Classification of Boolean forms in 8 variables

Gregor Leander, Philippe Langevin

ASI, Moscow, September 2007.

Boolean forms

A Boolean form of degree k in m variable

$$h(x) = \sum_{|S|=k} a_S X_S, \quad X_S = \prod_{i \in S} x_i, \quad a_s \in \mathbb{F}_2.$$

It is a space of dimension $\binom{m}{k}$, denoted by

$$\mathrm{RM}^*(k, m) = \mathrm{RM}(k, m)/\mathrm{RM}(k-1, m).$$

You must think h has a Boolean form, not has a Boolean function, for all Boolean functions g with $\deg(g) < k$ the symboles h and $h + g$ represent the same object!

Action of GL(2,m)

The translation group acts trivially on the Boolean forms in the sense that

$$h(x+u) \equiv h(x) \mod \mathrm{RM}(k-1,m).$$

The group $\mathrm{GL}(2,m)$ acts over the forms.

$$h \sim g \iff \exists A \in \mathrm{GL}(2,m), \quad h^A(x) = h(xA) = g(x) + f(x)$$

where $\deg f < k$.

- $n(k,m)$ the rank of $\mathrm{GL}(2,m)$ over $\mathrm{RM}^*(k,m)$.
- $\mathrm{cl}(k,m)$ a set of representatives.
- $\mathrm{orb}_m^k(h)$ the orbit of h.
- $\mathrm{fix}(h) = \{A \mid h^A = h\}$.

$$2^{\binom{m}{k}} = \sum_{f \in \mathrm{cl}(k,m)} \sharp \mathrm{orb}(f)$$

Baby example

The linear group has order

$$\sharp \mathrm{GL}(2,m) = \prod_{i=0}^{m-1} (2^m - 2^i) \approx 0.33 \times 2^{m^2}$$

and is generated by the "rotation" R and a transvection T.
In dimension 3,

$$T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad R = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Consider the quadratic form

$$h(x,y,z) = xz + xy \in \mathrm{RM}^*(2,3).$$

$$h^T(x,y,z) = h\big((x,y,z)T\big) = h(x,y,x+z) = x(x+z)+xy = xz+xy$$

$$h^S(x,y,z) = h\big((x,y,z)S\big) = h(y,z,x) = yx + yz$$

Classification

By a classification of $\mathrm{RM}^*(k, m)$ under the action of $\mathrm{GL}(2, m)$, we understand the determination of :

1. The number of class $n(k, m)$
2. A set of representative $\mathrm{cl}(k, m)$
3. The size of each orbits $\mathrm{orb}_m^k(f)$, $f \in \mathrm{cl}(k, m)$
4. A system of generators of the groups $\mathrm{fix}(f)$
5. Eventualy, the sructure of these groups...

In general, it is impossible to proceed randomly because of the (probable) existence of (unknown) orbits of small size.

Compound matrix

The action of A on the monomial X_S is given

$$X_S{}^A(x) = \prod_{i \in S} \left(\sum_{j=1}^{m} a_{ij} x_j \right) \tag{1}$$

$$= \sum_{T} \sum_{j \colon S \to T} \prod_{i \in S} a_{ij(i)} X_T \tag{2}$$

$$= \sum_{T} \det A_{S,T} X_T \tag{3}$$

The compound k-th matrix of A

$$C^r(A) = \left(\det A_{S,T} \right)$$

where $A_{S,T}$ is the square matrix of order k obtained keeping the lines of index $i \in S$ and columns of index $j \in T$.

Baby example (continuing)

The 2-th compound of

$$T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

in the basis (yz, xz, xy) is

$$C^2(T) = \begin{pmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{pmatrix}$$

Baby example (continuing)

The 2-th compound of

$$T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

in the basis (yz, xz, xy) is

$$C^2(T) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

In particular xz and yx are invariants, also h!

