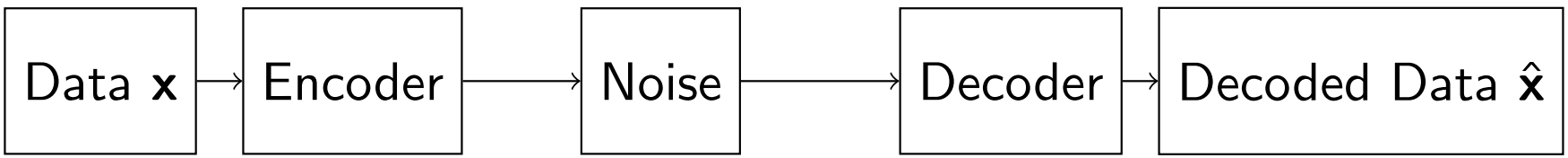
Lecture 2: Lattices from Codes

F. Oggier

Journées Algébriques du Gabon, Libreville, Mars 2025

Behind Codes

Transmitter Channel Receiver



$$\mathbf{x} = (x_1, \ldots, x_k) \mapsto \underbrace{\mathbf{c} = (c_1, \ldots, c_n)}_{\text{codeword, } n \geq k} \longrightarrow \mathbf{y} \to \hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_k)$$

where $\mathbf{y}$ is the noisy channel output.

- The vector $\mathbf{x} = (x_1, \ldots, x_k)$ of information symbols, the codeword $\mathbf{c} = (c_1, \ldots, c_n)$ both have coefficients in a given alphabet, typically a finite field $\mathbb{F}_q$ or the set $\mathbb{Z}_q = \{0, 1, \ldots, q-1\}$ of integers modulo q, for $q \geq 2$ a positive integer.
- If $q = p$ is a prime, $\mathbb{Z}_p = \mathbb{F}_p$ is a finite field, otherwise we have a finite ring.
- The arithmetic that follows depends on the choice of $\mathbb{F}_q$ or $\mathbb{Z}_q$.

- In $\mathbb{Z}_4$:

$+$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$\cdot$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

- In $\mathbb{F}_4$:

$+$	0	1	ω	ω^2
0	0	1	ω	ω^2
1	1	0	ω^2	ω
ω	ω	ω^2	0	1
ω^2	ω^2	ω	1	0

$\cdot$	0	1	ω	ω^2
0	0	0	0	0
1	0	1	ω	ω^2
ω	0	ω	ω^2	1
ω^2	0	ω^2	1	ω

for an element ω which is a zero of $X^2 + X + 1 \pmod 2$, so $\omega \neq 0, 1$, $\omega^2 = \omega + 1$, $\omega^3 = \omega(\omega + 1) = \omega^2 + \omega = 1 \pmod 2$.

- We assume the encoding of a vector of information symbols into a codeword

$$\mathbf{x} = (x_1, \ldots, x_k) \mapsto \mathbf{c} = (c_1, \ldots, c_n), n \geq k$$

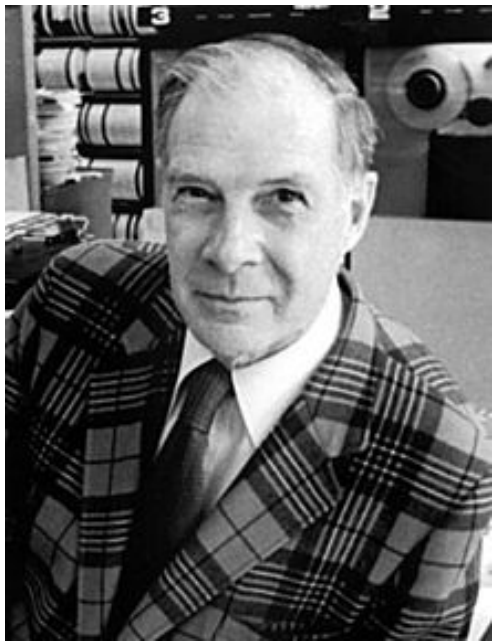
  is linear.
- The set of all codewords, denoted by C, is called a linear code.
- If we chose $\mathbb{F}_q$, codewords belong to $\mathbb{F}_q^n$, which is a vector space.
- Since the encoding is linear, the sum of two codewords is again a codeword, a multiple of a codeword is again a codeword, the whole zero codeword $\mathbf{0} \in C$ and C forms a linear subspace of $\mathbb{F}_q^n$. It thus has a dimension, namely k, and we call n the length.
- We say that we consider an (n, k) linear code C over $\mathbb{F}_q$.

Since an (n, k) linear code C is a subspace, it has a basis, containing k vectors, that we stack as rows of a $k \times n$ matrix G, called a generator matrix for C.

- There are many generator matrices.
- There is a unique generator matrix of the form $G = [\mathbf{I}_k | A]$ where $\mathbf{I}_k$ is the identity matrix. The code is said to be in systematic form.

$$\underbrace{(x_1, \ldots, x_k)}_{\text{information data}} \underbrace{\begin{bmatrix} & a_{11} & a_{1,n-k} \\ \mathbf{I}_k & \vdots & \vdots \\ & a_{k,1} & a_{k,n-k} \end{bmatrix}}_{\text{generator matrix } G} = \underbrace{(x_1, \ldots, x_k, c_{k+1}, \ldots, c_n)}_{\text{codeword}}$$

- For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{F}_q^n$, their Hamming distance $d_H(\mathbf{x}, \mathbf{y})$ is the number of coordinates in which they differ.
- For $\mathbf{x} \in \mathbb{F}_q^n$, its weight $wt_H(\mathbf{x})$ counts the number of nonzero coordinates of $\mathbf{x}$.
- If C is linear, then $d_H(\mathbf{c}, \mathbf{c}') = wt_H(\mathbf{c}'')$ for some $\mathbf{c}''$ in C.



