

Title	A Stochastic Construction of Golay Code
Author(s)	Itoh, Yoshiaki; Jimbo, Masakazu
Citation	数理解析研究所講究録 (1987), 607: 76-82
Issue Date	1987-02
URL	<a href="http://hdl.handle.net/2433/99703">http://hdl.handle.net/2433/99703</a>
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

A Stochastic Construction of Golay Code

Yoshiaki Itoh

The Institute of Statistical Mathematics  
4-6-7 Minami-Azabu Minato-ku,  
Tokyo 106, Japan

Masakazu Jimbo

Department of Information Sciences  
Science University of Tokyo  
Noda City, Chiba 278,  
Japan

§1. Golay binary code and random packing

Leech (1964) gave the densest known packing of spheres in 24-dimensional space on the 24-digit binary sequences based on the Golay code, which is generated by 12 binary sequences. Hence the set is made up of  $2^{12}$  or 4096 points of binary sequences called code words. From each point of the 4096 points, there are 759 points of Hamming distance 8, 2576 points of Hamming distance 12, 759 points of Hamming distance 16, and one point of Hamming distance 24. The proof is given for example in the book by MacWilliams and Sloane (1977). Here we will discuss a stochastic algorithm to generate Golay code introduced by Itoh (1986).

Random packing by Hamming distance is discussed by Itoh and Solomon (1986). Consider a random sequential packing into the  $2^{24}$  points. At first we choose one point (24 coordinates) at random and

we record it. Choose another and record it if its Hamming distance is 8, 12, 16 or 24, otherwise reject it. Now, choose the next point at random and record it if the Hamming distance from each of the previously chosen 2 points is 8, 12, 16 or 24. When the points  $I_1, I_2, \dots, I_k$  are already chosen, the next point  $I_{k+1}$  will be chosen if each of the Hamming distance from each of the previously chosen  $I_1, I_2, \dots, I_k$  is 8, 12, 16 or 24. We continue this procedure until there is no possible point to record among the  $2^{24}$  points and we now have a set  $S$  of recorded points. The histogram of the number of recorded points of  $S$  is given in Fig. 1. Out of 550 trials, 11 trials produce Golay code of 4096 points. Without loosing generality, we can choose the point  $(0,0,\dots,0)$  as the first point. We get 12 rows from the second to the 13th by the random packing procedure of Hamming distance 8, 12, 16 or 24. Then we make the set of all possible sums of the binary sequences by the addition of modulo 2. If the set consists of  $2^{12}$  points and the distribution of Hamming distance from each point has that of the Golay code, then the 12 rows produce a linear code of minimum distance 8, which is equivalent to the code used by Leech to construct the Leech lattice. Hence this will give a stochastic construction of Golay code.

## §2. Random packing and finite field

Let  $V_n$  be an  $n$ -dimensional space over a finite field  $GF(2)$ . The weight of a vector  $a$  is denoted by  $wt(a)$ . If the weight of every code word of a linear code  $C$  is a multiple of 4 then the code is said to be doubly even. The following lemmas is obvious.

Lemma 1. For two vectors  $a$  and  $b$  in  $V_n$ , we have  $wt(a + b) = wt(a) + wt(b) - 2a \cdot b$ , where  $a \cdot b$  is the number of coordinates in which both elements of  $a$  and  $b$  are 1.

Lemma 2. Let  $a$  and  $b$  are vectors of  $V_n$  such that

$$wt(a) \equiv wt(b) \equiv d(a,b) \equiv 0 \pmod{4}.$$

Then we have  $a \cdot b \equiv a \cdot b \equiv 0 \pmod{2}$ .

