通信情報保護學會誌
第1卷, 第3號, 1991. 12

# 完全符號의 存在性에 關한 考察

李 敏 燮*

## 1. 서 론

통신채널을 통하여 전달되어지는 데이타의 신뢰성(Data Reliability)을 증대시키기 위한 노력은 Shannon에 의하여 시작된 오류수정 부호이론(Error-Correcting Coding Theory)의 정립과 응용을 통하여 이루어져 왔다. 이와같은 오류수정부호 이론의 응용은 데이타통신 분야에서 뿐만 아니라 전산기 기억장치(Memory Systems), 화일시스템, 신경망(Neural-Networks) 및 암호체계(Cryptosystems) 분야에서도 이루어지고 있다. 특히 데이타의 보완성을 유지하기 위한 암호체계의 응용은 Goppa 부호를 이용한 공개키(Public Key) 암호체계의 실현이 McEliece에 의하여 이루어진 이후 부터 활발하게 시작되어, 암호체계에서 오류수정부호를 이용한 암호화는 중요한 비중을 차지하게 되었다. 이와같이 오류수정부호를 연구하는 학문을 부호이론이라 한다.

이와 같은 부호이론의 역사는 암호학의 역사와는 달리 오래되지 않은 40餘年 前의 일이었다. 1948年 C. E. Shannon이 'A Mathematical Theory of Communication'에서 情報理論(imformation theory)에 確率概念을 도입하여 정보를 비트(bit)로 표현하여 계산된 정보의 量 및 정보전달 속도가 通信路(channel)의 容量을 넘지 않는 범위에서 誤謬(error)에 구애됨이 없어 정보전달이 가능하다는 부호이론을 제창하면서 시작되었다. 이러한 부호이론은 1960년대에 이르러 符號化回路(encoding circuit) 및 復號回路(decoding circuit)의 개발 연구가 활발히 진행되었고, 1970년대에 실용화되기 시작되었다고 할 수 있다. 이와같이 짧은 연륜을 가진 부호이론은 대수학, 확률 및 통계이론을 배경으로 발전되고 있다.

일반적으로 부호체계에서 수신된 모든 電文벡터(기호)들이 복호되는 것은 아니다. 이런 관점에서 모든 전문벡터를 복호할 수 있는 부호인 완전부호는 많은 장점을 가진 부호이다. 본고에서는 이와같은 완전부호의 존재성에 관하여 고찰한다.

## 2. 부 호

전달하여야 할 내용을 傳文(message)이라 하고 이 전문이 通信路를 통하여 전달될 수 있도록 통신로가 인지할 수 있는 기호로 전환하는 과정을 符號化(encoding)라 하고 전환하는 장치를 符號器(encoder)라 한다. 또 부호화된 언어를 符號語(co-

* 단국대학교 수학과 부교수

deword)라 하고, 하나의 전달 체계에 속하는 모든 부호어들의 집합을 符號(code)라고 한다. 이 전문은 통신로를 통하여 受信者가 있는 곳으로 전달되어, 수신자는 이것을 이해할 수 있는 언어로 바꾸어 전달 받은 내용을 알게 된다. 여기서, 수신된 기호를 受信벡터(received vector)라 하고 이 자료를 수신자가 이해할 수 있는 내용으로 전환하는 과정을 復號(decoding)라 하며, 전환하는 장치를 復號器(decoder)라 한다. 그러므로 정보의 전달과정에 부호화와 복호가 필요하며 전달체계는 다음과 같다.

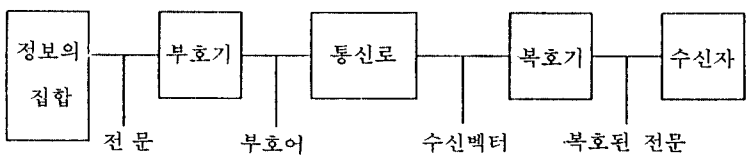



그림 1. 1

일반적으로, 이와같은 전달체계에서 부호화 및 복호는 전자계산기를 통하여 기계적으로 처리되지만 不良한 통신로나 소음이 심한 조건 등에서는 전문이 제대로 전달되지 않을 때가 있다. 이와같이 전문이 잘못 전달되는 경우에 誤謬(error)가 발생하였다고 한다. 부호이론은 수신벡터로 부터 이와 같은 오류를 檢出(detection)하고 修正(correction)하여 정확한 정보를 얻게 함과 동시에 정보의 부호화 및 복호를 쉽게 할 수 있게 하는 부호를 만드는 이론을 연구하는 학문이다.

정보를 부호화하는 방법은 크게 나누어 두 가지, 즉 블럭부호(block code)와 convolution 부호가 있다. 여기에서 언급하는 부호는 수학적인 이론 특히 대수학 이론이 많이 쓰이는 부호로서, 부호기가 k 비트의 전문을 n 비트의 부호어로 만드는 블럭부호이다. 이와같은 부호를 엄밀하게 정의하려면 전달하고자 하는 내용을 부호화 하는데, 쓰이는 q개의 서로 다른 기호들이 필요하며, 이 기호들의 집합을 alphabet이라 하고 편의상 q를 법(modulo)으로 한 집합 $\mathbf{Z}_q=\{0, 1, \cdots, (q-1)\}$로 쓴다. 또한 블럭부호를 간단히 부호라 하며 정의는 다음과 같다.