(source: wikipedia)

For an (n, k) linear code C over $\mathbb{F}_q$, its minimum distance $d_H(C)$ is the minimum weight of the nonzero codewords of C.

- An (n, k) linear code C over $\mathbb{F}_q$ with minimum Hamming distance $d_H(C) = d$ can recover from $d - 1$ erasures.
- Any two codewords differ in d coordinates.
- If up to $d - 1$ coordinates are missing, there is still at least one left to recognize the codeword.

The $(n, 1)$ repetition code

- Dimension: $k = 1$.
- Length: n.
- Encoding: $(x_1) \mapsto (x_1, \ldots, x_1) \in \mathbb{F}_q^n$.
- Its minimum distance is $d_H(C) = n$: if x_1 is not zero, then $(x_1, \ldots, x_1)$ has weight n.

Generator matrix:

$$(x_1)[1, \ldots, 1] = (x_1 \ldots, x_1).$$

The $(n, n-1)$ single parity check code

- Dimension: $k = n - 1$.
- Length: n
- Encoding: $(x_1, \ldots, x_k) \mapsto (x_1, \ldots, x_k, \sum_{i=1}^{k} x_i) \in \mathbb{F}_q^n$.
- Its minimum distance is $d_H(C) = 2$: the codeword $(1, 0, \ldots, 0, 1)$ has weight 2. It is not possible to have a codeword with weight 1. If there is a single data symbol which is not zero, then the parity symbol is also not zero. If there are at least two data symbols, then the weight is at least 2.

Generator matrix:

$$(x_1, \ldots, x_k) \begin{bmatrix} & & 1 \\ & \mathbf{I}_k & \vdots \\ & & 1 \end{bmatrix} = (x_1 \ldots, x_k, x_1 + \ldots + x_k).$$

The $(4, 3, 2)$ single parity check code

- $(x_1, x_2, x_3) \mapsto (x_1, x_2, x_3, x_1 + x_2 + x_3) \in \mathbb{F}_2^4$.
- Over $\mathbb{F}_2$, codewords are

$$(0,0,0,0), (1,0,0,1), (0,1,0,1), (1,1,0,0),$$
$$(0,0,1,1), (1,0,1,0), (0,1,1,0), (1,1,1,1).$$

- If we receive $(0, *, 1, 0)$,

The $(4, 3, 2)$ single parity check code

- $(x_1, x_2, x_3) \mapsto (x_1, x_2, x_3, x_1 + x_2 + x_3) \in \mathbb{F}_2^4$.
- Over $\mathbb{F}_2$, codewords are

$$(0,0,0,0), (1,0,0,1), (0,1,0,1), (1,1,0,0),$$
$$(0,0,1,1), (1,0,1,0), (0,1,1,0), (1,1,1,1).$$

- If we receive $(0, *, 1, 0)$, we must have sent $(0, 1, 1, 0)$.
- If we receive $(0, *, 1, *)$,

The $(4, 3, 2)$ single parity check code

- $(x_1, x_2, x_3) \mapsto (x_1, x_2, x_3, x_1 + x_2 + x_3) \in \mathbb{F}_2^4$.
- Over $\mathbb{F}_2$, codewords are

$$(0,0,0,0), (1,0,0,1), (0,1,0,1), (1,1,0,0),$$
$$(0,0,1,1), (1,0,1,0), (0,1,1,0), (1,1,1,1).$$

- If we receive $(0, *, 1, 0)$, we must have sent $(0, 1, 1, 0)$.
- If we receive $(0, *, 1, *)$, we could have sent $(0, 1, 1, 0)$ or $(0, 0, 1, 1)$.

The Singleton Bound

For $d_H = d \leq n$, $k \leq n - d + 1$. Equivalently $k \leq n - d + 1 \iff d \leq n - (k - 1)$.

- Project all the codewords on the first $k - 1$ coordinates. Since there are q^k different codewords, by the pigeon-hole principle, at least two of them should agree on these $k - 1$ coordinates.
- These then disagree on at most the remaining $n - (k - 1)$ coordinates. Hence the minimum distance d of the code is $d \leq n - (k - 1)$.

Codes whose parameters reach the Singleton bound are called maximum distance separable (MDS).