Proof. By Lemma 1,

$$\begin{aligned} 2\mathbf{a}'\mathbf{b} &\equiv 2\mathbf{a}\cdot\mathbf{b} = \text{wt}(\mathbf{a}) + \text{wt}(\mathbf{b}) \\ -\text{wt}(\mathbf{a} + \mathbf{b}) &= \text{wt}(\mathbf{a}) + \text{wt}(\mathbf{b}) \\ -d(\mathbf{a},\mathbf{b}) &\equiv 0 \pmod{4}. \end{aligned}$$

**Theorem 1.** Let  $\mathbb{I}_1, \dots, \mathbb{I}_k$  be vectors of  $V_n$ . If the weight of every vector  $\mathbb{I}_i$  is a multiple of 4 and the Hamming distance between any two vectors  $\mathbb{I}_i$  and  $\mathbb{I}_j$  ( $i \neq j$ ) is also a multiple of 4, then the code generated by  $\mathbb{I}_1, \dots, \mathbb{I}_k$  is doubly even.

Proof. Let  $S$  be a set of vectors including zero vector.

Assume the distance of any two vectors are multiples of 4 for any two vectors  $\mathbf{a}$  and  $\mathbf{b}$  of  $S$ , we have

$$\text{wt}(\mathbf{a} + \mathbf{b}) \equiv d(\mathbf{a},\mathbf{b}) \equiv 0 \pmod{4}.$$

And

$$\mathbf{a}'\mathbf{b} \equiv \mathbf{a}\cdot\mathbf{b} \equiv 0 \pmod{2}$$

is obtained by Lemma 2. Hence for any vector  $\mathbf{c}$  of  $S$ ,

$$\begin{aligned} d(\mathbf{a} + \mathbf{b}, \mathbf{c}) &= \text{wt}(\mathbf{a} + \mathbf{b} + \mathbf{c}) \\ &\equiv \text{wt}(\mathbf{a} + \mathbf{b}) + \text{wt}(\mathbf{c}) - 2(\mathbf{a} + \mathbf{b})'\mathbf{c} \\ &\equiv \text{wt}(\mathbf{a} + \mathbf{b}) + \text{wt}(\mathbf{c}) - 2(\mathbf{a}'\mathbf{c} + \mathbf{b}'\mathbf{c}) \\ &\equiv 0 \pmod{4} \end{aligned}$$

holds, which shows that the set  $S' = S \cup \{\mathbf{a} + \mathbf{b}\}$  also has the property that the distance of any two vectors of  $S'$  are multiples of 4. Hence, by setting  $S = \{\mathbb{I}_1, \dots, \mathbb{I}_k, \mathbf{0}\}$  the linear code generated by  $S$  is shown to be doubly even.

Let  $S$  be a set of points (code words) obtained from the above random packing procedure and let  $\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_k$  be linearly independent code words which generate linear space  $C$  over a finite field  $GF(2)$ .

**Theorem 2.**

$$S \subset C^\perp \quad \text{where } C^\perp \text{ is the dual space of } C$$

Proof. For any  $\mathbf{J} \in S$

$$\mathbf{J}\cdot\mathbb{I}_i \equiv \mathbf{J}\cdot\mathbb{I}_i \equiv 0 \pmod{2}, \text{ which means } \mathbf{J} \in C^\perp.$$

We have  $C \subset C^\perp$ , as given in MacWilliams and Sloane (1977).

Hence,  $C = C^\perp$  if  $\dim(C) = k = \frac{n}{2}$ .

Hence when there are 12 linearly independent vectors  $\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_{12} \in S$ , each of the remained vectors of  $S$  are represented by a linear combination of the 12 vectors. This gives a simple algorithm for our random packing procedure. We continue the random packing procedure until we get 12 linearly independent vectors  $\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_{12}$ . Then we make the random packing into the  $2^{12}$  linear combinations of  $\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_{12}$ . This makes computational time of our simulation much shorter.

### §3. Weight distributions and clusters

The enumeration of self-dual codes by Conway and Pless (1980) gives the following possible weight distribution for  $n = 0, 1, 2, 3, 4, 5, 7, 11$ ,

0	4	8	12	16	20	24
1	$6n$	$759-24n$	$2576-136n$	$759-24n$	$6n$	1.

