Hadamard Matrices and Reed-Muller Codes

Hadamard Matrices. In the 19th century, Hadamard considered the sizes of the determinants of $n \times n$ matrices $A$ with all entries in $[-1, 1]$. Since the norm of each row is at most $\sqrt{n}$ and the absolute value of the determinant is a measure of the volume of the box formed by its row vectors in $\mathbb{R}^n$, it is natural to conclude the determinant is at most $n^{n/2}$ and the row vectors should be orthogonal. For example, let row 1 be 1 1 and row 2 to be 1 −1, then the area of the square formed by these two vectors is 2. Matrices that have +1 or −1 as entries with orthogonal rows and orthogonal columns are important in various applications.

Definitions. Let $I_n$ be the $n \times n$ identity matrix. A $n \times n$ matrix $H$ is a Hadamard matrix (of order $n$) if and only if its entries are ±1 and it satisfies $HH^T = nI$. Two Hadamard matrices are equivalent if and only if one of them can be obtained by the other after permuting rows or columns or multiplying rows or columns by $-1$. A Hadamard matrix is normalized if and only if all entries of its first row and first column are +1. (Clearly, every Hadamard matrix is equivalent to a normalized one.) Often the entries of a Hadamard matrix are written as + or −, which correspond to 1 or −1 respectively.

Example. $(+)\begin{pmatrix} + & + & + & + \\ + & + & - & - \\ + & - & + & - \\ + & - & - & + \end{pmatrix}$ are normalized Hadamard matrices of orders 1, 2, 4 respectively.

Theorem. If $H$ is a Hadamard matrix of order $n$, then $n = 1, 2$ or $n \equiv 0 \pmod{4}$.

Proof. The cases $n < 4$ are easy to check. For $n \geq 4$, first normalize $H$. Since the top 2 rows are orthogonal, row 2 contains $n/2$ +'s and $n/2$ −'s. By permuting columns, we may assume the +'s in row 2 are in the first $n/2$ entries and the −'s are in the last $n/2$ entries. For row 3, let there be $a$ +’s under those columns with +, + as top 2 entries, $b$ −’s under those columns with +, + as top 2 entries, $c$ +’s under those columns with +, − as top 2 entries, $d$ −’s under those columns with +, − as top 2 entries.

\[
\begin{array}{cccc}
+ & + & + & + \\
+ & + & + & + \\
+ & + & + & + \\
+ & + & + & + \\
a columns & b columns & c columns & d columns \\
\end{array}
\]

Then $a + b = n/2$ and $c + d = n/2$. Taking inner product of row 1 and row 3, we get $a - b + c - d = 0$. Taking inner product of row 2 and row 3, we get $a - b - c + d = 0$. Solving the 4 equations of $a, b, c, d$, we get $n = 4a = 4b = 4c = 4d$. 

To produce Hadamard matrices of large orders, we introduce some auxiliary concepts.
**Definition.** Let $A$ be a $m \times n$ matrix with entries $a_{ij}$ and $B$ be another matrix. The **Kronecker product** (or **tensor product**) of $A$ and $B$ (denoted by $A \otimes B$) is the matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

**Example.** For $A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 4 \\ 0 & -1 \end{pmatrix}$, $A \otimes B = \begin{pmatrix} 1B & 0B \\ 2B & 3B \end{pmatrix} = \begin{pmatrix} 2 & 4 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 4 & 8 & 6 & 12 \\ 0 & -2 & 0 & -3 \end{pmatrix}$.

**Lemma.** (1) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$. There exist $A, B$ with $A \otimes B \neq B \otimes A$.

(2) $(A \otimes B)^T = AT \otimes BT$. If $A, B, C, D$ are $m \times n, q \times r, n \times p, r \times s$ matrices respectively, then $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$. (These will be left as exercises.)

**Notation.** Let $H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and let $H_{2^m}$ denote the tensor product of $m$ $H_2$ matrices.