정의 2. 1 전달할 정보가 $\mathbf{Z}_q$의 n개의 기호들의 블럭으로 부호화될 때 이 부호를 블럭부호(block code)라 한다. 이 부호의 원소인 블럭을 符號語(codeword)라 하고 n을 이 부호의 길이(length)라고 한다. 이 때 블럭부호에서 부호어들의 집합을 q 進符號(q-ary code)라 한다.

위의 정의로 부터 길이 n인 부호 C는 집합

$$\mathbf{Z}_q^n=\{(x_1, x_2, \cdots, x_n) \mid x_i \in \mathbf{Z}_q,\ 1 \leq i < n\}$$

의 부분집합임을 알 수 있다. 이 때에 $\mathbf{Z}_q^n$을 n 次元空間(n dimensional space)이라고 한다. 따라서 부호어는 $\mathbf{X}=(x_1, \cdots, x_n)$으로 나타내며 편의에 따라 $\mathbf{X}=x_1 \cdots x_n$ 또는 $\mathbf{X}=[x_1 \cdots x_n]$으로 쓰기도 한다. 집합 $\mathbf{Z}_q$의 원소의 개수 q는 일반적으로 임의의 양의 정수로 제한이 없으나 q=2인 경우가 흔히 이용된다.

예 2. 2 두 가지 정보 "yes"와 'no"를 두 개의 기호 '0'과 '1'을 사용하여 다음과 같이 3가지 방법으로 부호화하였다고 하자.

(1) 0 : yes (2) 00 : yes (3) 000 : yes
1 : no 11 : no 111 : no

그러면 alphabet은 $\mathbf{Z}_2=\{0, 1\}$이고

$C_1=\{0, 1\}$, $C_2=\{00, 11\}$ 및 $C_3=\{000, 111\}$

은 각각 부호이다. 또한 0과 1은 $C_1$의, 00과 11은 $C_2$의, 000과 111은 $C_3$의 부호어이다. 따라서 $C_1$, $C_2$와 $C_3$는 각각 길이가 1, 2 및 3인 이진 부호이다.

특히, 3가지 부호 모두 부호어가 0의 반복 또는 1의 반복으로 이루어져 있다. 이와 같은 부호를 反復符號(repetition code)라고 한다.

예 2. 2에서 수 0과 1이 통신로를 통하여 전달될 때에 잘못 전달되어 오류가 발생할 확률이 일반적으로 $\frac{1}{2}$ 보다 작은 것으로 가정하면

(1)의 경우에 0을 보냈는데 1을 받으면 잘못된 정보를 받게 되는데, 받는 사람이 오류의 發生 有無를 알지 못하기 때문에 잘못된 정보를 얻게 되므로, 유용한 부호가 되지 못한다.

(2)의 경우에 01 또는 10을 받으면 정보가 잘못 전달되었음을 알지만 00인지 11인지 알 수 없다. 이와 같은 경우에는 오류가 일어났음을 알게 되지만 원래의 정보는 알 수 없다. 즉, 한개의 오류를 검출하지만 이를 수정 할 수는 없다.

(3)의 경우에 001을 받았다면 이것은 확률적으로 111 보다 000이 보내진 부호어일 확률이 크므로 000으로 복호하게 된다. 즉, 이 부호는 오류가 한 개일 때에 오류를 수정하여 000을 받게 되어 확률적으로 전문이 바르게 전달된 것으로 생각한다. 이와같이 복호하는 방법을 最近方復號法(nearest neighborhood decoding method)이라고 한다. 또한 이 부호는 오류가 두 개일 때, 오류가 일어났음을 알지만 잘못된 정보를 얻게 됨을 유의하자.

위와같이 오류를 검출하고 수정하기 위하여 통신로는 다음에 정의하는 대칭통신로이어야 한다.

정의 2. 3 부호어가 통신로를 통하여 전달될 때 일어나는 오류가 다음의 성질을 가지면 이 통신로를 對稱通信路(symmetric channel)라 한다.

(I) 부호어를 이루는 각각의 기호에서 오류가 발생할 확률은 같다.

(II) 수신된 기호가 잘못 전달되었을 때, 그 기호를 제외한 다른 기호일 확률은 각각 같다.

부호어가 통신로를 통하여 전달될 때에 수신벡터를 최근방복호법으로 복호하는 경우 수정할 수 있는 오류의 개수의 범위를 알기 위하여 부호의 최소거리를 정의한다.

정의 2. 4 길이가 n인 부호 C의 두 부호어 $\mathbf{X}=x_1x_2\cdots x_n$과 $\mathbf{Y}=y_1y_2\cdots y_n$에 대하여

$$d(\mathbf{X}, \mathbf{Y})=|\{i : x_i \neq y_i\}|$$

를 **X**와 **Y** 사이의 Hamming 距離(Hamming distance between **X** and **Y**) 또는 간단히 距離라 한다. 또한, 부호 C의 最少距離(minimum distance)를

$$d(C)=\min\{d(\mathbf{X}, \mathbf{Y}) : \mathbf{X}, \mathbf{Y}\in C\text{이고 } \mathbf{X}\neq\mathbf{Y}\}$$

라 한다.

알파베트 $\mathbf{Z}_q=\{0, 1, \cdots, q-1\}$ 위에서 부호의 길이가 n, 최소거리가 d이고 부호어의 계수가 M개인 부호를 q 進 (n, M, d)-符號(q-ary (n, M, d)-code)라 한다. 일반적으로 q 진 (n, M, d)-부호에서 검출 및 수정할 수 있는 오류의 개수의 범위에 관하여 다음정리가 성립한다.

