Error-Correcting Codes
by
W. Wesley Peterson & E.J. Weldon, Jr.

Chapter 2. Introduction to Algebra

2 Introduction to Algebra

Structure is desirable in error-correcting codes for two reasons: It facilitates finding properties of a code, and even more important, it makes instrumentation of the codes practical. Algebraic structure has been the basis of the most important known codes.

This chapter consists of two parts. The first defines the most significant algebraic structures and gives a few examples of each. The rest of the chapter reviews some of the theory of vector spaces and matrices. Chapter 6 is also purely mathematical, dealing with the theory of rings and finite fields. These two chapters are in no sense complete mathematical presentations but rather barely minimum mathematical prerequisites for the discussion of codes.

Algebraic systems are systems that satisfy certain rules or laws, and for the most part, these are the same laws as apply to our ordinary number system. Thus a *group* is a system with one operation and its inverse, such as addition and its inverse, subtraction, or multiplication and its inverse, division. A *ring* has two operations, addition and multiplication, and the inverse operation, subtraction, for the first. A *field* has the two operations, both with inverses.

2.1 Groups

A group G is a set of objects, or elements, for which an operation is defined and for which Axioms G.1 to G.4 hold. Let $a, b, c, \ldots$ be elements of the group. The operation is a single-valued function of two variables, and might well be denoted $f(a, b) = c$ but is customarily denoted $a + b = c$ or $ab = c$ and called addition or multiplication, even though it may not be the addition or multiplication of the arithmetic of ordinary numbers.

> Axiom G.1. (*Closure*). *The operation can be applied to any two group elements to give a third group element as a result.*

> Axiom G.2. (*Associative Law*). *For any three elements a, b, and c of the group, $(a + b) + c = a + (b + c)$ if the operation is written as addition, or $a(bc) = (ab)c$ if the operation is written as multiplication.*

The associative law means that the order of performing operations is immaterial, and so parentheses are unnecessary.

> Axiom G.3. *There is an identity element.*

If the operation is called addition, the identity element is called zero and written 0 and is defined by the equation $0 + a = a + 0 = a$ for every element of a of the group. If the operation is called multiplication, the identity is one, written 1, and is defined by the equation $1a = a1 = a$.

> Axiom G.4. *Every element of the group has an inverse element.*

If the operation is addition, the inverse element corresponding to a is denoted $-a$ and is defined by the equation $a + (-a) = (-a) + a = 0$. If the operation is multiplication, the inverse of a is denoted a^{-1} and is defined by the equation $aa^{-1} = a^{-1}a = 1$.

In addition to the above laws, a group may satisfy the commutative law; that is, $a + b = b + a$, or if the operation is multiplication, $ab = ba$. Such a group is called *Abelian* or *commutative*.

In developing a general theory of groups, the multiplicative notation is used in this book.

> Theorem 2.1. *The identity element in a group is unique, and the inverse of each group element is unique.*

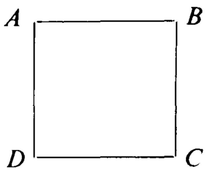
Proof. The identity element is unique, for if there were two identity elements, 1 and 1′, $(1)(1') = 1 = 1'$. Similarly, inverses are unique, for if a group element g were to have two inverses g^{-1} and g_1^{-1}, then $g^{-1} = 1g^{-1} = g_1^{-1}gg^{-1} = g_1^{-1}\,1 = g_1^{-1}$, so that they must be equal. Q.E.D.

Note that the inverse of a product is the product of inverses *in* *reverse order*, for $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a1a^{-1} = aa^{-1} = 1$, and therefore $b^{-1}a^{-1} = (ab)^{-1}$.

Examples. The set of all real numbers is a group under the operation of ordinary addition. The set of all positive and negative integers and zero is also a group under addition. The set of all real numbers excluding zero is a group under the operation of ordinary multiplication. All these groups are Abelian. The set of all nonsingular $n \times n$ matrices is a non-Abelian group, under the operation matrix multiplication.

