London Mathematical Society Impact150 Stories 1 (2016) 51–58
doi:10.1112/i150lms/t.0003

Error-correction and the binary Golay code

R.T.Curtis

Abstract

Linear algebra and, in particular, the theory of vector spaces over finite fields has provided a rich source of error-correcting codes which enable the receiver of a corrupted message to deduce what was originally sent. Although more efficient codes have been devised in recent years, the 12-dimensional binary Golay code remains one of the most mathematically fascinating such codes: not only has it been used to send messages back to earth from the Voyager space probes, but it is also involved in many of the most extraordinary and beautiful algebraic structures. This article describes how the code can be obtained in two elementary ways, and explains how it can be used to correct errors which arise during transmission

1. *Introduction*

Whenever a message is sent - along a telephone wire, over the internet, or simply by shouting across a crowded room - there is always the possibly that the message will be damaged during transmission. This may be caused by electronic noise or static, by electromagnetic radiation, or simply by the hubbub of people in the room chattering to one another. An error-correcting code is a means by which the original message can be recovered by the recipient even though it has been corrupted en route.

Most commonly a message is sent as a sequence of 0s and 1s. Usually when a 0 is sent a 0 is received, and when a 1 is sent a 1 is received; however occasionally, and hopefully rarely, a 0 is sent but a 1 is received, or vice versa. The first objective of an error-correcting code is to alert the recipient to the fact that something isn't quite right, that is to *detect an error*. The second and more demanding objective is to *correct* the message received to what was originally sent. For instance, a simple approach which we all use when we shout across a crowded room is to repeat the message. If the recipient hears 'Jane' and then when I repeat it hears 'Jean' he knows that something is wrong - he has detected an error - but he doesn't know which is correct. Mathematically, if we send a sequence of three 0s and 1s, and then repeat it, the message may arrive as 110100. The recipient knows that there is an error, since $110 \neq 100$, but he doesn't know which is correct.We could, of course, simply repeat the message until it becomes clear what was originally sent, but this would be cumbersome and inefficient. In practice the aim is reliably to send as much information as possible given the length of the messages. For comprehensive accounts of error-correcting codes see, for instance, MacWilliams and Sloane [**5**] or Assmus and Key [**1**]. Conway and Sloane [**2**] gives a wealth of information about the connections between codes, sphere-packing in n-dimensions and lattices.

Received 06 May 2015.

2010 *Mathematics Subject Classification* 11T71 (primary), 20D08 (secondary).