Sylvester introduced the following useful construction formula, namely if $K$ is a Hadamard matrix, then the matrix $H_2 \otimes K = \begin{pmatrix} K & K \\ K & -K \end{pmatrix}$ is also a Hadamard matrix. The next theorem is a generalization of this fact.

**Theorem.** If $H$ and $K$ are Hadamard matrices of orders $m$ and $n$ respectively, then $H \otimes K$ is a Hadamard matrix of order $mn$.

**Proof.** By part (2) of the lemma and the fact $I_m \otimes I_n = I_{mn}$, we get $(H \otimes K)(H \otimes K)^T = (H \otimes K)(HT \otimes K^T) = (HH^T) \otimes (KK^T) = mI_m \otimes nI_n = mnI_{mn}$.

**The Fast Hadamard Transform Theorem.** For $i = 1, 2, \ldots, m$, let

$$M_{2^m}^{(i)} = I_{2^{m-i}} \otimes H_2 \otimes I_{2^{i-1}}.$$

Then $H_{2^m} = M_{2^m}^{(1)}M_{2^m}^{(2)} \cdots M_{2^m}^{(m)}$ is a Hadamard matrix of order $2^m$.

**Proof.** Induct on $m$. Case $m = 1$ is clear. For $1 \leq i \leq m$, since $I_{rs} = I_r \otimes I_s$, we see

$$M_{2^m+i}^{(i)} = I_{2^{m+i-1}} \otimes H_2 \otimes I_{2^{i-1}} = I_2 \otimes I_{2^m-i} \otimes H_2 \otimes I_{2^{i-1}} = I_2 \otimes M_{2^m}^{(i)}$$

and

$$M_{2^m+i}^{(m+1)} = H_2 \otimes I_{2^m}.$$

Using the formula $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$, we have

$$M_{2^m+1}^{(1)}M_{2^m+1}^{(2)} \cdots M_{2^m+1}^{(m+1)} = (I_2 \otimes M_{2^m}^{(1)})(I_2 \otimes M_{2^m}^{(2)}) \cdots (I_2 \otimes M_{2^m}^{(m)})(H_2 \otimes I_{2^m}) = H_2 \otimes (M_{2^m}^{(1)}M_{2^m}^{(2)} \cdots M_{2^m}^{(m)})I_{2^m} = H_2 \otimes H_{2^m} = H_{2^{m+1}}.$$

2
So far we can write down $H_{2^m}$. We may ask if Hadamard matrices of order $2^kj$ with \( j > 1 \) odd and \( k \geq 2 \) exist. Indeed, we have the fact that Hadamard matrices of order 12 exist and they are all equivalent to the following matrix. (See W. A. Coppel's book, *Number Theory: An Introduction to Mathematics*, 2ed., pp.252-254.)

For this kind of Hadamard matrices, we have the following definition and theorem.

**Definition.** A \( n \times n \) matrix \( C \) is a *conference matrix of order* \( n \) if and only if the entries on its main diagonal are 0’s and the rest of the entries are \( \pm 1 \) such that \( CC^T = (n - 1)I \).

**Theorem.** (1) If \( C \) is a symmetric (i.e. \( C^T = C \)) conference matrix of order \( n \), then \( H = \left( \begin{array}{cc} I + C & -I + C \\ -I + C & -I - C \end{array} \right) \) is a Hadamard matrix of order \( 2n \). (So \( 4 \mid 2n \), i.e. \( n \) is even.)

(2) If \( C \) is an antisymmetric (i.e. \( C^T = -C \)) conference matrix, then \( H = I + C \) is a Hadamard matrix. (The orders of \( H \) and \( C \) are divisible by 4.)

**Proof.** Just multiply \( H \) with \( H^T \) in (1) and (2). Use \( C^T = -C \) in (1) and \( C^T = C \) and \((\pm I \pm C)^T = \pm I \pm C^T\) in (2).