- If $G = [\mathbf{I}_k | A]$ is a generator matrix for the (n, k) code C, then $H = [-A^T | \mathbf{I}_{n-k}]$ is called the corresponding parity check matrix.
- We have

$$G^T = \begin{bmatrix} \mathbf{I}_k \\ A^T \end{bmatrix}$$

  and $HG^T = -A^T + A^T = \mathbf{0}$.
- If $\mathbf{c} \in C$, $\mathbf{c} = \mathbf{x}G$ and $H\mathbf{c}^T = HG^T\mathbf{x}^T$.
- Thus C is contained in the kernel of the linear map $\mathbf{v} \mapsto H\mathbf{v}^T$. As H has rank $n - k$, this map has a kernel of dimension k, which is the dimension of C.
- Two view points:

$$C = \{\mathbf{c} = \mathbf{x}G,\ \mathbf{x} \in \mathbb{F}_q^k\},\ C = \{\mathbf{v} \in \mathbb{F}_q^n,\ H\mathbf{v}^T = \mathbf{0}\}$$

The $(n, 1)$ repetition code

Its generator matrix in systematic form:

$$[1, \underbrace{1 \ldots, 1}_{A}] = [\mathbf{I}_k | A].$$

The corresponding parity check matrix in systematic form:

$$[-A^T | \mathbf{I}_{n-k}] = \begin{bmatrix} -1 & \\ \vdots & \mathbf{I}_{n-k} \\ -1 & \end{bmatrix}$$

- Given an (n, k) linear code C over $\mathbb{F}_q$, the dual code $C^\perp$ is the $(n, n-k)$ linear code generated by the rows of its parity check matrix H.
- Equivalently: $C^\perp = \{\mathbf{v} \in \mathbb{F}_q^n,\ \mathbf{c} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{c} \in C\}$.

Indeed:

$$\begin{aligned} C^\perp &= \{\mathbf{v} \in \mathbb{F}_q^n,\ \mathbf{c} \cdot \mathbf{v} = \mathbf{0} \text{ for all } \mathbf{c} \in C\} \\ &= \{\mathbf{v} \in \mathbb{F}_q^n,\ \mathbf{x}G \cdot \mathbf{v} = \mathbf{0} \text{ for all } \mathbf{x} \in \mathbb{F}_q^k\} \end{aligned}$$

Let $\mathbf{g}_i$, $i = 1, \ldots, k$ be the rows of G. Then $0 = \sum_{i=1}^k x_i \mathbf{g}_i \cdot \mathbf{v}$ for any x_i implies $0 = \mathbf{g}_i \mathbf{v}^T$ for every i and

$$C^\perp = \{\mathbf{v} \in \mathbb{F}_q^n,\ G\mathbf{v}^T = \mathbf{0}\}.$$

- A self-orthogonal code C is a code which is included in its dual.
- A self-dual code C is a code which is equal to its dual.

- For $q \geq 2$ a positive integer, consider the set $\mathbb{Z}_q = \{0, 1, \ldots, q-1\}$ of integers modulo q.
- A linear code C in $\mathbb{Z}_q^n$ is by definition a subset which is an additive subgroup of $\mathbb{Z}_q^n$.
- If $q = p$ is a prime, $\mathbb{Z}_p^n = \mathbb{F}_p^n$ is a vector space over the field $\mathbb{Z}_p = \mathbb{F}_p$. A linear code is a subspace of dimension k of the vector space $\mathbb{Z}_p^n = \mathbb{F}_p^n$ (called an (n, k) code).
- In what follows, we will focus on $\mathbb{Z}_q$, meaning that we will consider only finite fields of the form $\mathbb{F}_p$ (though there is an analogous theory of what will be shown for the case $\mathbb{F}_q$).

Let

$$\rho : \mathbb{Z} \to \mathbb{Z}_q = \{0, 1, \ldots, q-1\},\ x \mapsto x \pmod{q},$$

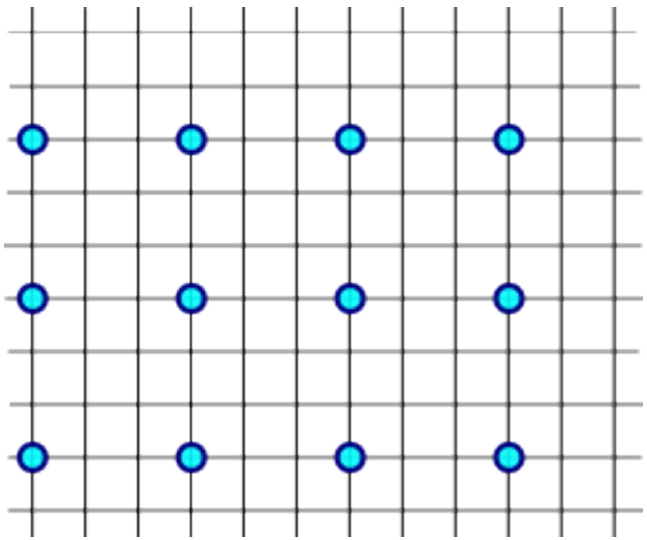
be the map of reduction modulo q. Given $a \pmod{q}$, its pre-image $\rho^{-1}(a)$ is the set of integers that are mapped to a by ρ, that is $\rho^{-1}(a) = \{a + bq,\ b \in \mathbb{Z}\}$.

The pre-image $\rho^{-1}(a) \subset \mathbb{Z}$ of $a \pmod{5}$.