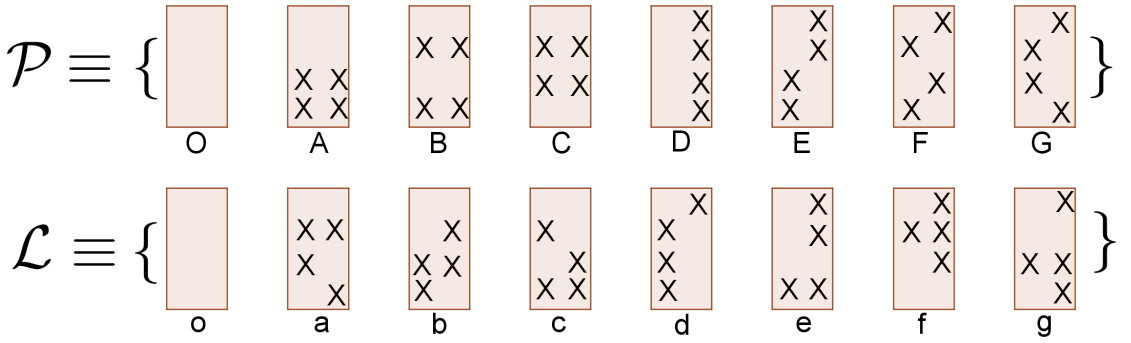


FIGURE 1. *The Point Space and the Line space*

2. *Binary linear codes*

In binary linear codes the symbols 0 and 1 are taken to be the elements of $\mathbb{Z}_2$, the integers modulo 2, thus $1+1=0$, and the code is a subspace C of an n-dimensional vector space V over $\mathbb{Z}_2$ with regard to a fixed basis. Thus its *codewords* are n-tuples of 0s and 1s; the positions in which a vector has non-zero entries is its *support*. If the subspace C has dimension k then the code is said to be an $[n,k]$ code. The *weight* of a codeword is the number of 1s it contains, and the *minimal weight* of non-zero codewords in C is of great importance in coding theory; it is normally denoted by d. Thus an $[n,k,d]$ code has length n, dimension k and minimal weight d. If the *Hamming distance* between two codewords is defined to be the number of positions in which they differ then, by the linearity, the minimal Hamming distance between codewords will also be d. For this reason we have the important result

Theorem 1 *A linear binary $[n,k,d]$ code can correct $\lceil \frac{1}{2}(d-1) \rceil$ errors. If d is even it can correct $\frac{1}{2}(d-2)$ errors and detect $d/2$.* †

To illustrate this result, consider the top row of Figure 1 the 4×2 arrays of which represent 8-dimensional vectors over $\mathbb{Z}_2$ in which the Xs indicate the positions in which 1s occur, there being 0s elsewhere. This set of vectors is closed under addition where, for instance, $A+B=C$ or more symmetrically $A+B+C=O$. They thus form a 3-dimensional subspace $\mathcal{P}$, the *point space*, of the vector space $V \cong \mathbb{Z}_2^8$ whose 2-dimensional subspaces are given by the seven *lines* of the *Fano plane*, see Figure 2. Extending this space by adjoining the all 1s vector I we obtain a 4-dimensional subspace $\mathcal{P}'$ consisting of the all 0s vector, the all 1s vector and 14 vectors (codewords) of weight 4 (namely $A-G$ above and $A'-G'$, where $X'=X+I$ denotes the complement of X). The resulting [8,4,4]-code is an *extended Hamming code*, and the 14 supports of weight 4 codewords form a *Steiner system* S(3,4,8), which is to say that any 3 of the 8 coordinate positions lie together in precisely one of them. According to Theorem 1 this code can correct any $\frac{d-2}{2}=1$ error and detect any 2 errors. This is readily confirmed: if a weight 3 vector is received, then it extends to a weight 4 codeword by converting a unique 0 to a 1; if a weight 5 vector is received, then just one of these 5 positions completes the complementary

† the symbol $\lceil x \rceil$ denotes the greatest integer less than or equal to x.

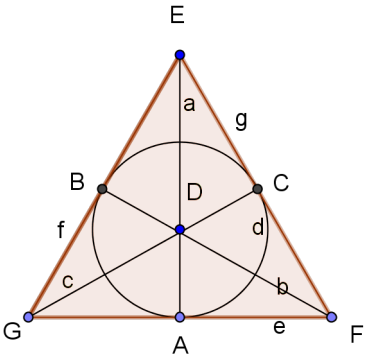


FIGURE 2. *The Fano plane showing the Points and Lines*

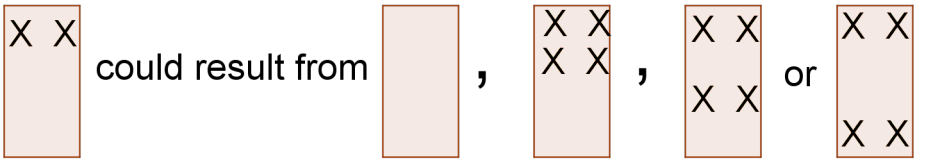


FIGURE 3. *Given 2 errors then the original message could be any one of four possibilities.*

three positions to a codeword and so replacing this 1 by a 0 gives a codeword. If a weight 1 or weight 7 vector is received then, assuming just one error, we may deduce that respectively the all 0s or all 1s vector was sent. However, if two errors have been introduced during transmission then we can deduce that something is wrong, but we cannot deduce the original message. For suppose we receive 1s in the top two positions of the array and 0s elsewhere. Then we know that there have been (at least) 2 errors during transmission, but we cannot say what those errors were, see Figure 3.