정리 2. 5 부호 C가 q 진 (n, M, d)-부호일 때,

(1) $d\geq s+1$이면 s개 까지의 오류를 검출할 수 있다.

(2) $d\geq 2t+1$이면 t개 까지의 오류를 수정할 수 있다.

정의 2. 6 n차원 공간 $\mathbf{Z}_q^n$의 임의의 벡터 $\mathbf{X}=x_1\cdots x_n$에 대하여

$$\omega(\mathbf{X})=|\{i : x_i\neq 0\}|$$

을 벡터 **X**의 Hamming 무게(weight) 또는 간단히 무게라 한다.

또한 부호 C에 대하여

$$\omega(C)=\min\{\omega(\mathbf{X}) \mid \mathbf{X}\in\mathbf{C}\setminus\{0\}\}$$

을 C의 Hamming 무게 또는 간단히 最少무게(minimum weight)라 한다.

## 3. 선형부호

일반적으로 부호에서 수정할 수 있는 誤謬의 범

위와 부호의 最少距離의 관계를 2절에서 살펴보았다. 또한 주어진 q 진 (n, M)－부호의 최소거리를 구하려면 M개의 부호어에서 서로 다른 두 개의 부호어의 거리를 구하여 그들의 최소값을 취하여야 하므로 매우 복잡하다. 또한, 부호를 나타낼 때에 모든 符號語를 나열하여야 하므로 부호어의 개수가 많은 경우는 곤란할 뿐만 아니라 부호화 및 復號에 많은 시간을 소비해야 한다. 이와같은 문제를 效率的으로 解決하기 위하여 부호에 代數學의 理論을 導入한다. 이제부터 다음과 같은 몇가지 記號를 가정하고 시작한다.

(Ⅰ) q를 素數의 거듭제곱으로 가정하고 q개의 서로 다른 기호들의 집합(알파베트)으로 Galois 體를 사용하고 기호로 $\mathbf{F}_q$를 사용한다.

(Ⅱ) n 次元 空間 $\mathbf{F}_q^n$은 n차원 벡터공간이다.

(Ⅲ) n 차원 벡터공간 $\mathbf{F}_q^n$의 元素인 벡터 **X**를 $\mathbf{X}=x_1\cdots x_n$ 또는 $\mathbf{X}=(x_1\ \cdots\ x_n)$으로 편리한 대로 사용한다. 일반적으로 行列의 演算을 시행할 때에는 $[x_1\ x_2\ \cdots\ x_n]$을 주로 사용한다.

**정의 3. 1** n차원 벡터공간 $\mathbf{F}_q^n$의 부분공간을 $\mathbf{F}_q$ 위에서의 q 진 線型符號(linear code) 또는 群符號(group code)라 한다.

선형부호는 n 차원 벡터공간 $\mathbf{F}_q^n$의 부분공간이므로 그의 차원은 n차 이하이고 또한 모든 부분공간은 덧셈에 관한 항등원을 영벡터로 가지므로, 선형부호는 영벡터 **0**를 부호어로 항상 가진다. 한편, 부호의 길이가 n이고 k 차원 부분공간인 선형부호를 q 진 [n, k]－부호라 나타내고, 최소거리 d가 주어질 때 q 진 [n, k, d]－부호라 쓴다. 또한 k 차원 부분공간의 원소의 개수는 $q^k$이므로 q 진 [n, k, d]－부호이면 q 진 $(n, q^k, d)$－선형부호이다.

제 2 절에서 n 차원공간의 임의의 두 벡터 **X**와 **Y**의 거리 d(**X**, **Y**) 및 부호 C의 최소거리 d(C), 벡터 **X**의 Hamming 무게 ω(**X**) 및 부호 C의 최소무게 ω(C)를 정의하였다. 선형부호에서 d(C)와 ω(C)의 관계는 다음과 같다.

**정의 3. 2** 부호 C가 선형부호이면 d(C)＝ω(C)이다.

일반적으로 q 진 (n, M, d)－부호의 최소거리를 구하려면 앞에 설명한 것처럼 M(M－1)/2 쌍의 거리를 구하여 최소값을 구하므로 복잡하지만, q 진 (n, M, d)－선형부호에서는 정리 3. 2에 의하여 **0**을 제외한 M－1개의 부호어들의 Hamming 무게를 구함으로서 최소거리를 쉽게 구할 수 있다.

**예 3. 3** 다음의 세 부호는 모두 선형부호이고 각각의 최소거리는 1, 2 및 3이다.

| 정보 부호 | $C_1$ | $C_2$ | $C_3$ |
|---|---|---|---|
| 동 | 00 | 000 | 00000 |
| 서 | 01 | 011 | 01101 |
| 남 | 10 | 101 | 10110 |
| 북 | 11 | 110 | 11011 |

모든 벡터공간은 基底(basis)를 가지므로 선형부호에서 부호어를 모두 나열하는 대신에 기저의 원소들만 나열하여도 부호어 전체를 알 수 있다. 따라서 부호어를 많이 가지는 부호를 표현하는데 효율적이다. 다음에 설명하는 生成行列(generator matrix)을 이용하면 선형부호의 모든 부호어를 나타내는데 편리하다.