Consider

$$\rho : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}_q \times \mathbb{Z}_q, \ (x_1, x_2) \mapsto (x_1 \ (\mathrm{mod}\, q), x_2 \ (\mathrm{mod}\, q)).$$

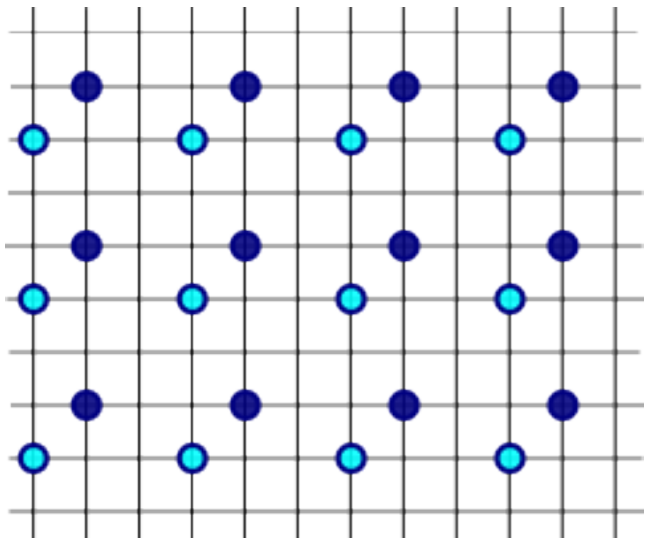
The pre-image $\rho^{-1}((a_1, a_2))$ is now the set of 2-dimensional vectors with integer entries that is mapped to (a_1, a_2) by ρ.



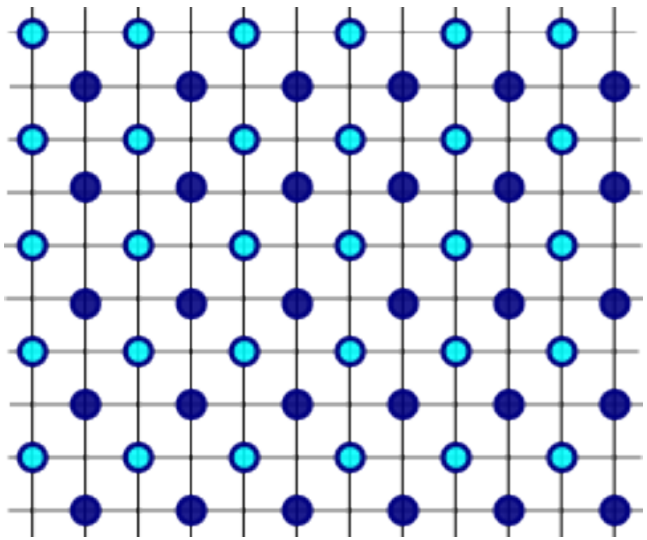
The pre-image $\rho^{-1}((a_1, a_2)) \subset \mathbb{Z}^2$ of $(a_1 \ (\mathrm{mod}\ 3), a_2 \ (\mathrm{mod}\ 3))$.

$$\rho : \mathbb{Z}^n \to \mathbb{Z}_q^n, \ \mathbf{x} \mapsto \rho(\mathbf{x}).$$

One may take any arbitrary subset S of $\mathbb{Z}_q^n$ and compute $\rho^{-1}(S)$, but what if S has a structure? how is it carried over to $\rho^{-1}(S)$?



The pre-image $\rho^{-1}(S) \subset \mathbb{Z}^2$ of $S = \{(a_1 \pmod 3, a_2 \pmod 3), (a_1 + 1 \pmod 3, a_2 + 1 \pmod 3)\}$.



The pre-image $\rho^{-1}(C) \subset \mathbb{Z}^2$ of the linear binary code $C = \{(0,0), (1,1)\}$.

A well known connection between linear codes in $\mathbb{Z}_q^n$ and lattices: Given a subset $S \subset \mathbb{Z}_q^n$ then $\rho^{-1}(S)$ is a lattice in $\mathbb{R}^n$ if and only if S is a linear code in $\mathbb{Z}_q^n$.

A linear code C in $\mathbb{Z}_q^n$ is an additive subgroup of $\mathbb{Z}_q^n$.

A lattice $L \subset \mathbb{R}^n$ is an additive subgroup of $\mathbb{R}^n$ given by

$$L = \{uM,\ u \in \mathbb{Z}^m\}$$

for M a generator matrix, whose rows are $m \leq n$ linearly independent vectors.

Given a subset $S \subset \mathbb{Z}_q^n$ then $\rho^{-1}(S)$ is a lattice in $\mathbb{R}^n$ if and only if S is a linear code in $\mathbb{Z}_q^n$.

Proof Idea.

- Suppose $S \subset \mathbb{Z}_q^n$ is a linear code. We need to check that $\rho^{-1}(C)$ is a discrete additive subgroup of $\mathbb{R}^n$.
- Since $\rho^{-1}(C) \subset \mathbb{Z}^n$, it is a discrete subset of $\mathbb{R}^n$.
- Take $\mathbf{x}, \mathbf{y}$ two arbitrary vectors in $\rho^{-1}(C)$. Their sum must belong to $\rho^{-1}(C)$. But $\mathbf{x}+\mathbf{y} \in \rho^{-1}(C)$ is equivalent to say that $\rho(\mathbf{x}+\mathbf{y})$ is a codeword in C. Now

$$\begin{aligned} \rho(\mathbf{x}+\mathbf{y}) &= (x_1 + y_1 \pmod q, \ldots, x_n + y_n \pmod q) \\ &= \rho(\mathbf{x}) + \rho(\mathbf{y}). \end{aligned}$$

- Similarly, $\mathbf{0} \in \rho^{-1}(C) \in \mathbb{Z}^n$ and $-\mathbf{x} \in \rho^{-1}(C)$ whenever $\mathbf{x} \in \rho^{-1}(C)$.
- Prove the converse similarly. $\square$

Construction A

Let C be a linear code in $\mathbb{Z}_q^n$, the integers modulo a positive integer $q \geq 2$, where q is either prime or composite. Let $\rho : \mathbb{Z}^n \to \mathbb{Z}_q^n$ be the componentwise reduction modulo q. Then the lattice $\Lambda_C = \rho^{-1}(C)$ is said to have been obtained via *Construction A*.