The code generated by the  $\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_{12}$  obtained in the previous section has one of the above 8 weight distributions. Hence the remained possible points after packing mutually independent 12 points should be not less than  $4096-(12n+1) \times 12$ , which will help to explain Fig. 2. The number of points obtained by our random packing procedure is not less than  $4096/(n+1)$ , if we can get the twelve independent vectors  $\mathbb{I}_1, \mathbb{I}_2, \dots, \mathbb{I}_{12}$ . When  $n=0$ , the twelve vectors generate Golay code. Each of the other clusters may correspond to one of the other seven weight distributions. Smaller  $n$  may correspond to a cluster of larger recorded numbers.

RECORDED POINTS	FREQUENCY (%)	HISTOGRAM OF NUMBERS OF RECORDED POINTS	NUMBERS OF RECORDED POINTS OBTAINED BY RANDOM SEQUENTIAL PACKING
1 - 10	0.1	740	740
11 - 20	0.2	740	740
21 - 30	0.3	740	740
31 - 40	0.4	740	740
41 - 50	0.5	740	740
51 - 60	0.6	740	740
61 - 70	0.7	740	740
71 - 80	0.8	740	740
81 - 90	0.9	740	740
91 - 100	1.0	740	740
101 - 110	1.1	740	740
111 - 120	1.2	740	740
121 - 130	1.3	740	740
131 - 140	1.4	740	740
141 - 150	1.5	740	740
151 - 160	1.6	740	740
161 - 170	1.7	740	740
171 - 180	1.8	740	740
181 - 190	1.9	740	740
191 - 200	2.0	740	740
201 - 210	2.1	740	740
211 - 220	2.2	740	740
221 - 230	2.3	740	740
231 - 240	2.4	740	740
241 - 250	2.5	740	740
251 - 260	2.6	740	740
261 - 270	2.7	740	740
271 - 280	2.8	740	740
281 - 290	2.9	740	740
291 - 300	3.0	740	740
301 - 310	3.1	740	740
311 - 320	3.2	740	740
321 - 330	3.3	740	740
331 - 340	3.4	740	740
341 - 350	3.5	740	740
351 - 360	3.6	740	740
361 - 370	3.7	740	740
371 - 380	3.8	740	740
381 - 390	3.9	740	740
391 - 400	4.0	740	740
401 - 410	4.1	740	740
411 - 420	4.2	740	740
421 - 430	4.3	740	740
431 - 440	4.4	740	740
441 - 450	4.5	740	740
451 - 460	4.6	740	740
461 - 470	4.7	740	740
471 - 480	4.8	740	740
481 - 490	4.9	740	740
491 - 500	5.0	740	740
501 - 510	5.1	740	740
511 - 520	5.2	740	740
521 - 530	5.3	740	740
531 - 540	5.4	740	740
541 - 550	5.5	740	740
551 - 560	5.6	740	740
561 - 570	5.7	740	740
571 - 580	5.8	740	740
581 - 590	5.9	740	740
591 - 600	6.0	740	740
601 - 610	6.1	740	740
611 - 620	6.2	740	740
621 - 630	6.3	740	740
631 - 640	6.4	740	740
641 - 650	6.5	740	740
651 - 660	6.6	740	740
661 - 670	6.7	740	740
671 - 680	6.8	740	740
681 - 690	6.9	740	740
691 - 700	7.0	740	740
701 - 710	7.1	740	740
711 - 720	7.2	740	740
721 - 730	7.3	740	740
731 - 740	7.4	740	740
741 - 750	7.5	740	740
751 - 760	7.6	740	740
761 - 770	7.7	740	740
771 - 780	7.8	740	740
781 - 790	7.9	740	740
791 - 800	8.0	740	740
801 - 810	8.1	740	740
811 - 820	8.2	740	740
821 - 830	8.3	740	740
831 - 840	8.4	740	740
841 - 850	8.5	740	740
851 - 860	8.6	740	740
861 - 870	8.7	740	740
871 - 880	8.8	740	740
881 - 890	8.9	740	740