**정의 3. 4** 선형부호가 q 진 [n, k]－부호일 때 기저를 이루는 모든 부호어를 행으로 하는 k×n 행렬을 生成行列(generator matrix)이라 한다.

**정의 3. 5** 예 3. 3의 부호 $C_1$, $C_2$ 및 $C_3$의 생성행렬을 각각 $G_1$, $G_2$ 및 $G_3$로 나타내면 다음과 같다.

$$G_1=\begin{bmatrix}0&1\\1&0\end{bmatrix}\qquad G_2=\begin{bmatrix}0&1&1\\1&0&1\end{bmatrix}$$

$$G_3=\begin{bmatrix}0&1&1&0&1\\1&0&1&1&0\end{bmatrix}$$

앞에서 주어진 부호를 간단하게 생성행렬로 표현하는 예들을 살펴보았다. 한편, 생성행렬이 주어지면 부호를 구할 수 있다.

정리 3. 6 Galois 체 $F_q$ 위에 [n, k]-부호 C의 생성행렬을 G라 하고 G의 k개의 행벡터를 각각 $R_1$, $R_2$, …, $R_k$라 하면, 임의의 부호어는 행벡터들 $R_1$, $R_2$, …, $R_k$의 一次結合으로 唯一하게 표현된다. 즉,

$$C=\{a_1R_1+\cdots+a_kR_k \mid a_1, a_2, \cdots, a_k \in F_q\}$$

이다. 따라서 C는 $q^k$개의 부호어를 가진다.

예 3. 7 생성행렬 G가 Galois체 $F_3$ 위에서 다음과 같이 주어지면

$$G_3=\begin{bmatrix} 1 & 0 & 2 & 1 \\ 1 & 1 & 0 & 2 \end{bmatrix}$$

선형부호 C는 C={0000, 1021, 2012, 1102, 2201, 2120, 1210, 0222, 0111}이다.

정리 3. 8 행렬 G가 [n, k]-부호 C의 생성행렬이면, G에 다음과 같은 연산 (R1), (R2), (R3), (C1) 및 (C2)를 적당히 시행하여 표준형 $[I_k \mid A]$을 구할 수 있다.

(R1) 두 행을 서로 교환한다.

(R2) 한 행에 0이 아닌 $F_q$의 원소를 곱한다.

(R3) 한 행에 0이 아닌 $F_q$의 원소를 곱한 것을 다른 행에 합한다.

(C1) 두 열을 서로 교환한다.

(C2) 한 열에 0이 아닌 $F_q$의 원소를 곱한다.

이와 같은 생성행렬 $[I_k \mid A]$으로 부터 얻은 부호를 부호 C와 동치인 부호라 하고 생성행렬을 표준형으로 가지는 선형부호를 주로 이용한다.

정리 3. 9 부호 C를 $F_q$ 위의 [n, k]-부호라 하고 그의 생성행렬 G를 표준형 $G=[I_k \mid A]$이라 하자. 단, $A=[a_{ij}]$는 $k\times(n-k)$ 행렬이다. 또한, 전문벡터를 $U=u_1\cdots u_k$라 하고 그에 대응하는 부호를 $X=UG=x_1x_2\cdots x_n$이라고 하면 다음이 성립한다.

$$x_i=u_i(1\le i\le k),\quad x_{k+1}=\sum_{j=1}^{k} a_{ij}\,u_j\quad (1\le i\le n-k)$$

정리 3. 9를 이용하면 $q^k$개의 전문을 부호화할 때 $F_q^k$의 벡터들과 대응시키고 이 전문벡터 $U=x_1\cdots x_k$를 부호어 UG로 부호화한다. 이 때에 부호어 UG이 처음 k자리로 부터 전문의 내용을 알 수 있고 나머지 n-k 자리의 숫자는 잘 전달된 부호어인가를 확인하는데 사용된다. 이런 과점에서 $x_1x_2\cdots x_k=U$를 電文자리(message digit)라 하고 $x_{k+1}\cdots x_n$을 檢査자리(check digit)라 한다. 따라서 $q^k$개의 정보를 부호할 때 생성행렬을 표준형으로 가지는 선형부호를 이용하면 쉽다.

정의 3. 10 n 차원 벡터공간 $F_q^n$의 임의의 두 벡터 $X=x_1x_2\cdots x_n$와 $Y=y_1y_2\cdots y_n$에 대하여

$$X\cdot Y=x_1y_1+x_2y_2+\cdots+x_ny_n$$

을 X와 Y의 內積(inner product)이라 한다. 특히, $X\cdot Y=0$일 때, 벡터 X와 Y는 直交(orthogonal, perpendicular) 또는 垂直이라고 말한다.

이 정의로 부터 다음 정리를 알 수 있다.

정리 3. 11 벡터공간 $F_q^n$의 임의의 세 벡터 X, Y, Z와 $F_q$의 임의의 두 원소 α, β에 대하여 다음이 성립한다.

(I) $X\cdot Y=Y\cdot X$

(II) $(\alpha X+\beta Y)\cdot Z=\alpha(X\cdot Z)+\beta(Y\cdot Z)$

주어진 q 진 [n, k]-부호 C에 대하여 부호 C의 모든 부호어와 직교인 벡터의 집합을 $C^p$라고 하자. 즉,

$$C^p=\{V\in F_q^n \mid C\text{의 임의의 부호어 } X\text{에 대하여 } X\cdot V=0\}$$

이고 정리 3. 11에 의하여 주어진 [n, k]-부호 C에 대하여 벡터공간 $F_q^n$의 임의의 벡터 V가 $C^p$의 원소일

필요충분조건은 V가 생성행렬 G의 모든 행벡터와 직교한다. 즉, $V \in C^p \Leftrightarrow VG^T = 0$이다.