- Sometimes $\Lambda_C = \frac{1}{\sqrt{q}}\rho^{-1}(C)$ is called Construction A.
- There are other constructions, Construction B,C,D...

q-ary lattices

The lattice Λ_C is also known as a q-ary lattice or modulo q lattice.

- Since $\mathbf{0} \in C$, $q\mathbf{e}_i \in \Lambda_C$, for all canonical vectors $\mathbf{e}_i$ and hence we have that $q\mathbb{Z}^n$ is a sublattice of Λ_C and the lattice inclusions
$$q\mathbb{Z}^n \subset \Lambda_C \subset \mathbb{Z}_q^n.$$
- On the other hand any lattice Λ in $\mathbb{R}^n$ satisfying $q\mathbb{Z}^n \subset \Lambda \subset \mathbb{Z}_q^n$ is obtained from the code $C = \rho(\Lambda)$ via Construction A.

q-ary lattices are used in lattice-based cryptography.

Property 1. If Λ_C is the q-ary lattice associated to the code $C \subseteq \mathbb{Z}_q^n$, then: $\left|\frac{\Lambda_C}{q\mathbb{Z}^n}\right| = \frac{q^n}{v(\Lambda_C)} = |C|$, where $|C|$ is the number of codewords of C and $v(\Lambda_C)$ is the volume of Λ_C, defined by the positive square root of the determinant of a Gram matrix.

Proof. That $|C| = \left|\frac{\Lambda_C}{q\mathbb{Z}^n}\right|$ follows from the isomorphism between $\Lambda_C/q\mathbb{Z}^n$ and C.
Since $q\mathbb{Z}^n$ is a sublattice of Λ_C, we have that a generator matrix M' of $q\mathbb{Z}^n$ is of the form $M' = AM$ for M a generator matrix of Λ_C and A an integral matrix, so

$$\left|\frac{\Lambda_C}{q\mathbb{Z}^n}\right| = |\det(A)| = \frac{|\det(M')|}{|\det(M)|}.$$

□

Property 1. If Λ_C is the q-ary lattice associated to the code $C \subseteq \mathbb{Z}_q^n$, then: $\left|\frac{\Lambda_C}{q\mathbb{Z}^n}\right| = \frac{q^n}{v(\Lambda_C)} = |C|$, where $|C|$ is the number of codewords of C and $v(\Lambda_C)$ is the volume of Λ_C, defined by the positive square root of the determinant of a Gram matrix.

- If q is prime, a code C is a subspace of dimension $k \leq n$ of $\mathbb{Z}_q^n = \mathbb{F}_q^n$ and hence has q^k codewords. From Property 1, we have that $v(\Lambda_C) = q^{n-k}$.

Property 2. Any full rank ($m = n$) integer lattice $\Lambda \subseteq \mathbb{Z}^n$ is q-ary for $q = v(\Lambda)$.

Proof. Since $\Lambda \subset \mathbb{Z}^n$, it follows that its volume $v(\Lambda) \in \mathbb{Z}$. Taking a generator matrix M for Λ and $q = v(\Lambda) = |det(M)|$, the linear system $M\mathbf{x} = q\mathbf{z}$ has an integer solution for any $\mathbf{z} \in \mathbb{Z}^n$. Indeed, $M^{-1} = \det(M)^{-1}\tilde{M}$ ($\tilde{M}$ contains the minors) and

$$\mathbf{x} = q \det(M)^{-1}\tilde{M}\mathbf{z} \in \mathbb{Z}$$

so every lattice point $q\mathbf{z}$ of $q\mathbb{Z}^n$ is a lattice point of Λ, therefore $q\mathbb{Z}^n \subset \Lambda$ (Λ is a q-ary lattice).

□

Generator matrices

- If p is prime, the linear (n, k) code C over $\mathbb{Z}_p = \mathbb{F}_p$ has a basis, formed by k vectors, which are placed in a generator matrix. Up to coordinate permutation, any code has a generator matrix in the reduced systematic form

$$\begin{bmatrix} \mathbf{I}_k & A \end{bmatrix}$$

  where $\mathbf{I}_k$ is the k-dimensional identity matrix.
- What if q is not a prime?

Consider the linear codes

$$C_1 = \{(2a, 2b, a+b),\ a, b \in \mathbb{F}_3\},\ C_2 = \{(2a, 2b, a+b),\ a, b \in \mathbb{Z}_4\}.$$

- The code C_1 over $\mathbb{F}_3$ has length $n = 3$, and generator matrices

$$M = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix},\ R = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}.$$

- The code C_2 over $\mathbb{Z}_4$ has length $n = 3$, and contains 8 codewords:

  $(0,0,0), (2,0,1), (0,0,2), (2,0,3), (0,2,1), (2,2,2), (0,2,3), (2,2,0).$

  The above matrix M is again a generator matrix for C_2.

In $\mathbb{Z}_4$, 2 is not invertible, so no systematic form, and the vectors $(2,0,1)$ and $(0,2,1)$ do not form a basis, because

$$\lambda_1(2,0,1) + \lambda_2(0,2,1) = 0$$

does not imply $\lambda_1 = \lambda_2 = 0$ since it could also be $\lambda_1 = \lambda_2 = 2$.