Many important groups are sets of transformations of some space, with the operation called multiplication, defined as follows: The transformation ab is the result of performing the transformation b followed by the transformation a. For example, the set of rotations of n-dimensional Euclidean space is a group. Note that the rotations of two-dimensional space form an Abelian group, while the rotations of three dimensional space are not commutative.

As a first example of a finite group, consider a transformation of a plane which maps a square onto itself. A transformation is completely determined if its effect on the four vertices is specified.



For example, one possible mapping is a 90° counterclockwise rotation of the square, which maps A onto D, B onto A, C onto B, and D onto C. It can be described in the notation sometimes used for permutations:

$$\begin{pmatrix} ABCD \\ DABC \end{pmatrix}.$$

There are eight such transformations in all:

$$1 = \begin{pmatrix} ABCD \\ ABCD \end{pmatrix}, \quad a = \begin{pmatrix} ABCD \\ DABC \end{pmatrix}, \quad b = \begin{pmatrix} ABCD \\ CDAB \end{pmatrix}, \quad c = \begin{pmatrix} ABCD \\ BCDA \end{pmatrix},$$

$$d = \begin{pmatrix} ABCD \\ BADC \end{pmatrix}, \quad e = \begin{pmatrix} ABCD \\ ADCB \end{pmatrix}, \quad f = \begin{pmatrix} ABCD \\ DCBA \end{pmatrix}, \quad g = \begin{pmatrix} ABCD \\ CBAD \end{pmatrix}.$$

The multiplication table is

	1	a	b	c	d	e	f	g
1	1	a	b	c	d	e	f	g
a	a	b	c	1	e	f	g	d
b	b	c	1	a	f	g	d	e
c	c	1	a	b	g	d	e	f
d	d	g	f	e	1	c	b	a
e	e	d	g	f	a	1	c	b
f	f	e	d	g	b	a	1	c
g	g	f	e	d	c	b	a	1

The fact that each element has an inverse can be easily seen from the multiplication table. Although the associative law could be verified from the multiplication table, this would be a very tedious job, but it should be clear from the definition of the group that the associative law does hold.

There is a group with only one element. That element must be the identity element by Axiom G.3, and it is easy to verify that the other axioms hold. There is also a group with two elements. One must be the identity element, 0. Let us call the other element a. Then a must have an inverse, and since $a + 0 = a \neq 0$, $-a \neq 0$, so $-a = a$. Thus the addition table *must be* $0 + 0 = 0$, $0 + a = a + 0 = a$, $a + a = 0$, and a set of two elements with addition defined in this way satisfies all the axioms G.1 to G.4. In fact, the only group with two elements is also Abelian.

2.2 Rings

A *ring R* is a set of elements for which two operations are defined. One is called addition and denoted $a + b$, and the other is called multiplication and denoted ab, even though these operations may not be ordinary addition or multiplication of numbers. In order for R to be a ring, the following axioms must be satisfied:

Axiom R.1. *The set R is an Abelian group under addition.*

Axiom R.2. (*Closure*). *For any two elements a and b of R, the product ab is defined and is an element of R.*

AXIOM R.3. (*Associative Law*). *For any three elements a, b, and c of R, $a(bc) = (ab)c$.*

AXIOM R.4. (*Distributive Law*). *For any three elements a, b, and c of R, $a(b + c) = ab + ac$ and $(b + c)a = ba + ca$.*

A ring is called *commutative* if its multiplication operation is commutative; that is, if for any two elements a and b, $ab = ba$.

THEOREM 2.2. *In any ring, for any elements a and b, $a0 = 0a = 0$ and $a(-b) = (-a)b = -(ab)$.*

Proof. In any ring, by Axiom R.4, for any a, $a(0 + 0) = a0 + a0$. But since $0 + 0 = 0$, $a0 = a0 + a0$. Next $a0$ must have an additive inverse, and adding this to both sides gives $0 = a0 + (-a0) = a0 + a0 + (-a0) = a0 + 0 = a0$, so in any ring $a0 = 0$. Similarly $0a = 0$. Then $0 = a0 = a(b + (-b)) = ab + a(-b)$, so $a(-b) = -(ab)$. Similarly $(-a)b = -(ab)$. Q.E.D.

