# 自己双対符号の Mass Formula の一般化

宗政 昭弘
九大・数理

## Definitions

- $\mathbf{F}_q$: finite field of $q$ elements.
- A (linear) code of length $n$ is a linear subspace of $\mathbf{F}_q^n$.
- The equivalence of codes is defined by permutation of coordinates.
- inner product: $\mathbf{F}_q^n \ni u, v$,

$$u \cdot v = \sum_{j=1}^{n} u_j v_j$$

- $\mathbf{F}_q^n \supset C$: a code of length $n$,

$$C^{\perp} = \{u \in \mathbf{F}_q^n | u \cdot v = 0 \ \forall v \in C\}.$$

- $C$ is self-dual if $C = C^{\perp}$.

## Weights

- $\mathbf{F}_q^n \ni u$,

$$\mathrm{wt}(u) = \#\{j | 1 \leq j \leq n,\ u_j \neq 0\}.$$

- minimum weight of $C$

$$\min(C) = \min\{\mathrm{wt}(u) | u \in C,\ u \neq 0\}.$$

**Fundamental Problem in Coding Theory.** Given positive integers $n, k$, find a code $C \subset \mathbf{F}_q^n$ with $\dim C = k$ such that $\min(C)$ is as large as possible.

In this talk, however, we consider global properties of the **set** of all self-dual codes.

Note: the set of all codes of length $n$ and dimension $k$ is a Grassmann space.

Assume $q = 2$ (binary codes).

- $C$ is doubly-even if for $\forall u \in C$,

$$\mathrm{wt}(u) \equiv 0 \pmod 4.$$

- A doubly-even self-dual (d.e.s.d.) code exists iff $n \equiv 0 \pmod 8$.

Note: if $C$ is a self-dual binary code, then

$$\mathrm{wt}(u) \equiv 0 \pmod 2$$

for $\forall u \in C$. Requiring the divisibility by 4 is the only meaningful divisibility condition (Gleason–Pierce).

## Extremal Codes

- If $C$ is doubly-even self-dual (d.e.s.d.) code of length $n$, then

$$\min(C) \leq 4\left[\frac{n}{24}\right] + 4.$$

A code achieving this bound is called **extremal**.

| $n$ | bound | example | number |
|---|---|---|---|
| 8 | 4 | Hamming | 1 |
| 16 | 4 | Hamming$^2$ | 2 |
| 24 | 8 | Golay | 1 |
| 32 | 8 | Reed–Muller | 5 |
| 40 | 8 | | $\geq 11395$ |
| 48 | 12 | QR | $\geq 1$ |
| 56 | 12 | | $\geq 166$ |
| 64 | 12 | | $\geq 3270$ |
| 72 | 16 | ? | ? |
| $\geq 3952$ | | $\not\exists$ | 0 |

Again

- $q$: an arbitrary prime power,
- $\mathbf{F}_q$ a finite field of $q$ elements.

In characteristic$\neq$ 2,

$$f(u) = \frac{1}{2}(u_1^2 + u_2^2 + \cdots + u_n^2)$$

is a quadratic form with associated bilinear form

$$f(u+v) - f(u) - f(v) = u \cdot v.$$

In characteristic 2, there is no quadratic form defined on $\mathbf{F}_q^n$ whose associated bilinear form is $u \cdot v$.

Over $\mathbf{F}_2$, one needs to restrict to $\langle \mathbf{1} \rangle^\perp$, where

$$\mathbf{1} = (1, 1, \ldots, 1) \in \mathbf{F}_2^n.$$

Note

$$\begin{aligned}\langle \mathbf{1} \rangle^{\perp} &= \{u \in \mathbf{F}_2^n | u \cdot \mathbf{1} = 0\} \\ &= \{u \in \mathbf{F}_2^n | u \cdot u = 0\} \\ &= \{u \in \mathbf{F}_2^n | \, \mathrm{wt}(u) = 0 \pmod 2\}\end{aligned}$$

Thus, for every self-dual code $C$,

$$C \subset \langle \mathbf{1} \rangle^{\perp}$$

hence

$$C = C^{\perp} \supset \langle \mathbf{1} \rangle$$

Define $f : \langle \mathbf{1} \rangle^{\perp} \to \mathbf{F}_2$ by

$$f(u) = \frac{\mathrm{wt}(u)}{2} \pmod 2$$

Then $f$ is a quadratic form on $\langle \mathbf{1} \rangle^{\perp}$ with associated bilinear form $u \cdot v$.