정리 3. 12 주어진 q 진 [n, k]-부호 C에 대하여 $C^p$는 [n, n-k]-부호이다. 이 때 $C^p$을 C의 雙對符號(dual code) 또는 直交符號라 하고 $C^p$의 생성행렬 H를 부호 C의 parity 檢査行列(parity-check matrix)이라 한다.

예 3. 13 (1) 선형부호 C가 C={0000, 1100, 0011, 1111}이면 C의

parity 검사행렬 H는 $H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ & & & \\ 0 & 0 & 1 & 1 \end{bmatrix}$ 이다.

정리 3. 14 행렬 $G=[I_k \mid A]$이 [n, k]-부호 C의 생성행렬이면, $H=[-A^T \mid I_{n-k}]$는 부호 C의 parity 검사행렬이다. 단 $A^T$는 A의 전치행렬이다.

## 4. 완전부호

정리 2. 5에 의하여 부호의 최소거리가 크면 클수록 오류를 수정할 수 있는 개수의 범위가 커짐으로서 정확한 정보를 얻을 수 있는 확률이 높아진다. 부호의 길이가 길면 길수록 전문을 전송하는데 걸리는 시간이 길어진다. 또한 부호어의 개수가 많으면 많은 정보를 표현할 수 있다. 즉, 부호의 길이는 짧을수록, 부호의 최소거리는 클수록 그리고 부호어의 개수는 많을수록 좋을 것이다. 다시 말하면 좋은 q 진 (n, M, d)-부호란 가능한 한 작은 n, 그리고 가능한 한 큰 M과 큰 d를 갖는 부호이다. 실제로 이것은 불가능하고, 두 변수를 고정시켰을 때 나머지 변수를 최적화하는 것과 부호화 및 복호를 쉽게 할 수 있는 부호를 개발하는 일이 부호이론의 중요한 문제이다.

부호의 길이 n과 최소거리 d가 주어질 때에 부호어의 개수 M을 최대화하는 문제를 생각하여 보자. 이 때 최대값 M을 기호로 $A_q(n, d)$라고 쓰면 다음 정리를 쉽게 알 수 있다.

정리 4. 1 (1) $A_q(n, 1) = q^n$ (2) $A_q(n, n) = q$

동치부호의 개념을 이용하면 $A_2(5, 3) = 4$은 어렵지 않게 구할 수 있다. 일반적으로 $A_q(n, d)$를 구하는 문제는 복잡하다. 다음 표 4-1은 이진법에서 부호의 길이 n과 최소거리 d의 값이 주어졌을 때 부호어의 개수 M의 값의 범위를 구하여 놓은 것이다.

표 4-1

| n | d = 2 | d = 5 | d = 7 |
|---|---|---|---|
| 5 | 4 | 2 | |
| 6 | 8 | 2 | |
| 7 | 16 | 2 | 2 |
| 8 | 20 | 4 | 2 |
| 9 | 40 | 6 | 2 |
| 10 | 72-79 | 12 | 2 |
| 11 | 144-158 | 24 | 4 |
| 12 | 256 | 32 | 4 |
| 13 | 512 | 64 | 8 |
| 14 | 1024 | 128 | 16 |
| 15 | 2048 | 256 | 32 |
| 16 | 2560-3276 | 256-340 | 36-37 |

부호의 최소거리가 주어졌을 때 수정할 수 있는 오류의 개수의 범위에 관한 정리를 2절에서 살펴보았다. 여기에서는 모든 수신벡터가 최근방복호법으로 복호되는가를 생각하여 보자.

예 4. 2 최소거리가 3인 부호 C={00000, 01101, 10110, 11011}에서 수신벡터 01010을 얻었다면 수신벡터와 각 부호어와의 거리가 각각 2, 3, 3 및 2가 되어 오류의 개수가 2개 이상되어 복호할 수가

없다.

위 예에서와 같이 수신벡터에 대하여 복호할 수 없는 부호들이 있다. 이러한 이유는 다음의 정의 및 정리로 부터 알 수 있다.

정의 4. 3 n 차원 공간 $\mathbf{Z}_q^n$에 속하는 임의의 벡터 **U**와 임의의 음이 아닌 정수 t에 대하여 집합

$\{\mathbf{V}\in\mathbf{Z}_q^n \mid d(\mathbf{U}, \mathbf{V})\leq t\}$을 中心이 **U**이고 반지름이 t인 n次元 球(n dimmensional sphere with center U and radius t)라 하고 기호로 S(**U**, t)라 쓴다.

반지름이 t인 n 차원 구 S(**U**, t)에 대하여

$$|S(\mathbf{U}, t)| = \binom{n}{0}+\binom{n}{1}(q-1)+\cdots+\binom{n}{t}(q-1)^t.$$

이고, 최소거리가 d=2t+1인 q 진 부호 C의 임의의 두 부호어를 **X**와 **Y**라면 $S(\mathbf{X}, t)\cap S(\mathbf{Y}, t)=\phi$이므로 다음 정리가 성립한다.

정리 4. 4 임의의 q 진 (n, M, 2t+1)-부호 C는 다음식을 만족한다.

$$M\{\binom{n}{0}+\binom{n}{1}(q-1)+\binom{n}{2}(q-1)^2+\cdots+\binom{n}{t}(q-1)^t\}\leq q^n \cdots \quad (4.1)$$

이 부등식 (4. 1)을 Hamming 上界(Hamming bound)라 한다. 식 (4. 1)의 등호가 성립할 때에 부호 C를 完全符號(perfect code)라 한다.