Examples. The set of all real numbers is a ring under the operations of ordinary addition and multiplication. The set of all positive and negative integers and zero is also a ring under ordinary addition and multiplication. Both these rings are commutative. The set of all $n \times n$ matrices with either integer or real-number elements is a ring under the operations matrix addition and matrix multiplication, and this ring is noncommutative. The set of all polynomials in one indeterminant, or variable, with integer coefficients is a commutative ring.

A set consisting of a zero element only is a ring, with the rules $0 + 0 = 0$, $(0)(0) = 0$. There are two different rings with two elements. One element must be the additive identity 0. The other element a must satisfy $a + a = 0$. Since $(0)(0) = 0a = a0 = 0$ by Theorem 2.2, the only question is, what is the value of aa? It turns out that either $aa = a$ or $aa = 0$ satisfies both the distributive and associative laws, and thus either choice gives a ring, and clearly these two choices give rings of different structure.

2.3 Fields

A *field* is a commutative ring with a unit element (multiplicative identity) in which every nonzero element has a multiplicative inverse.

A noncommutative ring in which every nonzero element has an inverse is usually called a *division ring* or a *skew field.*

Note that the nonzero elements of a field satisfy all the axioms for a group and thus form a group under the operation multiplication.

Examples. The set of all real numbers form, a field, as do also the set of all rational numbers and the set of all complex numbers.

The minimum number of elements a field can have is two, for it must have both an additive identity 0 and a multiplicative identity 1. They have to satisfy the addition and multiplication tables given in Table 2.1, for there is only one possible addition table for a group of two elements. Also, it was shown that for rings in general, $0a = 0$ for any a, and, since 1 is a unit element, $(1)(1) = 1$. It can be verified easily that the set 0 and 1 with the operations defined earlier satisfy all the axioms for a field.

Table 2.1. Addition and Multiplication Tables for the Field with Two Elements

+	0	1
0	0	1
1	1	0

.	0	1
0	0	0
1	0	1

It can be shown that for every number q that is a power of a prime number there is a field with q elements. The proof of this fits in better with the material of Chapter 6 and is presented there. However, it might be well to point out here that a field with p elements can be formed by taking the integers modulo p, provided p is a prime. *The integers modulo q do not form field if q is not a prime and the fields with $q = p^m$ elements $(m > 1)$ are not formed by taking integers modulo q.* For use in examples, addition and multiplication tables for fields with three and four elements are given in Tables 2.2 and 2.3. The field of four elements described in Table 2.3 is *not* the integers modulo 4.

Table 2.2. Addition and Multiplication Table for the Field with Three Elements

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

.	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Table 2.3. Addition and Multiplication Table for the Field with Four Elements

+	0	1	*a*	*b*
0	0	1	*a*	*b*
1	1	0	*b*	*a*
a	*a*	*b*	0	1
b	*b*	*a*	1	0

.	0	1	*a*	*b*
0	0	0	0	0
1	0	1	*a*	*b*
a	0	*a*	*b*	1
b	0	*b*	1	*a*

2.4 Subgroups and Factor Groups

A subset of elements of a group G is called a *subgroup* H if it satisfies all the axioms for the group itself with the same operation. To determine whether H is a subgroup, it is necessary to check only for closure (that is, if a and b are in H, then ab must be in H) and for inverses (that is, if a is in H, then a^{-1} must be also). If a set is closed under the group operation and the inverse is present, the identity must be present also, and the associative law must hold in the subgroup if it does in the group.

> *Example.* In the group of eight transformations of a square given previously, the sets $(1, a, b, c)$ and $(1, d)$ are both subgroups.
>
> In the group of all integers, the set of all integers that are even multiples of a given integer m is a subgroup for every m.