Since $f(u) = 0$ iff $\mathrm{wt}(u) \equiv 0 \pmod 4$, $C$ is doubly-even iff $f(u) = 0$ for all $u \in C$.

If $n \equiv 0 \pmod 4$, then $f$ induces a nondegenerate quadratic form $\overline{f}$ on $\langle \mathbf{1} \rangle^{\perp} / \langle \mathbf{1} \rangle$.

The orthogonal group

$$O(\overline{f}) \subset GL(\langle \mathbf{1} \rangle^{\perp}/\langle \mathbf{1} \rangle) = GL(n-2, 2)$$

acts transitively on doubly-even codes of any given dimension containing $\mathbf{1}$ (Witt's extension Theorem).

If $n \equiv 0 \pmod 8$, then the orthogonal group $O(\overline{f})$ acts transitively on the set of d.e.s.d. codes of length $n$.

The symmetric group on $n$ letters, permuting coordinates, is a subgroup of $O(\overline{f})$.

Study the **set** (homogeneous space) of d.e.s.d. codes.

Define the graph $\Gamma$:

vertices = d.e.s.d. codes of length $n$.

edges = $(C, C')$ with $\dim(C \cap C') = n/2 - 1$.

Then $\Gamma$ has diameter $n/2 - 1$, and $O(\overline{f})$ acts distance-transitively on $\Gamma$.

Define the graph $\Delta$:

vertices = equivalence classes of
d.e.s.d. codes of length $n$.

edges = $([C], [C'])$ with $\dim(C \cap C') = n/2 - 1$.

What is the diameter of $\Delta$?

| $n$ | #vertices | diameter |
|---|---|---|
| 8 | 1 | 0 |
| 16 | 2 | 1 |
| 24 | 9 | 4 |
| 32 | 85 | ? |
| 40 | $> 17492$ | ? |

## Mass Formula (MacWilliams–Sloane–Thompson 1973)

$$\#\text{doubly-even self-dual codes of length } n$$
$$= \prod_{j=0}^{n/2-2} (2^j + 1)$$
$$= \sum_{C\text{: up to equivalence}} \frac{n!}{|\mathrm{Aut}(C)|}.$$

This gives a method to verify that a classification of d.e.s.d. codes of length $n$ is complete, for small $n$.

## Generalization

(1) Mass formula for weight enumerators.

(2) Mass formula for Type II codes over $\mathbf{F}_{2^r}$.

## The weight enumerator

$$W_C(x,y) = \sum_{u \in C} x^{n-\mathrm{wt}(u)} y^{\mathrm{wt}(u)}$$
$$= x^n + \cdots$$

Its mass formula

$$\sum_{\substack{C\text{: d.e.s.d.} \\ \text{length } n}} W_C(x,y)$$
$$= \prod_{j=0}^{n/2-3} (2^j + 1)$$
$$\times (2^{n/2-2}(x^n + y^n) + \sum_{4|k} \binom{n}{k} x^{n-k} y^k).$$

This formula can be used to show the existence of extremal code of length 40.

$$\sum_{\substack{C\text{: d.e.s.d.} \\ \text{length } n}} W_C(1,y)$$

$$= \prod_{j=0}^{n/2-3} (2^j+1)(2^{n/2-2} + \binom{n}{4} y^4 + \cdots)$$

If $n = 40$, then

$$2^{n/2-2} > \binom{n}{4}$$

Each $W_C(1,y)$ has constant term 1.

Each $W_C(1,y)$ has integral coefficients.