예 4. 2의 이진 (5, 4, 3)-부호 C={00000, 01101, 10110, 11011}에서 식 (4. 1)의 좌변은 t=1 이므로 24이고 우변은 $2^5=32$이므로 (4. 1)의 등호가 성립하지 않는다. 즉, 이 부호는 완전부호가 아니며, 8개의 수신벡터는 임의의 부호어와 거리가 2이상이 되어 복호할 수 없다. 한편, 정리 4. 4에 있는 완전부호의 정의에 의하여 완전부호에서는 모든 수신벡터가 복호되는 것을 알 수 있을 뿐만 아니라, 실제로 n과 d=2t+1이 주어진 경우 $A_q(n, d)=M$을 만족하는 부호임을 쉽게 알 수 있다.

예 4. 5 (1) $F_q^n$

(2) n이 홀수일 때 이진 반복부호

$$C=\{0\cdots0,\ 1\cdots1\}$$

정리 4. 6 q 진 [n, k]-부호 C의 parity 검사행렬을 H라 하면 부호 C의 최소거리가 d(C)=d일 필요충분조건은 다음과 같다.

(I) 행렬 H의 임의의 d-1개의 열벡터는 일차독립이다.

(II) 일차종속인 d개의 열벡터가 존재한다.

정리 4. 6은 실제로 예 4. 7과 예 4. 9에서 Hamming 부호의 최소거리가 3임을 증명하는데, 이용된다.

예 4. 7 정수 r(≥2)에 대하여 벡터공간 $\mathbf{F}_2^r$의 영벡터를 제외한 $2^r-1$개의 서로 다른 벡터를 열벡터로 가지는 $r\times(2^r-1)$ 행렬을 parity 검사행렬로 가지는 부호를 이진 Hamming 부호라 하고 Ham(r, 2)로 나타낸다. 또한 Ham(r, 2)의 parity 검사행렬은 $r\times(2^r-1)$ 행렬이므로 부호 Ham(r, 2)의 부호의 길이는 $n=2^r-1$ 이다. 따라서 정리 3. 14에 의하여 Ham(r, 2)는 $[2^r-1, 2^r-1-r, 3]$-완전부호이다. 이를테면, r=3 이면 $n=2^3-1=7$ 이고 parity 검사행렬 H는

$$H=\begin{bmatrix}0&0&0&1&1&1&1\\0&1&1&0&0&1&1\\1&0&1&0&1&0&1\end{bmatrix}$$

이고 이 행렬의 열들의 순서를 적당히 바꾸면 다음과 같은 표준형 H를

$$H=\begin{bmatrix}0&1&1&1&1&0&0\\1&0&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$$

얻을 수 있다. 따라서 이 표준형을 parity 검사행렬로 가지는 선형부호 C의 생성행렬 G는 정리 3. 14에 의하여

$$G=\begin{bmatrix}1&0&0&0&0&1&1\\0&1&0&0&1&0&1\\0&0&1&0&1&1&0\\0&0&0&1&1&1&1\end{bmatrix}$$

이다.

정리 4. 8 r차원 벡터공간 $F_q{}^r$에는 다음 두 조건을 만족하는 부분집합 A가 존재한다.

(1) A는 $\frac{q^r-1}{q-1}$ 개의 벡터을 가진다.

(2) A의 임의의 두 벡터는 일차독립이다.

예 4. 9 정리 4. 8의 집합 A의 모든 벡터들을 열벡터로 가지는

$r\times\frac{q^r-1}{q-1}$ 행렬을 parity 검사행렬로 가지는 선형부호를 q 進 Hamming 符號라 하고 Ham(r, q)로 나타낸다(r≥2). 그러면 Ham(r, q)가

$$\left[\frac{q^r-1}{q-1},\ \frac{q^r-1}{q-1}-r,\ 3\right]-\text{완전부호}$$

이다. 이를테면 Ham(2, 3)의 parity 검사행렬 H는

$$H=\begin{bmatrix}0&1&1&1\\1&0&1&2\end{bmatrix}$$

이므로 이것의 표준형을 구하면

$$H=\begin{bmatrix}1&1&1&0\\1&2&0&1\end{bmatrix}\qquad G=\begin{bmatrix}1&0&2&2\\0&1&2&1\end{bmatrix}$$

이고 행렬 G에 의하여 생성된 부호가 Ham(2, 3)이다.

예 4. 10 생성행렬 G가

$$G=[I_{12}\mid A]=\begin{bmatrix}
1&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1&1&1&1\\
0&1&0&0&0&0&0&0&0&0&0&0&1&1&1&0&1&1&1&0&0&0&1\\
0&0&1&0&0&0&0&0&0&0&0&0&1&1&0&1&1&1&0&0&0&1&0\\
0&0&0&1&0&0&0&0&0&0&0&0&1&0&1&1&1&0&0&0&1&0&1\\
0&0&0&0&1&0&0&0&0&0&0&0&1&1&1&0&0&0&1&0&1&1\\
0&0&0&0&0&1&0&0&0&0&0&0&1&1&1&0&0&0&1&0&1&1&0\\
0&0&0&0&0&0&1&0&0&0&0&0&1&1&0&0&0&1&0&1&1&0&1\\
0&0&0&0&0&0&0&1&0&0&0&0&1&0&0&0&1&0&1&1&0&1&1\\
0&0&0&0&0&0&0&0&1&0&0&0&1&0&0&1&0&1&1&0&1&1&1\\
0&0&0&0&0&0&0&0&0&1&0&0&1&0&1&0&1&1&0&1&1&1&0\\
0&0&0&0&0&0&0&0&0&0&1&0&1&1&0&1&1&0&1&1&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&1&1&0&1&1&0&1&1&1&0&0&0
\end{bmatrix}$$

로 주어진 2 진 부호 $G_{23}$은 [23, 12, 7]-완전부호이다. 이 부호를 2 진 Golay 부호라 한다.