Next we will look at a way of producing conference matrices of large orders. Let \( q = p^n \), where \( p \) is a prime and \( n \in \mathbb{N} = \{1, 2, 3, \ldots \} \). A *field* is a set, like \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \), containing 0 and 1 such that we can define the 4 operations, namely addition, subtraction, multiplication and division (by nonzero denominators) with usual properties. While \( \mathbb{Q}, \mathbb{R}, \mathbb{C} \) are fields with infinitely many elements, we would like to point out there are also finite fields. For example, \( \mathbb{F}_2 = \{0, 1\} \) with usual properties of the 4 operations except \( 1 + 1 = 0 \).

In algebra, it is known that for \( q \) of the form \( p^n \) as above, there exists a finite field \( \mathbb{F}_q \) with \( q \) elements. Also, in \( \mathbb{F}_q, |\{x^2 : x \in \mathbb{F}_q \setminus \{0\}\}| = |\{y : y \neq x^2, x \in \mathbb{F}_q\}| \), i.e. the number of nonzero squares equals the number of nonsquares. Define \( \chi : \mathbb{F}_q \rightarrow \{0, 1, -1\} \) by

\[
\chi(a) = \begin{cases}
0 & \text{if } a = 0, \\
1 & \text{if } a \text{ is a nonzero square in } \mathbb{F}_q, \\
-1 & \text{if } a \text{ is a nonsquare in } \mathbb{F}_q
\end{cases}
\]

It can be used to define a useful \( q \times q \) matrix \( Q \) as follows. Let the elements of \( \mathbb{F}_q \) be \( a_0, a_1, \ldots, a_{q-1} \) with \( a_0 = 0 \). Define the \( ij \)-entry of \( Q \) to be \( Q_{ij} = \chi(a_i - a_j) \), where
0 \leq i, j < q. Then Q satisfies $QQ^T = qI - J, QJ = JQ = O$, where $J$ is the $q \times q$ matrix with 1 in all entries. In 1933, Paley observed that the $(q+1) \times (q+1)$ matrix

$$C = \begin{pmatrix}
0 & 1 & \cdots & 1 \\
\pm1 & \pm1 & & \\
\vdots & & & Q \\
\pm1 & & & \\
\end{pmatrix}$$

is a conference matrix of order $q+1$ where the $\pm$ signs are chosen in such a way that $C$ is symmetric if $q \equiv 1 \pmod{4}$ or antisymmetric if $q \equiv 3 \pmod{4}$. These Paley matrices $C$ produce many Hadamard matrices of large orders.

**Paley’s Theorem (1933).** If $q = p^n$ for some odd prime $p$ and $n \in \mathbb{N}$, then a Hadamard matrix of order $q+1$ exists if $q \equiv 3 \pmod{4}$ and a Hadamard matrix of order $2(q+1)$ exists if $q \equiv 1 \pmod{4}$.

In the figure, $+$ means 1 and $-$ means $-1$. The Hadamard matrices of order 12 shown are constructed from the Paley matrices of order $11 + 1$ and $5 + 1$ (using parts (2) and (1) of the theorem before Paley’s theorem respectively).

**Reed-Muller Codes.** With the existence of large order of Hadamard matrices, they provided important applications in error correction of signals. In 1954, D. E. Muller and I. S. Reed introduced a code, which became famous in 1972 when it was used in transmitting pictures of Mars and Saturn taken from US spacecrafts. The pictures were divided into a $600 \times 600$ grid of pixels. Each pixel captured the shades of gray in a scale of 0 to $63 = 2^6 - 1$. So in binary, it is 6 bits of (0,1)-signals. For a picture, this took $6 \times 600^2 = 2,160,000$ bits and additional bits were introduced to detect and correct bit errors in transmission due to noisy channels.

To understand the error correction method by Reed-Muller, we will define some terms.