Thus $\exists C$ such that $W_C(1,y)$ has no term $y^4$. So $\min(C) \geq 8$. Such a code $C$ is extremal, since

$$4\left\lfloor \frac{n}{24} \right\rfloor + 4 = 8$$

## The Biweight enumerator

$$\mathrm{Biwt}_C(x_{00}, x_{01}, x_{10}, x_{11})$$
$$= \sum_{u,v \in C} x_{00}^{\mathrm{wt}_{00}(u,v)} x_{01}^{\mathrm{wt}_{01}(u,v)} x_{10}^{\mathrm{wt}_{10}(u,v)} x_{11}^{\mathrm{wt}_{11}(u,v)}$$

$$\begin{aligned} u = &\ (0 \ldots 0 \quad 0 \ldots 0 \quad 1 \ldots 1 \quad 1 \ldots 1) \\ v = &\ (\underbrace{0 \ldots 0}_{\mathrm{wt}_{00}} \quad \underbrace{1 \ldots 1}_{\mathrm{wt}_{01}} \quad \underbrace{0 \ldots 0}_{\mathrm{wt}_{10}} \quad \underbrace{1 \ldots 1}_{\mathrm{wt}_{11}}) \end{aligned}$$

More generally, one can define the weight enumerator of degree $g$

$$W_{g,C}(x_a; a \in \mathbf{F}_2^g) = \sum_{\mathbf{u} \in C^g} \prod_{a \in \mathbf{F}_2^g} x_a^{\mathrm{wt}_a(\mathbf{u})}.$$

$$W_C = W_{1,C}, \qquad \mathrm{Biwt}_C = W_{2,C}$$

The mass formula

$$\sum_{\substack{C: \text{ d.e.s.d.} \\ \text{length } n}} W_{g,C}(x_a; a \in \mathbf{F}_2^g)$$

can be computed in principle (Runge, 1996).

We have found a nice formula for the case $g = 2$ in terms of the root system $E_8$.

**Theorem (Ozeki–M.).**

$$\sum_{\substack{C:\ \text{d.e.s.d.}\\ \text{length } n}} \mathrm{Biwt}_C(\mathbf{x}) = 2^{n/2-7} \prod_{j=0}^{n/2-4} (2^j+1) \times \sum_{\alpha \in E_8^{\mathbf{C}}} \langle \alpha, \mathbf{x} \rangle^n$$

where

$$\langle \alpha, \mathbf{x} \rangle = \alpha_{00}x_{00} + \alpha_{01}x_{01} + \alpha_{10}x_{10} + \alpha_{11}x_{11}$$

$$\alpha = (\alpha_{00}, \alpha_{01}, \alpha_{10}, \alpha_{11}) \in E_8^{\mathbf{C}} \subset \mathbf{C}^4$$

$E_8^{\mathbf{C}}$ = (unique) embedding of the root system $E_8$ into $\mathbf{C}^4$ invariant under the multiplication by $\sqrt{-1}$

The root system $E_8$ consists of the 240 vectors

$$\frac{1}{\sqrt{2}}(\pm 2,0,0,0,0,0,0,0),\ldots,\frac{1}{\sqrt{2}}(0,0,0,0,0,0,0,\pm 2),$$

and

$$\begin{array}{cc}
\frac{1}{\sqrt{2}}(*,*,*,*,0,0,0,0), & \frac{1}{\sqrt{2}}(0,0,0,0,*,*,*,*),\\
\frac{1}{\sqrt{2}}(*,*,0,0,*,*,0,0), & \frac{1}{\sqrt{2}}(0,0,*,*,0,0,*,*),\\
\frac{1}{\sqrt{2}}(0,0,*,*,*,*,0,0), & \frac{1}{\sqrt{2}}(*,*,0,0,0,0,*,*),\\
\frac{1}{\sqrt{2}}(*,0,*,0,*,0,*,0), & \frac{1}{\sqrt{2}}(0,*,0,*,0,*,0,*),\\
\frac{1}{\sqrt{2}}(0,*,0,*,*,0,*,0), & \frac{1}{\sqrt{2}}(*,0,*,0,0,*,0,*),\\
\frac{1}{\sqrt{2}}(*,0,0,*,0,*,*,0), & \frac{1}{\sqrt{2}}(0,*,*,0,*,0,0,*),\\
\frac{1}{\sqrt{2}}(0,*,*,0,0,*,*,0), & \frac{1}{\sqrt{2}}(*,0,0,*,*,0,0,*),
\end{array}$$

where $*$ means $\pm 1$. The latter $14\times 16$ vectors, if the signs are discarded, give 14 hyperplanes of the affine geometry $AG(3,2)=\mathbf{F}_2^3$.