3. *The binary Golay code $\mathcal{C}$*

The remarkable binary Golay code $\mathcal{C}$ was discovered by Marcel Golay, a Swiss-born mathematician who spent most of his life in the United States. It is a $[24, 12, 8]$-code [†] consisting of the zero vector, 759 codewords of weight 8, 2576 codewords of length 12, 759 codewords of length 16 and the all 1s codeword. Of course $1 + 759 + 2576 + 759 + 1 = 2^{12}$. According to Theorem 1 it can detect any $\frac{d}{2} = 4$ errors, and can correct any $\frac{d-2}{2} = 3$ errors. For suppose a

[†]technically this is the *extended* Golay code, but we prefer it here as it possesses the greater mathematical symmetry

codeword u is sent, but that three of its entries are altered during transmission, so that $u+e$ is received, where the error e has weight 3. Then u remains the only codeword within Hamming distance 3 of $u+e$. To see this, suppose that $u+e=u'+e'$ where u and u' are codewords and both e and e' are vectors of weight at most 3; then $e+e'=u+u'\in\mathcal{C}$, a codeword of length at most 6. Since the minimum weight of non-zero codewords is 8, we must have $e=e'$ and so $u=u'$. Thus, provided no more than 3 errors were introduced, we may infer that the original message was u. If four errors were introduced then we can see that what we have received is not a codeword, but there are several (6 in fact) codewords within Hamming distance 4 of the message received.

3.1. *The geometry of the Point and Line spaces*

We may now make use of the Point space $\mathcal{P}$ and the Line space $\mathcal{L}$ to construct the binary Golay code. First note that the usual inner product $x.y=\sum x_iy_i$ means that two vectors of $\mathbb{Z}_2^8$ are orthogonal if, and only if, their supports intersect evenly. In particular, vectors of even weight have $x.x=0$; they are *isotropic*. Any two vectors of the extended point space $\mathcal{P}'$ have inner product 0 and so $\mathcal{P}'\subset\mathcal{P}'^{\perp}$. Since $\dim\mathcal{P}'=4$, we have $\mathcal{P}'^{\perp}=\mathcal{P}'$. Any vector v of weight 4 which is not a codeword will be orthogonal to a 2-dimensional subspace of $\mathcal{P}$, (and so orthogonal to a 3-dimensional subspace of $\mathcal{P}'$). The second row of Figure 1 displays a 3-dimensional subspace, the *line space* $\mathcal{L}$ which intersects $\mathcal{P}'$ in the zero vector. Its non-zero vectors all have weight 4; these must correspond one-to-one to the 2-dimensional subspaces of $\mathcal{P}$ since if v and w were orthogonal to the same 2-dimensional subspace then $v+w$ would be orthogonal to the whole of $\mathcal{P}'$ and thus lie in $\mathcal{P}'$, which it visibly does not. Figure 2 shows the correspondence between the vectors of $\mathcal{L}$ and the lines of $\mathcal{P}$. Note that $\dim(\mathcal{P}'+\mathcal{L})=4+3=7$ and that, since all supports of generating vectors are of even weight, every vector in V of even weight can be written uniquely as the sum of a vector in $\mathcal{P}'$ and a vector in $\mathcal{L}$. For $P\in\mathcal{P}$ and $l\in\mathcal{L}$ we write $P\in l$ if P is on the line l, which is to say the vectors corresponding to P and l are orthogonal. We see that if $P\in l$ then $|P+l|=4$ and if $P\notin l$ then $|P+l|=2$ or 6.