Complementary map

The complementary map

$$\mathrm{RM}^*(k,m) \ni X_S \to X_{\bar{S}} \in \mathrm{RM}^*(m-k,m)$$

We have the following commutative diagram

$$\begin{array}{ccc} \mathrm{RM}^*(k,m) & \xrightarrow{\text{comp}} & \mathrm{RM}^*(m-k,m) \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle A^{-1*}} \\ \mathrm{RM}^*(k,m) & \xrightarrow{\text{comp}} & \mathrm{RM}^*(m-k,m) \end{array} \tag{4}$$

In particular,

$$f \sim g \iff \mathrm{comp}\,(f) \sim \mathrm{comp}\,(g).$$

whence $n(k,m) = n(m-k,m)$.

- See X.-D. Hou papers.

Number of class

Using Burnside Lemma,

$$n(k,m)|\mathrm{GL}(2,m)| = \sum_{A \in \mathrm{GL}(2,m)} t(A)$$

where $t(A)$ is the number of forms fixed by A.

$$n(k,m) = \sum_{i=1}^{t} \frac{2^{\binom{m}{k} - \mathrm{rank}\left(C^k(A_i) - I\right)}}{\gamma(A_i)}$$

where A_i is a list of representatives of the conjugacy classes of $\mathrm{GL}(2,m)$, I the $\binom{m}{k} \times \binom{m}{k}$-identity matrix and $\gamma(A)$ the order of the centralizer of A in $\mathrm{GL}(2,m)$.

The enumeration of all the irreducible polynomials of degree less or equal to m allows the construction of all the possible invariant factors...

Numerical table

Table: Number of $\mathrm{GL}(2, m)$-orbits in $\mathrm{RM}^*(k, m)$.

$k \backslash m$	6	7	8	9	10
3	6	12	32	349	3691561
4	3	12	999	$\sim 10^{15}$	$\sim 10^{34}$

- Hou, Sugita, Kasami and Fujiwara classified $\mathrm{RM}^*(3, 8)$ (1996).
- Brier & PL classified $\mathrm{RM}^*(3, 9)$ (2003).
- In this talk, I explain the end of the classification of $\mathrm{RM}^*(4, 8)$, continuing the works of Rabizonni, Véron, Zanotti & PL (2006).

Invariant

An invariant of degree k in m variables is a mapping j such that

$$f \sim g \longrightarrow \mathrm{j}(f) = \mathrm{j}(g)$$

- A basic invariant

  $\mathrm{var}\,(f) =$ minimal number of variables of f up to equivalence

- Note that

  $$\mathrm{j} \circ \mathrm{comp}$$

  is an invariant of degree $m - k$.

Multiplicative invariants

Let $h \in \mathrm{RM}^*(k, m)$.
For all i, $0 \leq i \leq m - k$, we construct an invariant $\mathfrak{R}_{i,k}$ in considering the mulptiplication by h over the forms of degree i :

$$\mathrm{RM}^*(i, m) \ni f \xrightarrow{M_h^i} fh \in \mathrm{RM}^*(k + i, m) \tag{5}$$

$$\mathfrak{R}_{i,k}(h) = \dim \ker M_h^i$$

Clearly, it takes at most $\binom{m}{i}$ values.

A fundamental Invariant

Let $h \in \mathrm{RM}^*(2, m)$.
By definition, the rank of the quadratic form h is the dimension of the radical of h:

$$\mathrm{rk}\,(h) = \dim\{y \in \mathbb{F}_2^m \mid \forall x, \quad h(x+y) + h(x) + h(y) = 0\}$$

Of course,

$$h \mapsto \mathfrak{q}(h) = \mathrm{rk}\,(q)$$

is an invariant.
For parity reason, it takes at most $\frac{m}{2}$ values.

Invariant quadratic

Assume that $m = 2t$,

$$h = \sum_S a_S X_S \in \mathrm{RM}^*(t, m)$$

then

$$h \mapsto \mathfrak{Q}(h) = \sum_{\{S,\bar{S}\}} a_S a_{\bar{S}} \mod 2$$

in an invariant.