891 - 900	9.0	740	740
901 - 910	9.1	740	740
911 - 920	9.2	740	740
921 - 930	9.3	740	740
931 - 940	9.4	740	740
941 - 950	9.5	740	740
951 - 960	9.6	740	740
961 - 970	9.7	740	740
971 - 980	9.8	740	740
981 - 990	9.9	740	740
991 - 1000	10.0	740	740
1001 - 1010	10.1	740	740
1011 - 1020	10.2	740	740
1021 - 1030	10.3	740	740
1031 - 1040	10.4	740	740
1041 - 1050	10.5	740	740
1051 - 1060	10.6	740	740
1061 - 1070	10.7	740	740
1071 - 1080	10.8	740	740
1081 - 1090	10.9	740	740
1091 - 1100	11.0	740	740
1101 - 1110	11.1	740	740
1111 - 1120	11.2	740	740
1121 - 1130	11.3	740	740
1131 - 1140	11.4	740	740
1141 - 1150	11.5	740	740
1151 - 1160	11.6	740	740
1161 - 1170	11.7	740	740
1171 - 1180	11.8	740	740
1181 - 1190	11.9	740	740
1191 - 1200	12.0	740	740
1201 - 1210	12.1	740	740
1211 - 1220	12.2	740	740
1221 - 1230	12.3	740	740
1231 - 1240	12.4	740	740
1241 - 1250	12.5	740	740
1251 - 1260	12.6	740	740
1261 - 1270	12.7	740	740
1271 - 1280	12.8	740	740
1281 - 1290	12.9	740	740
1291 - 1300	13.0	740	740
1301 - 1310	13.1	740	740
1311 - 1320	13.2	740	740
1321 - 1330	13.3	740	740
1331 - 1340	13.4	740	740
1341 - 1350	13.5	740	740
1351 - 1360	13.6	740	740
1361 - 1370	13.7	740	740
1371 - 1380	13.8	740	740
1381 - 1390	13.9	740	740
1391 - 1400	14.0	740	740
1401 - 1410	14.1	740	740
1411 - 1420	14.2	740	740
1421 - 1430	14.3	740	740
1431 - 1440	14.4	740	740
1441 - 1450	14.5	740	740
1451 - 1460	14.6	740	740
1461 - 1470	14.7	740	740
1471 - 1480	14.8	740	740
1481 - 1490	14.9	740	740
1491 - 1500	15.0	740	740
1501 - 1510	15.1	740	740
1511 - 1520	15.2	740	740
1521 - 1530	15.3	740	740
1531 - 1540	15.4	740	740
1541 - 1550	15.5	740	740
1551 - 1560	15.6	740	740
1561 - 1570	15.7	740	740
1571 - 1580	15.8	740	740
1581 - 1590	15.9	740	740
1591 - 1600	16.0	740	740
1601 - 1610	16.1	740	740
1611 - 1620	16.2	740	740
1621 - 1630	16.3	740	740
1631 - 1640	16.4	740	740
1641 - 1650	16.5	740	740
1651 - 1660	16.6	740	740
1661 - 1670	16.7	740	740
1671 - 1680	16.8	740	740
1681 - 1690	16.9	740	740
1691 - 1700	17.0	740	740
1701 - 1710	17.1	740	740
1711 - 1720	17.2	740	740
1721 - 1730	17.3	740	740
1731 - 1740	17.4	740	740
1741 - 1750	17.5	740	740
1751 - 1760	17.6	740	740
1761 - 1770	17.7	740	740
1771 - 1780	17.8	740	740
1781 - 1790	17.9	740	740
1791 - 1800	18.0	740	740
1801 - 1810	18.1	740	740
1811 - 1820	18.2	740	740
1821 - 1830	18.3	740	740
1831 - 1840	18.4	740	740
1841 - 1850	18.5	740	740
1851 - 1860	18.6	740	740
1861 - 1870	18.7	740	740
1871 - 1880	18.8	740	740
1881 - 1890	18.9	740	740
1891 - 1900	19.0	740	740
1901 - 1910	19.1	740	740
1911 - 1920	19.2	740	740
1921 - 1930	19.3	740	740
1931 - 1940	19.4	740	740
1941 - 1950	19.5	740	740
1951 - 1960	19.6	740	740
1961 - 1970	19.7	740	740
1971 - 1980	19.8	740	740
1981 - 1990	19.9	740	740
1991 - 2000	20.0	740	740

Fig. 1 Histogram

(The 11 trials of 4096 recorded points are not given here.)