3.2. *Construction of $\mathcal{C}$*

In order to construct the binary Golay code $\mathcal{C}$ we take three copies of V (V_1,V_2,V_3 say), placed side by side, so that $V_1+V_2+V_3$ has dimension 24; the V_i are known as the three *bricks*. We now define a 12-dimensional subspace of this space:

$$\mathcal{C}=\{[P,Q,R]_l=[P+l,Q+l,R+l]\mid P,Q,R\in\mathcal{P}'\text{ with }P+Q+R=O\text{ or }I,\ l\in\mathcal{L}\}\,.$$

First note that $\mathcal{C}$ is closed under addition and is thus a subspace. To work out its dimension we see that, from the previous paragraph, we can place any of the 2^7 even subsets in V_1, but that the line l is then fixed; Q can now be chosen freely from $\mathcal{P}'$ with 2^4 choices, but now there are just 2 choices for R depending on whether $P+Q+R=O$ or I. Thus $|\mathcal{C}|=2^7.2^4.2=2^{12}$. Also recall from the previous paragraph that $|P+l|=2$ for $P\in\mathcal{P}',l\in\mathcal{L}$ can only occur if $P\notin l$. But any two distinct lines of the Fano plane intersect in a point (any two distinct 2-dimensional subspaces of a 3-dimensional space intersect in a 1-dimensional space), so it is not possible for a codeword of $\mathcal{C}$ to have weight 2 (or less) in each of V_1,V_2 and V_3 unless it is the zero vector. In other words weight 6 or less is impossible and the minimum weight must be 8. In Figure 4 we give examples of the four distinct ways in which a codeword of weight 8 can occur, and work out the number of each.

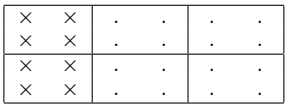
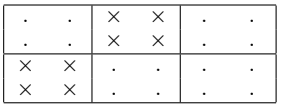
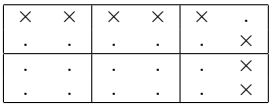
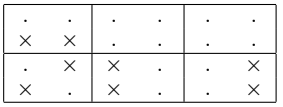
	Symbol	$\mathcal{C}$-set	Calculation	Number
(i)	$[O', O, O]_o$		3	3
(ii)	$[A, A', O]_o$		$3.7.2^2$	84
(iii)	$[D', D', O']_d$		7.3.4.2	168
(iv)	$[A, B, C]_a$		7.3.4.2	504
			Total	759

FIGURE 4. *Examples of the 759 codewords of $\mathcal{C}$ of weight 8*

(i) Choosing the zero element of $\mathcal{L}$ and O or I of $\mathcal{P}'$ we can have all 1s in one of the three bricks and 0s elsewhere.
(ii) We can again choose the zero of $\mathcal{L}$ and any P or P' in $\mathcal{P}'$ repeated (or complemented) in two of the three bricks; thus there are $7.3.2^2 = 84$ such codewords.
(iii) We may choose any non-zero $l \in \mathcal{L}$ and any $P \in \mathcal{P}'$ such that $P \notin l$. We take P or its complement such that $|P + l| = 2$ and place it in two of the bricks. Complementation on the third brick is allowed and so we obtain $7.4.3.2 = 168$ such codewords.
(iv) A fixed line $l \in \mathcal{L}$ will intersect any other line $\{P, Q, R\}$ in one point, P say. Then as above $|P + l| = 4$ and, choosing $Y = Q$ or $Y = Q'$, $Z = R$ or $Z = R'$ so that $|Y + l| = |Z + l| = 2$ we obtain a codeword of weight 8. There are thus $7.6.3.2.2 = 504$ such codewords.

The supports of these 759 codewords of weight 8 are known as *octads* and we may immediately deduce that if two distinct octads had more than 4 points in common then the sum of their corresponding vectors would have weight less than or equal to 6. Since no such codewords exist we conclude that any set of 5 points can lie in at most one octad. But the total number of 5 element subsets of a 24-element set is

$$\binom{24}{5} = 24.23.11.7 = 759.56 = 759.\binom{8}{5},$$

the number of 5-element subsets of some octad of the system. Thus every 5-element subset of the 24 points is contained in a unique octad: the octads form a *Steiner system* S(5,8,24). Note that the construction also implies that the symmetric difference of two octads which intersect in 4 points is another octad. This construction of the binary Golay code is taken from Curtis [**3**].