- $\exists g \in \mathrm{RM}^*(t-1, m)$ such that $h + g$ is bent $\Rightarrow \mathfrak{Q}(h) = 0$

Derivation

The derivation of $f \in \mathrm{RM}^*(k, m)$ is the homogeneous form of degree $k-1$ obained by reducing

$$\mathrm{der}_u(f) = f(x+u) + f(x) \equiv \sum_{i=1}^{m} f(x1, x2, \ldots, u_i, \ldots, x_m)$$

- D_0 if $f \neq 0$ then $\deg(\mathrm{der}_u f) < \deg(f)$
- D_1 $\mathrm{der}_u(f+g) = \mathrm{der}_u f + \mathrm{der}_u g$
- D_2 $\mathrm{der}_u(f \circ A) = (\mathrm{der}_{uA} f) \circ A$
- D_3 $\mathrm{der}_{u+v} f = \mathrm{der}_u f + \mathrm{der}_v f + \mathrm{der}_{u,v} f$

Remark : we will not use the notion of restriction.

Invariant Transvection

Recall that a transvection T is defined by a pair $(\phi, u) \in \mathbb{F}_2^{m*} \times \mathbb{F}_2^m$ such that $\phi(u) = 0$:

$$T(x) = x + \phi(x)u$$

Since the set of transvections is invariant by conjugaison in $\mathrm{GL}(2, m)$

$$h \mapsto \mathfrak{T}(h) = \sharp\{T \mid h^T = h\}$$

in an invariant. It is easy to compute since

$$h(x+\phi(x)u) = h(x)(1+\phi(x))+h(x+u)\phi(x) = \phi(x)\mathrm{der}_u h(x)+h(x)$$

so considering the mapping

$$\Psi_u : \phi \mapsto (\phi.\mathrm{der}_u h, \phi(u))$$

we have $\mathfrak{T}(h) = \sum_{u \in \mathbb{F}_2^m} \dim \ker \Psi_u$

Lift by derivation

Let j be an invariant of degree $k-1$ in m variables. We construct an invariant of degree k. Using the rule :

$$\mathrm{der}_u(h \circ A) = (\mathrm{der}_{uA} h) \circ A$$

- the distribution of the values of the mapping $u \mapsto (\mathrm{der}_u h)$, denoted $\mathrm{j}'(h)$ is constant over the class of h. It is the lift by derivation of j.
- Note that the distribution of the values of the Fourier coefficients is also constant. denoted $\widehat{\mathrm{j}'}(h)$.
- In practice, $\widehat{\mathrm{j}'}$ is more discriminant than j'.

Classification of $\mathrm{RM}^*(3,7)$

The lift by derivation of $\mathfrak{q}$ discriminates the 12 class of $\mathrm{RM}^*(3,7)$.

orb. size	fix. size	cubic
1	163849992929280	0
11811	13872660480	123
1763776	92897280	137+237+147+247+157+267+467
2314956	70778880	145+123
45354240	3612672	123+456
59527440	2752512	123+245+346
21165312	7741440	123+145+246+356+456
238109760	688128	124+235+346+457+561+267+137
444471552	368640	712+724+134+234+135+745+146
2222357760	73728	127+123+147+245+167
13545799680	12096	127+123+234+345+456+567+617
17777862080	9216	127+234+125+457+245+167+126

By the complementary map, we get the classification of $\mathrm{RM}^*(4,7)$.

Classification of $\mathrm{RM}^*(4,8)$

The space of quartic forms has order 2^{70}. We want apply combinaisons of the above invariants to this space.
Let $\mathfrak{L}$ the Fourier lift of $\mathfrak{q}'$ (double lift). This invariant takes 951 values.

Table: Number of values taken by the combinaisons of the invariants.

	$\mathfrak{Q}$	$\mathfrak{R}$	$\mathfrak{T}$	$\mathfrak{Q}\&\mathfrak{R}$	$\mathfrak{Q}\&\mathfrak{T}$	$\mathfrak{R}\&\mathfrak{T}$	$\mathfrak{Q}\&\mathfrak{R}\&\mathfrak{T}$
$\mathfrak{L}$	963	951	914	965	963	951	965
	2	14	53	22	56	90	118

The work factor for the computation of these invariants and their complementary is about 2^{20}, we have to reduce the space of quartics...