For C a linear code:

- Each codeword $\mathbf{a} \in C$ can be written using a set of generators, say $\mathbf{a} = \sum_{i=1}^{l} a_i \mathbf{v}_i$, $\mathbf{v}_i = (v_{i1}, \ldots, v_{in})$ for $i = 1, \ldots, l$ (and $l = k$ for the case of a linear (n, k) code over $\mathbb{F}_p$). Now

$$\mathbf{a} = \sum_{i=1}^{l} a_i \mathbf{v}_i \in C \iff \rho^{-1}(\mathbf{a}) = \sum_{i=1}^{l} a_i \mathbf{v}_i + \sum_{i=1}^{n} q h_i \mathbf{e}_i \in \mathbb{R}^n$$

  where $0 \leq a_i, v_{ij} \leq q - 1$ for all i, j, $\mathbf{e}_i$, $i = 1, \ldots, n$ form the canonical basis of $\mathbb{R}^n$ and $h_1, \ldots, h_n \in \mathbb{Z}$. In words, $\rho^{-1}(\mathbf{a})$ is an integral linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_l, q\mathbf{e}_1, \ldots, q\mathbf{e}_n$.

For $q = p$ prime, and

$$\begin{bmatrix} \mathbf{I}_k & A \end{bmatrix},$$

the generator matrix of C in systematic form, then a generator matrix of Λ_C is

$$\begin{bmatrix} \mathbf{I}_k & A \\ \mathbf{0} & q\mathbf{I}_{n-k} \end{bmatrix}.$$

- If q is not a prime, use the Hermite normal norm (HNF) instead.

For $C_1 = \{(2a, 2b, a + b),\ a, b \in \mathbb{F}_3\}$, with generator matrix

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

a generator matrix for Λ_{C_1} is

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}.$$

For $n \geq 2$, consider the linear code $C = \{(a_1, \ldots, a_{n-1}, \sum_{i=1}^{n-1} a_i),\ a_1, \ldots, a_{n-1} \in \mathbb{F}_2\}$ over $\mathbb{F}_2$. It has length n and dimension $n-1$. A systematic generator is

$$\begin{bmatrix} \mathbf{I}_{n-1} & \mathbf{1} \end{bmatrix}.$$

A generator matrix for Λ_C is thus

$$\begin{bmatrix} \mathbf{I}_{n-1} & \mathbf{1} \\ \mathbf{0} & 2 \end{bmatrix}.$$

Exercise. We have just constructed the lattice

$$D_n = \{(x_1, \ldots, x_n),\ \sum_{i=1}^{n} x_i \text{ is even}\}.$$

The Hamming distance

The *Hamming distance* d_H counts the number of coordinates in which two codewords differ. For $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$,

$$d_H(\mathbf{x}, \mathbf{y}) = |\{i; x_i \neq y_i\}|.$$

For example, in $\mathbb{F}_2^4$, $d_H((1,0,1,1),(0,1,0,1)) = 3$ and in $\mathbb{F}_5^3$, $d_H((1,0,3),(1,2,0)) = 2$.

- The minimum Hamming distance of a linear code C in $\mathbb{Z}_q^n$ is the minimum of all distances between two different vectors, which simplifies to the minimum Hamming weight w_H of a nonzero codeword.

Can we relate

- the minimum distance $d_H(C)$ of a linear binary code C,
- the shortest norm $\lambda = \min_{\mathbf{x} \in \Lambda_C, \mathbf{x} \neq \mathbf{0}} ||\mathbf{x}||$ of Λ_C?

Let C be a linear binary code and $\Lambda_C = \rho^{-1}(C)$. Then $2\mathbb{Z}^n \subseteq \Lambda_C$, so $(2, 0, \ldots, 0) \in \Lambda_C$ with norm 2.

- If we scale the lattice by a, i.e., $\Lambda_C = a\rho^{-1}(C)$, then $(2a, 0, \ldots, 0) \in \Lambda_C$ with norm $2a$.

In $\Lambda_C = a\rho^{-1}(C)$ for C a linear binary code and a a scaling factor, we have:

- $(2a, 0, \ldots, 0) \in \Lambda_C$ with norm $2a$,
- for $\mathbf{c} \in C$, $\mathbf{c} \in \mathbb{R}^n$ has the smallest norm among all $\mathbf{v}$ such that $\rho(\mathbf{v}) = \mathbf{c}$, and $||a\mathbf{c}|| = a||\mathbf{c}|| = a\sqrt{w_H(\mathbf{c})}$.

In order to have $a\sqrt{w_H(\mathbf{c})} \leq 2a$, we need $w_H(\mathbf{c}) \leq 4$.

Let C be a linear binary code with minimum distance $d_H(C)$, $\Lambda_C = \rho^{-1}(C)$, and $\lambda = \min_{\mathbf{x}\in\Lambda_C, \mathbf{x}\neq\mathbf{0}} ||\mathbf{x}||$. Then:

i) If $d_H(C) < 4$, $\lambda = \sqrt{d}$ and the minimum norm vectors of Λ_C are the codewords of C with weight d, as well as the vectors obtained by replacing one or more codeword coordinates set to 1 by -1.

ii) If $d_H(C) = 4$, $\lambda = 2$ and the minimum norm vectors of Λ_C are the codewords of C with weight equal to 4 as well as the vectors obtained by replacing one or more codeword coordinates set to 1 by -1.

iii) If $d_H(C) > 4$, $\lambda = 2$ and the minimum norm vectors of Λ_C are the ones which have ± 2 for their unique nonzero coordinate(s).