3.3. *The code $\mathcal{C}$ used in space missions*

The space probe Voyager 1 was launched by NASA on September the 5th 1977, following its sister ship Voyager 2 which had been launched on August the 20th the same year. The apparent discrepancy in their numbering is due to the fact that Voyager 1 travelled faster than Voyager 2 and overtook it, thus becoming the lead ship. Their original purpose was to explore Saturn and Jupiter, and to send back images of those planets. Transmitting messages over such vast distances is prone to interference and a robust and reliable error-correcting code was required. In fact the scientists in charge of the programme decided on the binary Golay code, and so $\mathcal{C}$ is responsible for some of the most beautiful and informative photographs of the planets in our solar system. In fact, although, the original mission was to these two planets, the Voyagers continued on their way and returned images of Uranus and Neptune. They remain the only spacecrafts to have visited these distant planets. In the last few months we have been told that they have now left our solar system and entered interstellar space, the first manmade objects to do so.

4. *The mathematical significance of $\mathcal{C}$*

The mathematical importance of the Golay code goes way beyond its error-correcting properties. In fact it is an extraordinarily symmetrical object which is preserved by the famous Mathieu group M_{24}, a group of 244823040 permutations of 24 letters. Moreover it is the main ingredient in the construction of the 24-dimensional *Leech lattice* Λ whose group of symmetries is the *Conway group* $\cdot\mathrm{O}$. Λ was discovered in connection with sphere-packing in n-dimensional space: it affords the best known *lattice* packing in $\mathbb{R}^{24}$, an arrangement of unit spheres in $\mathbb{R}^{24}$ such that their centres form a lattice. The *kissing number* or *Newton number* τ_n in n-space is the maximum number of unit spheres which can simultaneously touch a given unit sphere without overlapping one another. Determining τ_n is remarkably difficult even in low dimensions and it was only in 2007 that Musin proved that $\tau_4 = 24$. Values of τ_n when the problem is restricted to the centres of the spheres lying on a lattice have been determined up to dimension 8; in fact, the E_8 lattice affords a kissing number of 240 which has been shown to be the unrestricted maximum possible. However, beyond 8 no optimality is known until dimension 24 when the Leech lattice affords a kissing number of $\tau_{24} = 196560$ which is again the unrestricted maximum. Nor does the influence of $\mathcal{C}$ stop there as the Monster group M is usually constructed using properties of $\cdot\mathrm{O}$.

4.1. *The dodecahedron construction of $\mathcal{C}$*

The construction of $\mathcal{C}$ given here in terms of the Point space and the Line space lends itself to proving the important properties of the code. Moreover the resulting diagram of the code is the so-called *Miracle Octad Generator* or MOG which enables one to work with $\mathcal{C}$, M_{24}, Λ and $\cdot\mathrm{O}$. In fact the group M_{24} is also involved in the largest Janko group J_4 and in the Monster group M, and so the MOG is often used by people carrying out detailed investigations of those sporadic simple groups. However, there are many other ways of constructing $\mathcal{C}$ and we include here a brief description of how it can be obtained directly from the regular dodecahedron. In Figure 5 we give a diagram of the regular dodecahedron with its 12 faces labelled with the points of the projective line $\mathrm{P}_1(11)$. We may draw a graph Γ with 12 vertices corresponding to these faces, with two faces joined if they abut one another; thus Γ is a regular graph with