**Definitions.** (1) A $m$-ary word of length $n$ is sequence of $n$ symbols, where each symbol is an element in a set $S = \{s_1, s_2, \ldots, s_m\}$ called the alphabet. The set of all $m$-ary words of length $n$ is denoted by $S^n$ (or $H(n,m)$ called the Hamming space). Typically, we will take $S = \mathbb{F}_q$ for some $q$. 


(2) A **code** with \(M\) **codewords of length** \(n\) is a subset of \(S^n = \mathbb{F}_q^n\) with \(M\) elements. Typically, we consider binary (i.e. 2-ary) words and take \(q = 2\) so that the alphabet is \(\mathbb{F}_2 = \{0, 1\}\) and a codeword of length \(n\) is consisted of \(n\) 0 or 1 symbols that is in the code.

(3) The **Hamming metric** is the function \(d : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \{0, 1, 2, 3, \ldots\}\) defined by

\[
d(a_1 a_2 \ldots a_n, b_1 b_2 \ldots b_n) = |\{i : a_i \neq b_i, i = 1, 2, \ldots, n\}|.
\]

For all \(x, y, z \in \mathbb{F}_q^n\), the Hamming metric satisfies the property that (1) \(d(x, y) \geq 0\) with equality if and only if \(x = y\); (2) \(d(x, y) = d(y, x)\) and (3) \(d(x, z) \leq d(x, y) + d(y, z)\). Next, we define \(d(C) = \min\{d(x, y) : x \neq y \text{ for } x, y \in C\}\).

(4) A \((n, M, d)\)-**code** is a code with \(M\) codewords, each is of length \(n\) and \(d\) is the minimum distance between two distinct codewords. A code \(C\) in \(\mathbb{F}_q^n\) is **linear** if and only if \(x, y \in C\) implies \(x + y \in C\). Also, for codes in \(\mathbb{F}_q^n\), the **weight** of a word \(a_1 a_2 \ldots a_n\) is defined to be \(w(a_1 a_2 \ldots a_n) = |\{i : a_i \neq 0, i = 1, 2, \ldots, n\}|\) so that \(d(x, y) = w(x - y)\).

**Example.** Let \(n = 8\) and \(S = \mathbb{F}_2 = \{0, 1\}\). Then \(\mathbb{F}_2^8\) has \(2^8 = 256\) words and let \(C = \{\text{00000000}, \text{00001111}, \text{11110000}, \text{11111111}\}\) be the code with 4 codewords. The minimum distance \(d(C)\) between two distinct codewords is 4. The sum of two codewords is a codeword. So \(C\) is a binary linear \((8, 4, 4)\)-code.

Now \(11100000\) is a word in \(\mathbb{F}_2^8\), but it is not a codeword in the code \(C\). The minimum distance from \(11100000\) to a codeword in \(C\) is \(d(11100000, 11110000) = 1\). We say there is a one bit error in 11100000. In error correction schemes, 11100000 will be replaced by the codeword 11110000 as it is the closest codeword to 11100000.

**Theorem.** Let \(C\) be a code. For every word \(y \notin C\), let there be a \(x \in C\) with \(d(x, y) \leq t\).

1. If \(d(C) \geq t + 1\), then \(C\) can detect up to \(t\) errors.
2. If \(d(C) \geq 2t + 1\), then the code \(C\) can correct up to \(t\) errors in any codeword.

**Proof.** (1) If \(d(C) \geq t + 1\), then for all \(z \in C\) with \(z \neq x\), we must have \(d(z, y) \geq 1\) for otherwise \(d(x, z) \leq d(x, y) + d(y, z) < t + 1\), contradicting \(d(C) \geq t + 1\). So \(y\) contains at least 1 and at most \(t\) errors from every codeword.

(2) If \(d(C) \geq 2t + 1\), then for all \(z \in C\) with \(z \neq x\), we must have \(d(z, y) \geq t + 1\) for otherwise \(d(x, z) \leq d(x, y) + d(y, z) < 2t + 1\), contradicting \(d(C) \geq 2t + 1\). Therefore, \(x\) is the only codeword that can allow \(y\) to have at most \(t\) errors.