(from Itoh(1986))

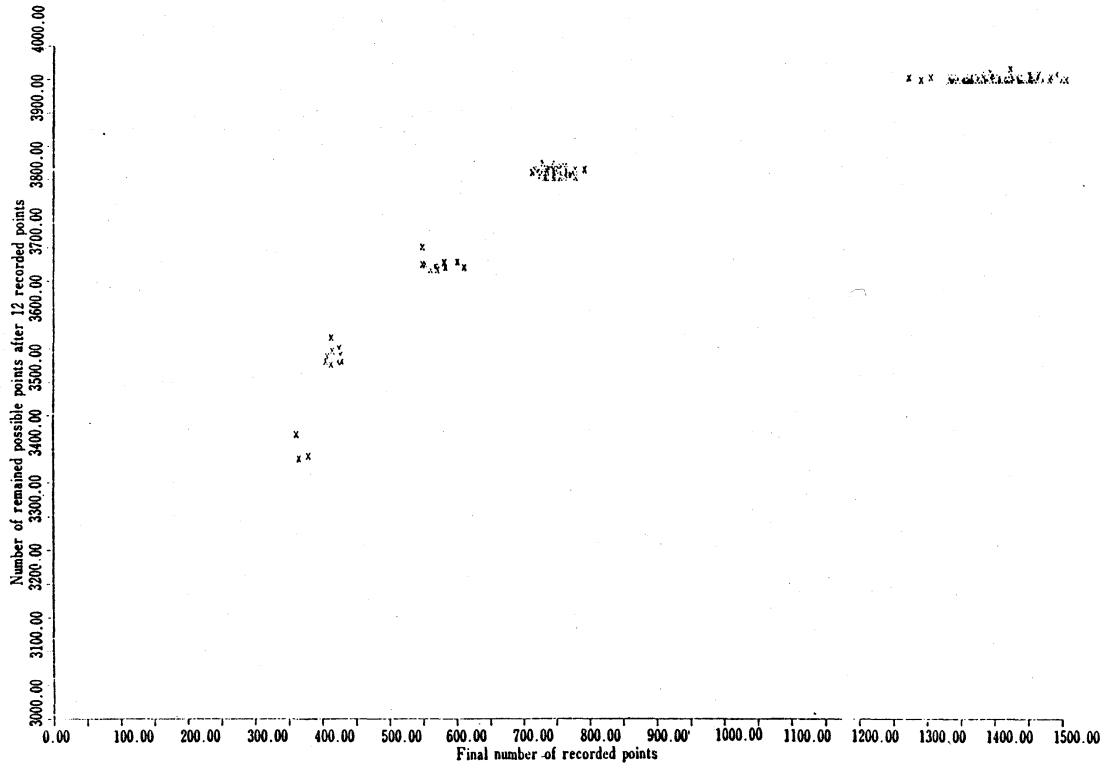


Fig. 2 The first 12 points and the clusters.

(from Itoh(1986))

## References

- Conway, J.H. and Pless, V (1980). On the enumeration of self-dual codes, *Journal of Combinatorial Theory, Series A* 28, 26-53.
- Itoh, Y.(1986). Golay code and random packing, *Ann. Inst. Statist. Math.* 38, 583-588.
- Itoh, Y. and Solomon, H.(1986). Random sequential coding by Hamming distance, *J. Appl. Prob.* 23, 688-695.
- Leech, J.(1964). Some sphere packings in Higher space, *Canad. J. Math.*, 16, 657-682.
- MacWilliams, E.J. and Sloane, N.J.A.(1977). *The theory of error-correcting codes*, I, II, North-Holland.