Since

$$\mathrm{Biwt}_C(x,0,0,y) = W_C(x,y)$$

we obtain

**Corollary.**

$$\sum_{\substack{C:\ \text{d.e.s.d.}\\ \text{length } n}} W_C(x,y) = \frac{1}{16} \prod_{j=0}^{n/2-3} (2^j+1) \sum_{\alpha \in D_4^{\mathbf{C}}} \langle \alpha, \mathbf{x} \rangle^n$$

where

$$\begin{aligned} \langle \alpha, \mathbf{x} \rangle &= \alpha_0 x + \alpha_1 y \\ \alpha &= (\alpha_0, \alpha_1) \in D_4^{\mathbf{C}} \subset \mathbf{C}^2 \\ D_4^{\mathbf{C}} &= \text{(unique) embedding of the root} \\ &\quad \text{system } D_4 \text{ into } \mathbf{C}^2 \text{ invariant} \\ &\quad \text{under the multiplication by } \sqrt{-1} \end{aligned}$$

The root system $D_4$ consists of 24 vectors

$$\pm\sqrt{2}e_1, \pm\sqrt{2}e_2, \pm\sqrt{2}e_3, \pm\sqrt{2}e_4,$$
$$\frac{1}{\sqrt{2}}(\pm 1, \pm 1, \pm 1, \pm 1).$$

## Modular Forms

- Broué–Enguehard (1972): If $C$ is a d.e.s.d. code of length $n$, then

$$W_C(\theta_3(2\tau), \theta_2(2\tau))$$

is a modular form of weight $n/2$ on $SL(2, \mathbf{Z})$, where

$$\theta_3(\tau) = \sum_{m \in \mathbf{Z}} q^{m^2}$$
$$\theta_2(\tau) = \sum_{m \in \mathbf{Z}} q^{(m+1/2)^2}$$

where $q = e^{\pi i \tau}$.

Indeed, this is the theta series of the even integral unimodular lattice constructed from the code $C$.

- More generally, Hermann (1991): If $C$ is a d.e.s.d. code of length $n$, then

$$W_{g,C}(f_a; a \in \mathbf{F}_2^g)$$

is a Siegel modular form of weight $n/2$ on $Sp(2g, \mathbf{Z})$ where $f_a$'s are theta constants defined by

$$f_a(\tau) = \sum_{x \in \mathbf{Z}^g} \exp 2\pi i(\tau[x + \frac{1}{2}a]),$$

and

$$\begin{aligned} \tau \in & \text{ Siegel upper half space} \\ & = \{\tau \in M_g(\mathbf{C}) \mid \tau^T = \tau,\ \mathrm{Im}\,\tau > 0\} \end{aligned}$$

Note that the mass formula

$$\sum_{\substack{C:\ \text{d.e.s.d.} \\ \text{length } n}} W_{g,C}(f_a; a \in \mathbf{F}_2^g) \qquad (\#)$$

also gives a Siegel modular form of weight $n/2$ on $Sp(2g, \mathbf{Z})$.

It is easier to compute (#) than to compute $W_{g,C}(f_a; a \in \mathbf{F}_2^g)$ for an individual code $C$!.

Not every modular form on $SL(2,\mathbf{Z})$ is a linear combination of

$$W_C(\theta_3(2\tau),\theta_2(2\tau))$$

The Eisenstein series of weight 6

$$E_6 = x^{12} - 33x^8y^4 - 33x^4y^8 + y^{12}$$

with

$$x = \theta_3(2\tau), y = \theta_2(2\tau),$$

can not be obtained from the weight enumerators since there is no d.e.s.d. code of length $n = 12$.