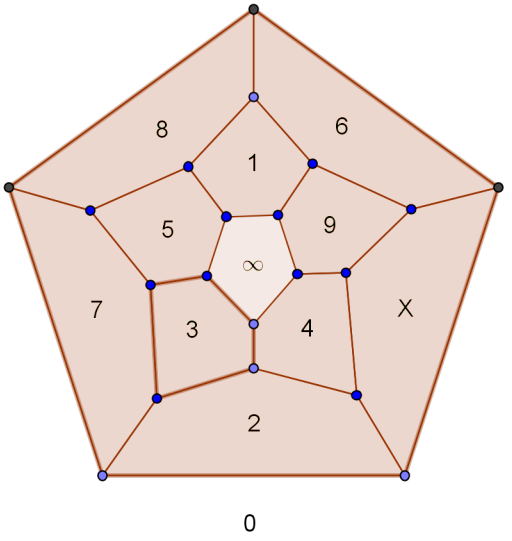


FIGURE 5. *Labelling of the faces of the dodecahedron*

	∞	0	1	2	3	4	5	6	7	8	9	X	∞	0	1	2	3	4	5	6	7	8	9	X
∞	1	.	.	.	.	.	.	.	.	.	.	.	1	1	.	1	.	.	.	1	1	1	.	1
0	.	1	.	.	.	.	.	.	.	.	.	.	1	1	1	.	1	1	1	.	.	.	1	.
1	.	.	1	.	.	.	.	.	.	.	.	.	.	1	1	1	1	1	.	.	1	.	.	1
2	.	.	.	1	.	.	.	.	.	.	.	.	1	.	1	1	.	.	1	1	.	1	1	.
3	.	.	.	.	1	.	.	.	.	.	.	.	.	1	1	.	1	.	.	1	.	1	1	1
4	.	.	.	.	.	1	.	.	.	.	.	.	.	1	1	.	.	1	1	1	1	1	.	.
5	.	.	.	.	.	.	1	.	.	.	.	.	.	1	.	1	.	1	1	1	.	.	1	1
6	.	.	.	.	.	.	.	1	.	.	.	.	1	.	.	1	1	1	1	1	1	.	.	.
7	.	.	.	.	.	.	.	.	1	.	.	.	1	.	1	.	.	1	.	1	1	.	1	1
8	.	.	.	.	.	.	.	.	.	1	.	.	1	.	.	1	1	1	.	.	.	1	1	1
9	.	.	.	.	.	.	.	.	.	.	1	.	.	1	.	1	1	.	1	.	1	1	1	.
X	.	.	.	.	.	.	.	.	.	.	.	1	1	.	1	.	1	.	1	.	1	1	.	1

FIGURE 6. *A generating matrix for $\mathcal{C}$ obtained from the dodecahedron*

valence 5. Let A denote the adjacency matrix of Γ, a 12×12 matrix with a 1 in the ijth position if vertex i is joined to vertex j, and 0s elsewhere, and interpret these entries as elements of the field $\mathbb{Z}_2$. Now let J denote the 12×12 all 1s matrix. Then, as is described in Curtis [**4**], the matrix

$$[I_{12} \mid A + J]$$

is a generating matrix for a copy of $\mathcal{C}$; that is to say, the 12 rows of this matrix generate a 12-dimensional subspace of $\mathbb{Z}_2^{24}$ which is isomorphic to $\mathcal{C}$. This 12×24 matrix is displayed in Figure 6.

From the above it is clear that the rich and complex combinatorial structure which lies at the heart of these 24-dimensional objects in some sense germinates in that most beautiful of the platonic solids. Indeed we have a direct route:

dodecahedron $\rightarrow$ $\mathcal{C}$ $\rightarrow$ M_{24} $\rightarrow$ the Leech lattice Λ $\rightarrow$ the Conway group $\cdot\mathrm{O}$ ($\rightarrow$ M).

References

1. E.F. Assmus Jr. and J.D. Key, "Designs and their Codes", Cambridge Tracts in Mathematics, vol. 103 (Cambridge, 1992)

2. J.H. Conway and N.J.A. Sloane, "Sphere Packings, Lattices and Groups", Grundlehren der Mathematischen Wissenschaften, vol. 290 (Berlin: Sringer, 1988)

3. R.T. Curtis, A new combinatorial approach to M_{24}, *Math. Proc. Cambridge Phil. Soc.*, **79** (1976), 25–42.

4. R.T. Curtis, The regular dodecahedron and the binary Golay code, *Ars Combinatoria* **29B** (1990), 55–64.

5. F. J. MacWilliams and N.J.A. Sloane, "The Theory of Error-Correcting Codes", North Holland Mathematical Library, vol. 16 (Amsterdam, 1977)

R.T.Curtis,
Department of Mathematics,
University of Birmingham.
UK

robcurtis.mog@gmail.com