Gamma(\tau) = \sum_{x \in \Gamma} q^{\frac{1}{2}x.x}, q = e^{2\pi i\tau}$
Steiner system, S(n,m,l): $(X, \mathcal{B})$, X a set, $|X| = n$, $\mathcal{B} \subset \mathcal{P}(X)$, $\forall U \in \mathcal{B}$, $|U| = m$; $\forall O \subset X, |O| = l \Rightarrow \exists! U \in \mathcal{B} \ni O \subset U$.