Suppose that the elements of a group G are $g_1, g_2, g_3, \ldots$, and the elements of a subgroup H are $h_1, h_2, h_3, \ldots$, and consider the array formed as follows: The first row is the subgroup, with the identity at the left and each other element appearing once and only once. The first element in the second row is any element not appearing in the first row, and the rest of the elements are obtained by multiplying each subgroup element by this first element on the left. Similarly a third, fourth, and fifth row are formed, each with a previously unused group element in the first column, until all the group elements appear somewhere in the array.

$$\begin{array}{llllll}
h_1 = 1, & h_2, & h_3, & h_4, & \ldots, & h_n \\
g_1 h_1 = g_1, & g_1 h_2, & g_1 h_3, & g_1 h_4, & \ldots, & g_1 h_n \\
g_2 h_1 = g_2, & g_2 h_2, & g_2 h_3, & g_2 h_4, & \ldots, & g_2 h_n \\
\vdots & \vdots & \vdots & \vdots & & \vdots \\
g_m h_1 = g_m, & g_m h_2, & g_m h_3, & g_m h_4, & \ldots, & g_m h_n
\end{array}$$

The set of elements in a row of this array is called a left *coset*, and the element appearing in the first column is called the *coset leader*. Right cosets could be similarly formed. The array itself is known as the *coset decomposition* of the group.

THEOREM 2.3. *Two elements g and g' of a group G are in the same left coset of a subgroup H of G if and only if $g^{-1}g'$ is an element of H.*

Proof. If g and g' belong to the coset whose leader is g_i, then $g = g_i h_j$ for some j, $g' = g_i h_k$ for some k, and $g^{-1}g' = (g_i h_j)^{-1} \cdot g_i h_k = h_j^{-1} g_i^{-1} g_i h_k = h_j^{-1} h_k$, which is in the subgroup. On the other hand, if $g = g_i h$, where g_i is the coset leader, and if $g^{-1}g' = h'$, then $g' = gh' = g_i hh'$, which is in the same coset, since hh' is in the subgroup.

Q.E.D.

THEOREM 2.4 *Every element of the group G is in one and only one coset of a subgroup H.*

Proof. Every element appears at least once, by the definition of the construction of the array. It must be shown that each element appears only once in the array. Suppose first that two elements in the same row, $g_i h_j$ and $g_i h_k$, are equal.Then multiplying each on the left by g_i^{-1} would give $h_j = h_k$, a result that is a contradiction, since each subgroup element was assumed to appear only once in the first row. Now suppose that two equal elements appear in different rows, $g_i h_j = g_k h_l$, and suppose that $i > k$. Then multiplying on the right by h_j^{-1} gives $g_i = g_k h_l h_j^{-1}$. Since $h_l h_j^{-1}$ is in the subgroup, this indicates that g_i is in the kth coset, a situation that contradicts the rule of construction that coset leaders should be previously unused. Q.E.D.

The number of elements in a group is called the *order* of the group. The number of cosets of G with respect to a subgroup H is called the *index* of G over H. Clearly,

(Order of H) (index of G over H) = (order of G).

A subgroup H of a group G is called *normal* if, for any element h of H and any element g of G, $g^{-1}hg$ is in H. In general, left cosets may not be right cosets, and vice versa. However, every left coset of a normal subgroup is also a right coset, and vice versa. In an Abelian group, every left coset is trivially a right coset, and also all subgroups are trivially normal. In this book the only use made of normal subgroups will be for Abelian groups, and therefore the foregoing result will not be proved in general.