Exercise. Show that the lattice E_8 can be obtained via Construction A from the extended Hamming code in $\mathbb{Z}_2^8$ given by

$$C = \{(a_1, a_2, a_3, a_4, a_2+a_3+a_4, a_1+a_3+a_4, a_1+a_2+a_4, a_1+a_2+a_3), a_i \in \mathbb{Z}_2\}$$

The Lee Distance

We consider here the set of integers modulo q in its typical representation, $\mathbb{Z}_q = \{0, 1, \ldots, q-1\}$. For a and b in $\mathbb{Z}_q$, their Lee distance is defined by

$$d_{Lee}(a, b) = min\{|a - b|, q - |a - b|\}.$$

In $\mathbb{Z}_q^n$, the Lee distance between $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{b} = (b_1, b_2, \ldots, b_n)$ is defined as

$$d_{Lee}(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} d_{Lee}(a_i, b_i).$$

For a and b in $\mathbb{Z}_q$,

$$d_{Lee}(a,b) = min\{|a-b|, q-|a-b|\}.$$

- In $\mathbb{Z}_{13}$, $d_{Lee}(1,4) = \min\{3,10\} = 3$ and $d_{Lee}(1,12) = \min\{11,2\} = 2$.
- If $q = 2$, either $a = b$ and $d_{Lee}(a,b) = min\{0, q-|a-b|\} = 0$ or $a \neq b$, in which case $d_{Lee}(a,b) = min\{1, 2-|a-b|\} = 1$, which matches the Hamming distance.

Back to $\mathbb{F}_q$. For speaking of duality, it is often preferred to work with $\Lambda_C = \frac{1}{\sqrt{q}}\rho^{-1}(C)$ (we will see why next).

$C^\perp = \{\mathbf{v} \in \mathbb{F}_q^n,\ \mathbf{c}\cdot\mathbf{v} = \mathbf{0}$ for all $\mathbf{c} \in C\}$	$L^* = \{\mathbf{y} \in \mathbb{R}^n,\ \mathbf{y}\cdot\mathbf{x} \in \mathbb{Z}$ for all $\mathbf{x} \in L\}$

We have $\frac{1}{\sqrt{q}}\Lambda_{C^\perp} = (\frac{1}{\sqrt{q}}\Lambda_C)^*$, or equivalently $\frac{1}{q}\Lambda_{C^\perp} = \Lambda_C^*$.

Proof. We show an inclusion first.

- For $(x_1, x_2, \ldots, x_n) \in \frac{1}{\sqrt{q}}\Lambda_{C^\perp}$, each x_i is of the form $\frac{x_i'}{\sqrt{q}}$, so there is $(a_1, a_2, \ldots, a_n) \in C^\perp$ such that $x_i' \equiv a_i \pmod{q}$, that is $\sqrt{q}x_i \equiv a_i \pmod{q}$.
- For $(y_1, y_2, \ldots, y_n) \in \frac{1}{\sqrt{q}}\Lambda_C$, there is $(b_1, b_2, \ldots, b_n) \in C$ such that $\sqrt{q}y_i \equiv b_i \pmod{q}$.
- Thus $q\sum_i x_i y_i \equiv \sum_i a_i b_i \pmod{q} \equiv 0 \pmod{q}$ and we have $\frac{1}{\sqrt{q}}\Lambda_{C^\perp} \subseteq (\frac{1}{\sqrt{q}}\Lambda_C)^*$.

We know $\frac{1}{\sqrt{q}}\Lambda_{C^\perp} \subseteq (\frac{1}{\sqrt{q}}\Lambda_C)^*$.
To prove the reverse inclusion, it is thus enough to show that $\det(\frac{1}{\sqrt{q}}\Lambda_{C^\perp}) = \det((\frac{1}{\sqrt{q}}\Lambda_C)^*)$.

- By Property 1, $\det(\frac{1}{\sqrt{q}}\Lambda_C) = \frac{q^n}{|C|^2}$ so $\det((\frac{1}{\sqrt{q}}\Lambda_C)^*) = \frac{|C|^2}{q^n}$.
- We also have $\det(\frac{1}{\sqrt{q}}\Lambda_{C^\perp}) = \frac{q^n}{|C^\perp|^2}$ and $|C||C^\perp| = q^n$, so $\det(\frac{1}{\sqrt{q}}\Lambda_{C^\perp}) = \frac{q^n}{|C^\perp|^2} = \frac{q^n}{(q^n/|C|)^2} = \frac{|C|^2}{q^n} = \det((\frac{1}{\sqrt{q}}\Lambda_C)^*)$.

This concludes the proof, and we have $\frac{1}{\sqrt{q}}\Lambda_{C^\perp} = (\frac{1}{\sqrt{q}}\Lambda_C)^*$.

We have $\frac{1}{\sqrt{q}}\Lambda_{C^\perp} = (\frac{1}{\sqrt{q}}\Lambda_C)^*$, or equivalently $\frac{1}{q}\Lambda_{C^\perp} = \Lambda_C^*$.