First reduction

Let $h \in \mathrm{RM}^*(4,8)$, we decompose:

$$h = f + x_8 g, \quad f \in \mathrm{RM}^*(4,7), \quad g \in \mathrm{RM}^*(3,7)$$

and $\mathrm{GL}(2,7)$ acts naturally over $\mathrm{RM}^*(4,8)$,

$$B \in \mathrm{GL}(2,7), \quad h \sim f^B + x_8 g^B = h^A, \quad A = \begin{pmatrix} B & 0 \\ 0 & 1 \end{pmatrix}$$

The set of pairs $(f,g) \in \mathrm{cl}(4,7) \times \mathrm{RM}^*(3,7)$ provides a reduction of $\mathrm{RM}^*(4,8)$ of size

$$12 \times 2^{\binom{7}{3}} = 12 \times 2^{35} = 412,316,860,416$$

each pair (f,g) represents $\sharp \mathrm{orb}_7^4(f)$ elements.
It is too large!

Second reduction

Let $h \in \mathrm{RM}^*(4,8)$, then

$$h = f + x_8 g \sim f + x_8 g + x_8 \mathrm{der}_v f$$

where $f \in \mathrm{RM}^*(4,7)$, $g \in \mathrm{RM}^*(3,7)$ and $v \in \mathbb{F}_2^{m-1}$
We introduce the space

$$\Delta(f) = \{\mathrm{der}_v f \mid v \in \mathbb{F}_2^{m-1}\},$$

Note that

$$\dim \Delta(f) = \mathrm{var}(f)$$

In particular, the set of pairs

$$(f, g) \in \mathrm{cl}(4,7) \times \mathrm{RM}^*(3,7)/\Delta(f)$$

give a new reduction.
Each pair (f, g) represents $\sharp \mathrm{orb}_7^4(f) \times 2^{\mathrm{var}(f)}$, but the size is about

$$6,442,450,944$$

Last reduction

- The group $\mathrm{fix}\,(f)$ acts over $\mathrm{RM}^*(k,m)/\Delta(f)$.

For all $f \in \mathrm{cl}(4,7)$, we compute a set of representatives of $\mathrm{RM}^*(3,7)/\Delta(f)$ under the action of $\mathrm{fix}\,(f)$.

The set

$$f + x_8 g, \quad f \in \mathrm{cl}(4,7), g \in \mathrm{RM}^*(3,7)/\Delta(f)/\mathrm{fix}\,(f)$$

provides a reduction $\mathrm{RM}^*(4,8)$, of size

68647.

Each pair (f,g) represents

$$\sharp\mathrm{orb}_7^4\,(f) \times 2^{\mathrm{var}\,(f)} \times \sharp\mathrm{orb}\,(g/\mathrm{fix}\,(f))$$

Applying invariants

The product *sayj* of invariants presented here takes 966 distincts values. It do not disriminate the 999 class.

$$X_v = j^{-1}(v)$$

There are a few collisions. Some of them are easy to detect, when

$$\sharp X_v \quad \text{does not divide} \quad \sharp \mathrm{GL}(2, m)$$

Detecting collisions

The cardinality of an orbit is smaller than 2^{64}. We can apply the birthday paradox to separate X_v in subclass. For all the representatives $h \in X_v$, we calculate partial orbits

$$h \mapsto h^A = f + x_8 g \mapsto f + x_8 g', \quad g' = g \mod \Delta(f)$$

taking A randomly in $\mathrm{GL}(2, m)$.

- The generated forms are in a set of size $\approx 2^{56}$, so computing about 2^{27} random actions in a set X_v, we have a good probability to prove equivalence between different representatives.
- We continue this random generation until 999 subclass are found.

30 Collisions

There are 30 collisions :

- 3 sets X_v containing 3 class
- 27 sets containing 2 class

The number of class is

$$966 + 27 + 6 = 999$$

Example of collision:
Inv=Q:0:R:8-24:T:0:L:26:Q:0:R:8-24:T:0:L:26
Q=2357+1457+1367+1467+2467+1248+1348+1458+3458+1568+
Orb=191002277471846400 Fix=28
Q=2357+1457+1367+1467+2467+1238+1348+3458+1568+
Orb=76400910988738560O Fix=7

Computation of fixators

If $A \in \mathrm{fix}\,(h)$ then A fixes the mapping

$$u \mapsto \mathrm{der}_u f\,.$$

In general, this map takes a lot of distincts values and it is possible to enumerate all the $A \in \mathrm{GL}(2, m)$ that fixes it. For a such A, we check if

$$h^A = h$$

If it is the case, we update a list of generators using the method of Schreier.