If a subgroup H of a group G is normal, it is possible to define an operation on cosets to form a new group for which the cosets are the elements. This group is called the *factor group* and denoted G/H. The coset containing g is denoted $\{g\}$. The definition of multiplication for cosets is

$$\{g_1\}\{g_2\} = \{g_1 g_2\}.$$

This is not a valid definition unless it happens that, no matter which element is chosen as a representive of each of the two cosets to be multiplied, the resulting coset is the same. In other words, it must be shown that if g_1 and g_1' are in the same coset, and g_2 and g_2' are in the same coset, then $g_1 g_2$ and $g_1' g_2'$ are also in the same coset. Assume that $g_1^{-1}g_1' = h_1$, $g_2^{-1}g_2' = h_2$, and then, since the subgroup is normal, $g_2'^{-1}h_1 g_2'$ must be an element of H, say h_3. Hence $(g_1 g_2)^{-1}g_1' g_2' = g_2^{-1}g_1^{-1}g_1' g_2' = g_2^{-1}h_1 g_2' = g_2^{-1}g_2' h_3 = h_2 h_3$, which is an element of H. Therefore $g_1 g_2$ and $g_1' g_2'$ are in the same coset, and the definition is consistent.

Now let us check that G/H is actually a group. The operation is clearly defined for all pairs of cosets, and therefore Axiom G.1 is satisfied. To check the associative law, note that

$$\begin{aligned}\{g_1\}(\{g_2\}\{g_3\}) = \{g_1\}\{g_2 g_3\} = \{g_1 g_2 g_3\} &= \{g_1 g_2\}\{g_3\}\\ &= (\{g_1\}\{g_2\})\{g_3\}.\end{aligned}$$

The identity element is the subgroup itself, $H = \{1\}$, since $\{1\}\{g\} = \{1g\} = \{g\}$ and $\{g\}\{1\} = \{g1\} = \{g\}$. Similarly the inverse coset of $\{g\}$ is the coset containing g^{-1}, $\{g^{-1}\}$, since $\{g\}\{g^{-1}\} = \{gg^{-1}\} = \{1\}$ and $\{g^{-1}\} \cdot \{g\} = \{g^{-1}g\} = \{1\}$. Also if the original group is Abelian, it is easily verified that the factor group is also.

Examples. Suppose that the group G is the group of eight transformations of the square, and H is the subgroup consisting of 1, a, b, c. Then the standard array of left cosets is, if d is chosen as the coset leader,

1	a	b	c
d	$da = g$	$db = f$	$dc = e$

There is only one coset consisting of all the elements of G except those in H, and so this must also be a right coset, and H must be normal. If the identity coset is called I and the other one D, then the multiplication table is $II = I$, $ID = \{1\}\{d\} = d = D$, $DI = D$, $DD = \{d\}\{d\} = \{dd\} = \{1\} = I$. This, of course, has the same structure as the only group of two elements.

As a more important example, let G be the group of all positive and negative integers and zero under addition, and let H be the subgroup that consists of all multiples of an integer n. All the numbers from zero to $n - 1$ inclusive are in different cosets, since for two elements a and b to be in the same coset, $(-a) + b$ must be in the subgroup and thus be a multiple of n. These can be taken as coset leaders, and it is easily seen that there are no other cosets. Since G is Abelian, addition of cosets can be defined, and the cosets form a group. For example, let $n = 3$. Then the cosets are

$$\begin{array}{llllllll} 0, & 3, & -3, & 6, & -6, & 9, & -9, & \ldots \\ 1, & 4, & -2, & 7, & -5, & 10, & -8, & \ldots \\ 2, & 5, & -1, & 8, & -4, & 11, & -7, & \ldots \end{array}$$

If these are called {0},{1}, and {2}, respectively, the addition table is

+	{0}	{1}	{2}
{0}	{0}	{1}	{2}
{1}	{1}	{2}	{0}
{2}	{2}	{0}	{1}

This may be recognized as addition modulo 3.

2.5 Vector Spaces and Linear Algebras

A set V of elements is called a *vector space* over a field F if it satisfies the following axioms:

AXIOM V.1. *The set V is an Abelian group under addition.*

AXIOM V.2. *For any vector* $\mathbf{v}$ *in V and any field element c, a product* $c\mathbf{v}$*, which is a vector in V, is defined. (Field elements are called scalars, elements of V are vectors.)*

AXIOM V.3. *(Distributive Law). If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in V and c is a scalar,* $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.