**Definition.** For \(m = 1, 2, 3, \ldots\), the **Reed-Muller code** \(R(1, m)\) is the span of the rows of the \((m + 1) \times 2^m\) generating matrix \(G\), where column \(j\) is \(2^m + j - 1\) in base 2 for \(j = 1, 2, \ldots, 2^m\). Below let \(1_n\) be the row vector with all \(n\) coordinates equal 1.

**Example.** \(R(1, 3)\) has generating matrix \(G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & \vdots \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & \vdots \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & \vdots \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & \vdots \end{pmatrix}^T\). Row 1 is the vector \(1_8\), row 2 is the vector \(v_3\), row 3 is the vector \(v_2\) and row 4 is the vector \(v_1\).
Remarks. $R(1, m)$ is a $(2^m, 2m+1, 2^{m-1})$ code since there are $M = 2^{m+1}$ codewords (consist of the sums of every $k$ rows for $k = 0, 1, \ldots, m + 1$), each has length $n = 2^m$ and minimum distance $d = 2^{m-1}$. By part (2) of the last theorem, $R(1, m)$ is capable of correcting $\lfloor (2^{m-1} - 1)/2 \rfloor = 2^{m-2} - 1$ bit errors.

Encoding Scheme. If each pixel is assigned one of the $2^{m+1}$ colors, then write the $j$-th color in base 2 as a row vector $v$, then $vG$ is the codeword corresponding to the color.

Decoding Scheme. If $vG$ was sent and a word $r$ is received (which may or may not be a codeword), then use the fast Hadamard transform to write down the $H_{2^m}$ Hadamard matrix and do the following steps:

Step 1. If $r = (r_1, r_2, \ldots, r_{2^m})$, then let $F = ((-1)^{r_1}, (-1)^{r_2}, \ldots, (-1)^{r_{2^m}})$.

Step 2. Let $x$ be a coordinate of $FH_{2^m}$ with largest absolute value. If $|x| \neq 2^m$, then let $a_m a_{m-1} \cdots a_1$ be $|x|$ in base 2 and go to step 3, otherwise, the codeword is $r$ and stop.

Step 3. If $x > 0$, then the codeword is $a_m v_m + a_{m-1} v_{m-1} + \cdots + a_1 v_1$, otherwise it is $1_{2^m} + a_m v_m + a_{m-1} v_{m-1} + \cdots + a_1 v_1$.

Example. In $R(1, 3)$ coding scheme, if a $vG$ was sent and $r = (10000011)$ is received, then $F = (-1, 1, 1, 1, 1, -1, -1)$ and $FH_8 = (2, -2, 2, -2, -2, -6, -2)$. The maximum absolute value of the coordinates of $FH_8$ is $| -6 | = 6$, which is 110 in base 2. The correct codeword is $1_8 + 1v_3 + 1v_2 + 0v_1 = (11000011)$. So the second bit of $r$ was an error.

Exercises. (1) Compute $H_8 = H_2 \otimes H_2 \otimes H_2$.

(2) Prove that $(A \otimes B) \otimes C = A \otimes (B \otimes C)$. Give an example $A \otimes B \neq B \otimes A$.

(3) Prove that $(A \otimes B)^T = A^T \otimes B^T$. If $A, B, C, D$ are $m \times n, q \times r, n \times p, r \times s$ matrices respectively, then prove that $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$.

(4) Prove that a Hadamard matrix of order $n$ exists, where $n$ is a multiple of 4 and at most 100 (except for 92). (Hint: Use Paley’s Theorem for $n = 12, 20, 28, 36, 44, 52, 60, 68, 76, 84, 100$. The remaining cases can be taken care of by using $H_{mn} = H_m \otimes H_n$.)