Number of class up to complementary

The previous algorithm can be modified to check the equivalence between two homogeneous forms.

	auto	not
$\mathfrak{Q} = 1$	294	168
$\mathfrak{Q} = 0$	300	236

In particular, there are 418 class of homogeneous forms h, up to complementary, with $\mathfrak{Q}(h) = 0$.

Covering radius of Reed-Muller code

The covering radius of Reed-Muller codes are not known in general. The handbook of coding theory says

$$44 \leq \rho(3, 8) \leq 67$$

Let $h \in \mathrm{RM}^*(k, m)$, and let $g \in \mathrm{RM}(k-1, m)$:

$$\mathrm{wt}\,(h+g) = 2^{m-1} - \frac{1}{2} S(h+g), \quad S(f) = \sum_{x \in \mathbb{F}_2^m} (-1)^f(x).$$

Adapting a recent trick of Carlet,

$$\begin{aligned} S(h+g)^2 &= 2^{2m} - 2 \sum_{u \in \mathbb{F}_2^m} \mathrm{wt}\,(\mathrm{der}_u h + \mathrm{der}_u g) \\ &\leq 2^{2m} - 2 \sum_{u \in \mathbb{F}_2^m} D(u) \end{aligned}$$

where D_u is the distance of $\mathrm{der}_u h$ to the code $\mathrm{der}_u \mathrm{RM}(3, 8)$.

This fact can be used to show that the distance from

$$Q{=}2345{+}1246{+}1356{+}2467{+}3467{+}2567{+}1348{+}1258{+}1358{+}2478{+}3578{+}1678$$

to the $\mathrm{RM}(3,8)$ satisfies

$$44 < 50 \leq \mathrm{nl}_{3,8}(Q) \leq 60$$

Improving the pevious bound.

Number of bent functions

Let $h \in \mathrm{RM}^*(4,8)$.
Let $\mathrm{NB}(h)$ be the number of bent functions of the form

$$h + g, \quad g \in \mathrm{RM}(3,8)/\mathrm{RM}(1,8). \tag{6}$$

In a previous work, we use the bound

$$\mathrm{NB}(h) \leq \sharp\mathrm{orb}(h) \times 2^{84-\mathrm{var}(h)-R_2(h)}$$

to obtain the estimation

$$\sharp\{\text{nb bent functions}\} \leq 2^{120.2} < 2^{143} \quad \text{(CK bound)}$$

For a given h, it is possible to compute all the bent functions of the form

$$h + g, \quad g \in \mathrm{RM}(3,8)/\mathrm{RM}(1,8). \tag{7}$$

in 18 days on a single standard PC.
Typically, for the 21 forms satisfying

$$\mathfrak{Q}(h) = 0, \quad \mathrm{fix}\,(h) = 1, \quad \mathrm{var}\,(h) = 8, \quad R_2(h) = 28$$

the estimation claims

$$\mathrm{NB}\,(h) \leq \sharp\mathrm{orb}\,(h) \times 2^{48}$$

The exact computations show

$$\mathrm{NB}\,(h) \leq \sharp\mathrm{orb}\,(h) \times 2^{8} \times B(h)$$

with $2^{20.2} \leq B(h) \leq 2^{20.3}$

Some bent functions

$$\mathrm{NB}(h) = \sharp\mathrm{orb}(h) \times 2^8 \times B(h)$$

1204440	1207840	1208048	1214024	1227312
1269888	1270048	1271008	1271784	1272504
1273808	1277872	1280960	1281032	1281568
1281816	1282136	1282792	1284328	1293024
1296616				

Table: The values of $B(h)$.

Note this correspond to 21×2^{92} non affine equivalent bent functions not intersecting the Maiorana-MacFarland class...