If $8|n$

$$\frac{1}{\prod_{j=0}^{n/2-3}(2^j+1)} \sum_{\substack{C\text{: d.e.s.d.} \\ \text{length } n}} W_C(x,y)$$

$$= 2^{n/2-2}(x^n+y^n) + \sum_{4|k} \binom{n}{k} x^{n-k}y^k \qquad \text{(Mac–S-T)}$$

$$= \frac{1}{16} \sum_{\alpha \in D_4^{\mathbf{C}}} \langle \alpha, \mathbf{x} \rangle^n \qquad \text{(Mu-O)}$$

If $n = 12$, then (Mu-O) reduces to

$$x^{12} - 33x^8y^4 - 33x^4y^8 + y^{12}$$

$$\overset{\text{Br–Eng}}{\longrightarrow} \quad \text{Eisenstein series of weight 6}$$

The subgroup of $GL(2,\mathbf{C})$ leaving $D_4^{\mathbf{C}}$ invariant is a unitary reflection group of order 96, and its ring of invariants is isomorphic to the ring of modular forms on $SL(2,\mathbf{Z})$ via the Broué–Enguehard map

$$x \mapsto \theta_3(2\tau), \qquad y \mapsto \theta_2(2\tau).$$

For a TOB $B = \{\alpha_1, \ldots, \alpha_r\}$ of $\mathbf{F}_{2^r}$ over $\mathbf{F}_2$, and $u \in \langle \mathbf{1} \rangle^\perp \subset \mathbf{F}_{2^r}^n$, define

$$f_B(u) = \sum_{j=1}^{r} \frac{\mathrm{wt}(\alpha_j u)}{2} \alpha_j^2 \in \mathbf{F}_{2^r}.$$

Then $f_B$ is a quadratic form on $\langle \mathbf{1} \rangle^\perp$, and

$$f_B(u+v) - f_B(u) - f_B(v) = u \cdot v$$

for all $u, v \in \langle \mathbf{1} \rangle^\perp$.

**Definition.** A self-dual code $C$ of length $n$ over $\mathbf{F}_{2^r}$ is called a Type II code with respect to a TOB $B$ if

$$f_B(u) = 0 \quad \text{for all } u \in C,$$

or equivalently, $\phi_B(C)$ is doubly-even.

Although the definition of Type II codes depends on the choice of a TOB, the classification of Type II codes over $\mathbf{F}_{2^r}$ obtained so far turned out to be independent of the choice of a TOB.

Recently Betsumiya proved:

**Theorem.** The definition of Type II codes is independent of the choice of a TOB.

In other words, if $C$ is a Type II code over $\mathbf{F}_{2^r}$ with respect to a TOB $B$, then $C$ is a Type II code over $\mathbf{F}_{2^r}$ with respect to any TOB.

One could consider self-dual codes over $\mathbf{F}_{2^r}$ with the property that $\phi_B(C)$ is doubly-even for (not necessarily trace-orthogonal) a basis $B$ of $\mathbf{F}_{2^r}$ over $\mathbf{F}_2$.

We conjecture that such a code is automatically Type II with respect to any TOB.

**Further Generalization.**

Combining the two generalizations:

(1) Mass formula for weight enumerators

(2) Mass formula for Type II codes over $\mathbf{F}_{2^r}$

we are lead to:

(3) Mass formula for the weight enumerators of Type II codes over $\mathbf{F}_{2^r}$

More precisely, find

$$\sum_{\substack{C\text{: Type II code} \\ \text{of length } n \\ \text{over } \mathbf{F}_{2^r}}} W_{\phi_B(C)}(x,y)$$

This becomes a modular form of weight $rn/2$ after the substitution

$$x \mapsto \theta_3(2\tau), \qquad y \mapsto \theta_2(2\tau).$$

In general,

$$\sum_{\substack{C\text{: Type II code} \\ \text{of length } n \\ \text{over } \mathbf{F}_{2^r}}} W_{\phi_B(C)}(x,y)$$

and

$$\sum_{\substack{C\text{: d.e.s.d. code} \\ \text{of length } rn}} W_C(x,y)$$

are linearly independent.

More generally, find

$$\sum_{\substack{C\text{: Type II code} \\ \text{of length } n \\ \text{over } \mathbf{F}_{2^r}}} W_{g,\phi_B(C)}(x_a; a \in \mathbf{F}_2^g).$$

This becomes a Siegel modular form of weight $rn/2$ after the substitution

$$x_a \mapsto f_a.$$