AXIOM V.4. *(Distributive Law). If* $\mathbf{v}$ *is a vector and c and d are scalars,* $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$.

AXIOM V.5. *(Associative Law). If* $\mathbf{v}$ *is a vector and c and d are scalars,* $(cd)\mathbf{v} = c(d\mathbf{v})$, *and* $1\mathbf{v} = \mathbf{v}$.

A set A of elements is called a *linear associative algebra* over a field F if it satisfies the following axioms:

AXIOM A.1. *The set A is a vector space over F.*

AXIOM A.2. *For any two elements* $\mathbf{u}$ *and* $\mathbf{v}$ *of A, there is a product* $\mathbf{uv}$ *defined that is in A.*

AXIOM A.3. (*Associative Law*) *For any three elements* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *of A*, $(\mathbf{uv})\mathbf{w} = \mathbf{u}(\mathbf{vw})$.

AXIOM A.4. (*Bilinear Law*) *If c and d are scalars in F and* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in A, then*

$\mathbf{u}(c\mathbf{v} + d\mathbf{w}) = c\mathbf{uv} + d\mathbf{uw}$ and $(c\mathbf{v} + d\mathbf{w})\mathbf{u} = c\mathbf{vu} + d\mathbf{wu}$.

An *n-tuple* over a field is an ordered set of n field elements and is denoted $(a_1, a_2, a_3, \dots, a_n)$, where each a_i is an element of the field. Addition of n-tuples is defined as follows:

$$(a_1, a_2, \dots, a_n) + (b_1, b_2, \dots, b_n) = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n).$$

Multiplication of an n-tuple by a field element is defined as follows:

$$c(a_1, a_2, \dots, a_n) = (ca_1, ca_2, \dots, ca_n).$$

With these two definitions it can be verified easily that the set of all n-tuples over a field form a vector space, and such vector spaces play a central role in coding theory. They are the subject of the remainder of this chapter.

Multiplication of n-tuples can be defined as follows:

$$(a_1, a_2, \dots, a_n)(b_1, b_2, \dots, b_n) = (a_1 b_1, a_2 b_2, \dots, a_n b_n);$$

with this definition the n-tuples form a linear algebra. This type of multiplication is occasionally useful. Another type of multiplication of n-tuples leading to a linear algebra is described in Chapter 6 and plays a more important role in coding theory.

The identity element of the vector space will be denoted $\mathbf{0}$; that is,

$$\mathbf{0} = (0, \dots, 0).$$

It is clearly true for n-tuples and in fact easily shown for vector spaces in general that for any vector $\mathbf{v}$, $0\mathbf{v} = \mathbf{0}$, and for any scalar c, $c\mathbf{0} = \mathbf{0}$. Also, $(-\mathbf{v}) = (-1)\mathbf{v}$, for $\mathbf{v} + (-1)\mathbf{v} = 1\mathbf{v} + (-1)\mathbf{v} = [1 + (-1)]\mathbf{v} = 0\mathbf{v} = \mathbf{0}$.

A subset of a vector space is called a *subspace* if it satisfies the axioms for a vector space. In order to check whether a subset of a vector space is a subspace, it is necessary only to check for closure under addition and multiplication by scalars. Note that, since $-\mathbf{v} = (-1)\mathbf{v}$, closure under multiplication by scalars assures that the inverse of each element is in the subspace. Then closure under addition is sufficient to ensure that it is a subgroup, and the associative and distributive laws must hold in the subspace if they hold in the original vector space.

A *linear combination* of k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a sum of the form

$$\mathbf{u} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k.$$

The a_i are scalars, that is, field elements.

THEOREM 2.5. *The set of all linear combinations of a set of vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ *of a vector space* V *is a subspace of* V.