예 4. 11 생성행렬 G가

$$G=[I_6 \mid A]=\begin{bmatrix} 1&0&0&0&0&0&0&1&1&1&1 \\ 0&1&0&0&0&0&1&0&1&2&2 \\ 0&0&1&0&0&0&1&1&0&1&2 \\ 0&0&0&1&0&0&1&2&1&0&1 \\ 0&0&0&0&1&0&1&2&2&1&0 \\ 0&0&0&0&0&1&1&1&2&2&1 \end{bmatrix}$$

로 주어진 3 진 부호 $G_{11}$은 [11, 6, 5]-완전부호이다. 이 부호를 3 진 Golay 부호라 한다.

## 5. 완전부호의 존재성

부호이론에 중요한 부호들 중에 하나는 4절에서 설명한 것처럼 모든 수신벡터를 復號할 수 있을 뿐만 아니라 부호어를 $A_q(n, d)$개를 가지는 完全符號이다. 제 4 절에서 지금까지 설명한 완전부호의 예는 다음과 같은 선형부호이다.

(i) $F_q^n$
(ii) n이 홀수일 때 이진 반복부호
$C=\{0\cdots0, 1\cdots1\}$
(iii) Hamming 부호 Ham(r, q), 단 q는 소수의 거듭제곱이다.
(iv) 이진 [23, 12, 7]-Golay 부호
(v) 3 진 [11, 6, 5]-Golay 부호

이 절에서는 위와 같은 부호들을 포함한 완전부호의 존재성을 알아 보자. 이와같은 완전부호를 찾으려는 시도는 1949년 M. Golay에 의해 시작되었다. 그는 q 진 (n, M, 2t+1) 부호가 완전부호가 되기 위한 필요충분조건 즉, 식 (4. 1)의 등호를 만족하는 변수 n, M 및 2t+1을 찾았다. 그 결과로 (i), (ii)와 이진 Hamming 부호를 만들었고, 식 (4. 1)의 등호를 만족하는 변수들은 (23, $2^{12}$, 7), (90, $2^{78}$, 5)와 (11, $3^6$, 5)의 3가지 밖에 없다고 제안하였다. 또한, (23, $2^{12}$, 7)-부호와 (11, $3^6$, 5)-부호 즉, 2 진 Golay 부호와 3 진 Golay 부호를 만들었고, (90, $2^{78}$, 5)-부호는 존재하지 않음을 증명하였다(정리 5. 1).

이것은 놀라웁게도 1967년 J. H. van Lint가 전자계산기를 이용하여

$n<1000$, $t \leq 1000$이고 $q\leq 100$

인 변수에서는 Golay의 주장이 옳다는 것을 확인하였다. 또한 1950년 Hamming은 Golay의 결과를 확장하여 q 진 Hamming 부호를 만들었다. 이후 선형완전부호에 관하여는 1973년 A. Tietäväinen과 van Lint에 의하여 앞에 나열한 5가지 경우 밖에 없음이 증명되었다[정리 5. 3과 따름정리 5. 5]. 같은 해 이 문제는 V. A. Zinov'ev와 V. K. Leont'ev에 의하여 독립적으로 증명되었다.

정리 5. 1 이진 [90, 78, 5]-완전부호는 존재하지 않는다.

Tietäväinen과 van Lint는 이 문제(5. 3)를 해결하기 위하여 1957년에 Llody가 증명한 정리 5. 2를 이용하였다.

정리 5. 2 q 진 완전 (n, M, 2t+1)-부호가 존재한다면 다항식 $L_t(x)$는 구간 [1, n]에서 t개의 서로 다른 정수근을 가진다.

단, $L_t(X)=\sum_{j=0}^{t}(-1)^j(q-1)^{t-j}\binom{x-1}{j}\binom{n-x}{t-j}$ 이다.

정리 5. 3 q가 소수의 거듭제곱일 때 (i)과 (ii)를 제외한 q 진 완전부호는 Hamming 부호 또는 Golay 부호와 같은 변수를 가진다.

S. P. Lloyd의 정리는 정리 5. 3의 증명에서 뿐만 아니라 폭넓게 응용된다. 특히, Lenstra와 A. M. Odlyzko에 의하면 Lloyd의 정리를 만족하려면 변수들은 $t\leq 11$, $q\leq 8$이고 $n\leq 485$이 되어야 한다는 사실이 밝혀졌다. 이 사실은 식 (4. 1)을 만족하는

변수의 범위를 제한하여 완전부호의 존재성의 연구에 크게 기여했다. 또한, Llody 정리를 이용하면 정리 5. 1을 다음 예에서와 같이 쉽게 증명할 수도 있다.

예 5. 4 이진 [90, 78, 5]-완전부호가 존재한다면 다항식 $L_2(x)$는 Llody 정리에 의하여

$$L_2(x)=\sum_{j=0}^{2}(-1)^j(2-1)^{2-j}\binom{x-1}{j}\binom{n-x}{t-j}$$

는 2개의 서로 다른 정수근을 가져야 한다. 한편, 방정식 $L_2(x)=0$을 정리하면

$$x^2-91x+2048=0$$

이고 이 이차방정식은 정수근을 갖지 않는다. 따라서 Llody 정리에 모순이므로 이진 [90, 78, 5]-완전부호는 존재하지 않는다.