The equivalent statement follows from remembering that $(\frac{1}{\sqrt{q}}\Lambda_C)^* = \sqrt{q}\Lambda_C^*$. $\square$

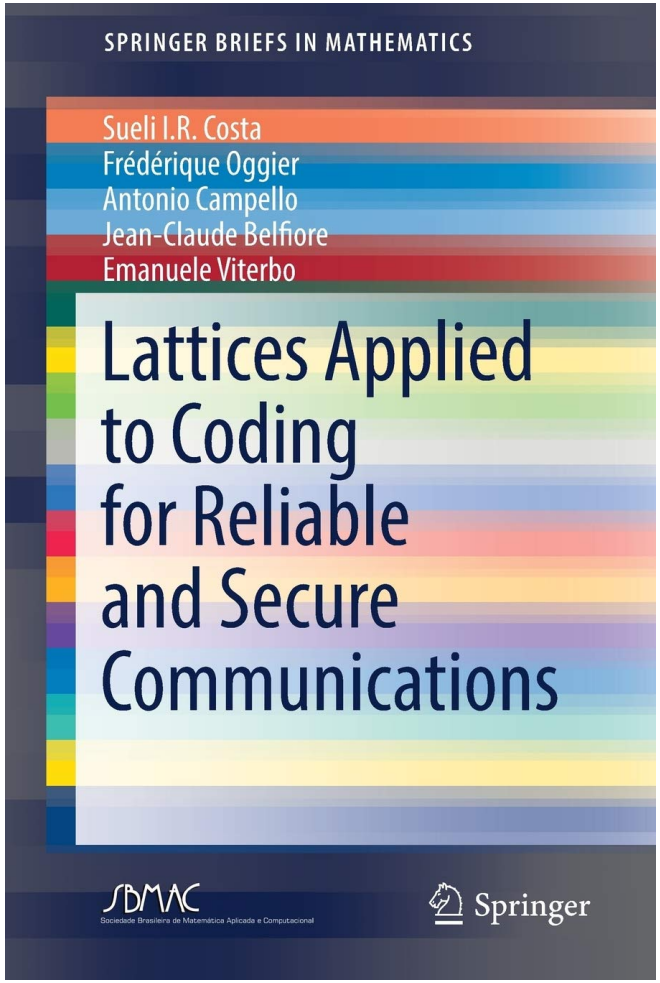
We have $\frac{1}{\sqrt{q}}\Lambda_{C^\perp} = (\frac{1}{\sqrt{q}}\Lambda_C)^*$.
Then $C = C^\perp$ if and only if $\frac{1}{\sqrt{q}}\Lambda_C = (\frac{1}{\sqrt{q}}\Lambda_C)^*$ if and only if $\frac{1}{\sqrt{q}}\Lambda_C$ is unimodular.
Normalizing gives a determinant of 1.

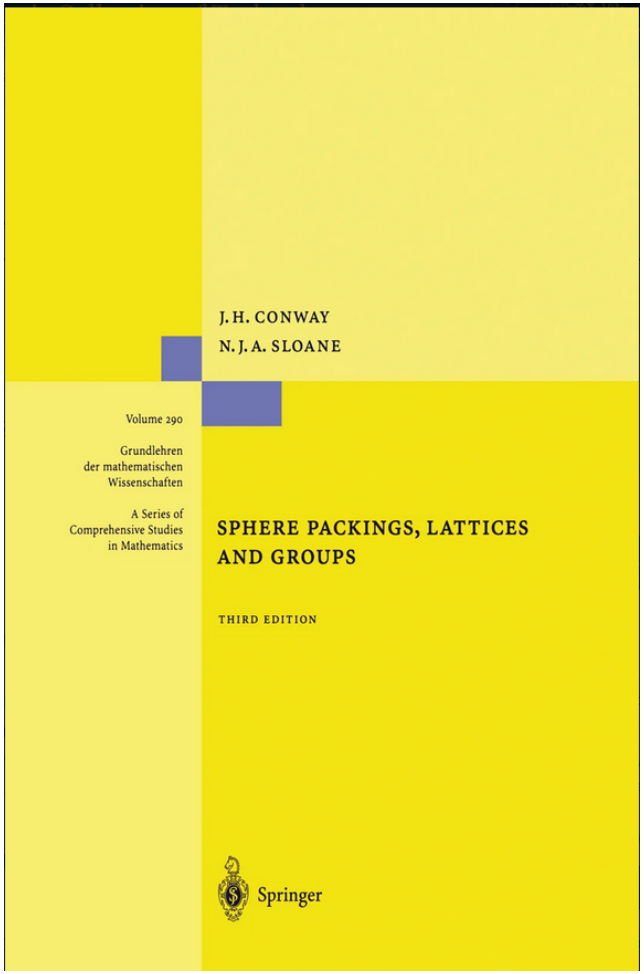
More ...

- weight enumerator for codes $\Longleftrightarrow$ theta series for lattices

Construction A is of interest:

- to mathematicians looking for "good" (e.g., unimodular, l-modular, extremal) lattices,
- to cryptographers (via q-ary lattices),
- to information theorists (for capacity achieving codes),
- to coding theorists (decoding).


SPRINGER BRIEFS IN MATHEMATICS
Sueli I.R. Costa
Frédérique Oggier
Antonio Campello
Jean-Claude Belfiore
Emanuele Viterbo
Lattices Applied to Coding for Reliable and Secure Communications
SBMAC
Sociedade Brasileira de Matemática Aplicada e Computacional
Springer


J. H. CONWAY
N. J. A. SLOANE
Volume 290
Grundlehren der mathematischen Wissenschaften
A Series of Comprehensive Studies in Mathematics
SPHERE PACKINGS, LATTICES AND GROUPS
THIRD EDITION
Springer