What's so special about the Golay codes?

David Wilding, March 2011

http://dpw.me/mathematics/

By the time NASA's 20th century *Voyager* missions reached Jupiter and Saturn, the weak radio signals carrying pictures taken by the two spacecraft were being plagued by interference. To address this problem a method of error correction was used, which allowed receivers back on Earth to recover the original, uncorrupted, pictures.

The fundamental concept in error correction is that of a *codeword*, which is just a string of digits. To correct errors we choose a collection of codewords that are quite different from each other, so that when one is received it is impossible to mistake it for any other codeword if it has been corrupted slightly. Such a set of codewords is called a q-ary *code*, where q determines how many different values the digits of each codeword can take. For example, the codewords in a 2-ary (or *binary*) code are strings of '0's and '1's, and the codewords in a 10-ary code are strings of the usual decimal digits '0' to '9'.

One of the most basic codes is the binary code comprising the two codewords '000' and '111'. If we send either of these codewords and at most one of the digits gets changed, so that '000' becomes '001' say, then the receiver can still recognise which codeword was sent. However, if at most two of the digits get changed then there is some ambiguity, and the receiver will not necessarily pick the correct codeword.

For codewords with a given number of digits, a good code should both have a relatively large number of codewords and yet be able to tolerate mistakes in many digits at once. These two goals are in conflict because the more codewords a code has, the less different they can all be from each other; when designing codes we always have to compromise.

The code used in the *Voyager* missions has codewords that are 24-digit long binary strings, and if at most three of the digits in a codeword get changed then we can still identify the codeword. This code is a variant of one introduced in 1949 by electrical engineer Marcel Golay. Golay's code is slightly different from the *Voyager* code in that the codewords are all one digit shorter, but they can still tolerate mistakes in up to three digits at once. An interesting feature of Golay's code is that if four or more of the digits in a codeword get changed then we are guaranteed to incorrectly identify the codeword.

Codes with this property are called *perfect*, and are rather rare. Golay's code is now known as the *perfect binary Golay code* and the *Voyager* code is called the *extended binary Golay code*.

One way to list all of the codewords in the extended binary Golay code is to use a *Steiner system*. In general, to form a Steiner system we begin with a set of n *points*. Then we collect together *blocks* of m of these points, with the property that every set of l points is contained in exactly one of these blocks. The resulting structure is called an $\mathrm{S}(l, m, n)$ system, and the classic example of such a system is the $\mathrm{S}(2, 3, 7)$ *Fano plane* in Figure 1. There are seven points in total, the seven blocks (all the lines, including the circle) each contain three points, and every pair of points is on exactly one line.

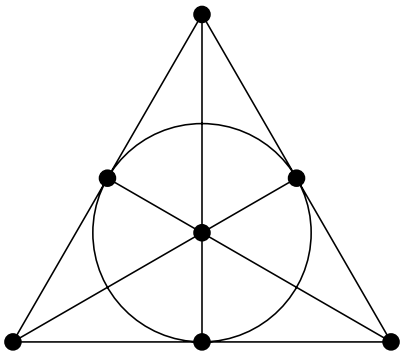

Figure 1: An $\mathrm{S}(2, 3, 7)$ Steiner system.

The large and small *Witt designs*, named after mathematician Ernst Witt, are two particularly interesting Steiner systems. The large Witt design is an $\mathrm{S}(5, 8, 24)$ system, and is the orbit of one of its blocks under an action of the *projective special linear group* $\mathrm{PSL}(2, 23)$. Similarly, the small Witt design is an $\mathrm{S}(5, 6, 12)$ system, and is the orbit of one of its blocks under an action of $\mathrm{PSL}(2, 11)$.

If we regard the small Witt design as a system whose points are the numbers 1 to 12 then it is possible to interpret its blocks as 12-digit long binary codewords. For instance, the block $\{1, 5, 7, 8, 10, 11\}$ corresponds to the codeword 100010110110 by writing '1' in the 1st, 5th, 7th, 8th, 10th and 11th places, and by writing '0' everywhere else. An analogous approach allows us to view the blocks of the large Witt design as 24-digit long binary codewords. Among these 24-digit codewords there are 12 that have exactly one '1' and eleven '0's in the first 12 places. These 12 codewords happen to be in the extended binary Golay code, and they are special because we can determine the remaining 4084 codewords in the extended binary Golay code using them. Specifically the 12 codewords are a *basis* for the extended binary Golay code.

A q-ary code whose codewords form a vector space under digit-wise addition and scalar multiplication (modulo q) is called *linear*, and every linear code has a basis of codewords. Linear codes are desirable because we can use techniques from linear algebra to manipulate them and reason about them. In particular a basis provides a compact description of the codewords in a linear code, which we can use to build a list of all the codewords.

By carefully selecting six of the 12-digit codewords that correspond to the blocks of the small Witt design we can obtain a basis for the *extended ternary Golay code*, which is a 3-ary code whose codewords are strings of '0's, '1's and '2's. This code is a variant of another code introduced by Golay, called the *perfect ternary Golay code*, and its codewords can tolerate mistakes in up to two digits at once.

We can use the codewords in the extended Golay codes to list the blocks of the two Witt designs. In the case of the extended binary Golay code, the codewords that have eight non-zero digits correspond to the blocks of the large Witt design. Similarly in the case of the extended ternary Golay code, the codewords that have six non-zero digits correspond to the blocks of the small Witt design. This relationship between the extended Golay codes and the Witt designs links the Golay codes to the *classification of the finite simple groups*—a remarkable achievement of 20th century pure mathematics.

The classification of the finite simple groups asserts that there are several families of finite simple groups, such as the cyclic groups of prime order and most of the alternating groups. It also says that there are 26 so-called *sporadic* finite simple groups that do not belong to any of the families. Five of these sporadic groups, the *Mathieu groups*, were discovered by mathematician Émile Mathieu in the 19th century, and are intimately linked to the Witt designs.

An *automorphism* of an $\mathrm{S}(l, m, n)$ Steiner system is a permutation of the n points that sends each block of the system to some (possibly different) block. The number of distinct automorphisms of a given Steiner system can be enormous: the large Witt design, for example, has almost a quarter of a billion!

The automorphisms of a Steiner system always form a group, called the *automorphism group* of the system, and the automorphism groups of the Witt designs turn out to be two of the Mathieu groups. We can obtain the other three Mathieu groups by considering smaller Steiner systems related to the Witt designs.

References

[1] P. J. Cameron and J. H. van Lint. *Designs, Graphs, Codes and their Links*. Cambridge University Press, Cambridge, 1991.

[2] T. M. Thompson. *From Error-Correcting Codes through Sphere Packings to Simple Groups*. The Mathematical Association of America, 1983.