한편 V. Pless는 1968년에 Golay 부호와 같은 변수를 가지는 선형부호는 Golay 부호와 동치임을 증명하였다. 이를 이용하면,

따름정리 5. 5 q가 소수의 거듭제곱일 때 (i)과 (ii)를 제외한 선형완전부호는 Hamming 부호 또는 Golay 부호와 동치이다.

지금까지 선형부호에서 완전부호의 존재성을 살펴보았다. 1962년 J. L. Vasil'ev가 이진 Hamming 부호와 같은 변수를 가지는 선형 부호가 아닌 완전부호의 집합들을 찾을 때 까지 잠시동안 완전부호는 오직 위에서 설명한 5가지만 존재하는 것으로 믿었다. 그러나 이후 q 진 Hamming 부호와 같은 변수를 가지는 선형부호가 아닌 완전부호를 1968년 J. Schönheim이 그리고 1969년 B. Lindström이 발견하는 것을 시작으로 선형부호가 아닌 완전부호의 존재성을 연구하게 되었다. 1973년 S. L. Snover는 이진 Golay 부호와 같은 변수를 가지는 완전부호는 유일하고, 1975년 P. Delsarte와 J. M. Goethals는 3 진 Golay 부호와 같은 변수를 가지는 완전부호는 유일하다는 것을 증명하였다. 이와는 대조적으로 1983년 K. T. Phels는 Ham(4, 2)와 같은 변수를 가지는 동치가 아닌 최소한 수천 개의 $(15, 2^{11}, 3)$-완전부호가 존재하는 것을 증명하였다. 또한 1982년 M. R. Best는 다음과 같은 정리를 증명하였다.

정리 5. 6 $t\geq 3$이고 $t\neq 6$, $t\neq 8$일 때 t개의 오류를 수정할 수 있는 완전부호는 이진 Golay 부호 뿐이다.

이 정리의 발표 이후 곧 바로 1984년 $t=6$과 $t=8$인 경우에 Y. Hong에 의하여 밝혀졌으며, $t=2$인 경우도 정리 되었다. 따라서 결국 $t=1$인 경우의 문제만 남게 되었다. 이 경우 q가 소수의 거듭제곱일 때는 Hamming 부호와 같은 변수를 가지는 부호로 위에 설명한 몇 가지가 알려져 있다. 그러나 q가 소수의 거듭제곱이 아닌 경우에는 거의 알려지지 않고 있다. q의 값과 t의 값이 아주 적은 특수한 경우인 $q=6$이고 $t=2$인 경우의 부호는 1958년 Golay에 의하여 제안되어, S. W. Golomb과 E. C. Posner에 의하여 1964년 존재하지 않음이 증명되었다. 즉, 6진 $(7, 6^5, 3)$-완전부호는 존재하지 않는다. 이 정리는 궁극적으로 1782년 Euler가 제안한

"Euler's 36 officers problem"

이라는 문제로 추론된다.

## 참 고 문 헌

1. Berlekamp, E. R., Algebraic coding theory, McGraw-Hill, New York, 1968.

2. Best. M. R., Binary codes with a minimum distance of four, IEEE Trans. Info. Theory 26, 738-42, 1980.

3. Best. M. R., A contribution to the nonexistence of perfect codes, Ph. D. dissertation, Universty of Amsterdam, 1980.

4. Cameron, P. J. and van Lint, J. H., Graphs, codes and designs, London Math. Soc. Lec-

ture Note Series, Vol. 43, Cambridge Univ. Press, Cambridge, 1980.

5. Hamming, R. W., Coding and information theory, Prentice-Hall, New Jersey, 1980.

6. Hill, R., A first course in coding theory, Clarendon Press, Oxford, 1985.

7. Lidl, R. and Niederreiter, H., Finite Fields, Encyclopedia of Mathematics and its Applications, Cam. Univ., Press, 1983.

8. Van Lint, J. H., A survey of perfect codes. Rocky Mountain J. of Mathematics 5, 199-224, 1975.

9. Van Lint, J. H., Introduction to coding theory, Springer-Verlag, New York, 1982.

10. Macwilliams, F. J. and Sloane, N. J. A., The theory of error-correcting codes, North-Holland, Amsterdam, 1977.

11. Pless, V., On the uniqueness of the Golay codes, J. Comb. Theory 5, 215-28, 1968.

12. Tietäväinen, A., On the nonexistence of perfect codes over finite fields, SIAM J. Appl. Math. 24, 88-96, 1973.

13. 金應泰·朴勝安, 線型代數學, 청문각, 1991.

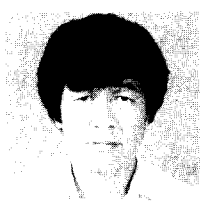
14. 金應泰·朴勝安, 整數論, 경문사, 1991.

15. 金應泰·朴勝安, 現代代數學, 경문사, 1991.

16. 朴勝安·李敏燮·李載學·申鉉容, 代數的符號理論, 체신부 연구보고서, 1991.

17. 李晩榮, 1984 符號理論, 喜重堂, 1984.

## □ 著者紹介



李 敏 燮(終身會員)

서울大學校 師範大學 數學科(理學士)
西江大學校 大學院 數學科(理學碩士)
University of Alabama 大學院 數學科(理學博士)
현재 檀國大學校 自然科學大學 數學科 副教授
忠淸 數學會 理事