Proof. Clearly every linear combination of vectors of V is also a vector of V. If the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is called S, and $\mathbf{w} = b_1\mathbf{v}_1 + \cdots + b_k\mathbf{v}_k$ and $\mathbf{u} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ are any two elements of S, then $\mathbf{w} + \mathbf{u}$ is also in S, for $\mathbf{w} + \mathbf{u} = (b_1 + c_1)\mathbf{v}_1 + \cdots + (b_k + c_k)\mathbf{v}_k$ is in S. Also any scalar multiple of $\mathbf{w}$, $a\mathbf{w} = ab_1\mathbf{v}_1 + \cdots + ab_k\mathbf{v}_k$ is in S. Since S is closed under addition and multiplication by scalars, S is a subspace of V. Q.E.D.

A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is *linearly dependent* if and only if there are scalars $c_1, \ldots., c_k$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

A set of vectors is *linearly independent* if it is not linearly dependent. A set of vectors is said to *span* a vector space if every vector in the vector space equals a linear combination of the vectors in the set.

THEOREM 2.6. *If a set of k vectors* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ *spans a vector space that contains a set of m linearly independent vectors* $\mathbf{u}_1, \ldots., \mathbf{u}_m$, *then* $k \geqq m$.

Proof. Since $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span the space, $\mathbf{u}_1$ can be expressed as a linear combination of the $\mathbf{v}_i$. Therefore, this equation can be solved for some one of the $\mathbf{v}_i$, say $\mathbf{v}_j$, in terms of $\mathbf{u}_1$ and the rest of the $\mathbf{v}_i$. Consequently, the set consisting of $\mathbf{u}_1$ and the rest of the $\mathbf{v}_i$ spans the vector space, since any linear combination of the $\mathbf{v}_i$ becomes a linear combination of $\mathbf{u}_1$ and all the $\mathbf{v}_i$ except $\mathbf{v}_j$ when the expression for $\mathbf{v}_j$ in terms of $\mathbf{u}_1$ and the other $\mathbf{v}_i$ is used to eliminate $\mathbf{v}_j$. Then $\mathbf{u}_2$ can be expressed as a linear combination of $\mathbf{u}_1$ and all the $\mathbf{v}_i$ except $\mathbf{v}_j$. Since the $\mathbf{u}_i$ are linearly independent, some $\mathbf{v}_i$ must have a nonzero coefficient, and therefore this $\mathbf{v}_i$ can be expressed in terms of $\mathbf{u}_1$, $\mathbf{u}_2$, and the remaining $k - 2$ vectors $\mathbf{v}_i$, and these k vectors span the space. The process can be continued until all m of the $\mathbf{u}_i$ vectors are used, and, since at each stage one $\mathbf{v}_i$ vector is replaced, the number of vectors $\mathbf{v}_i$ must have been at least as great as the number of vectors $\mathbf{u}_i$. Q.E.D.

THEOREM 2.7. *If two sets of linearly independent vectors span the same space, there are the same number of vectors in each set.*

Proof. If there are m vectors in one set and k in the other, then by Theorem 2.6, $m \geqq k$ and $k \geqq m$, and thus $m = k$. Q.E.D.

In any space, the number of linearly independent vectors that span the space is called the *dimension* of the space. A set of k linearly independent vectors spanning a k-dimensional vector space is called a *basis* of the space. It follows from Theorem 2.7 that every set of more than k vectors in a k-dimensional vector space is linearly dependent. It follows from Theorem 2.6 that no set of fewer than k vectors can span a k-dimensional space.

THEOREM 2.8. *If V is a k-dimensional vector space, any set of k linearly independent vectors in V is a basis for V.*

Proof. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be a set of linearly independent vectors in V. If they do not span V, there must be some vector $\mathbf{v}$ in V that is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Then the set $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ of $k + 1$ vectors in V is linearly independent. This contradicts Theorem 2.6, and therefore $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ must span V. Q.E.D.

THEOREM 2.9. *If a vector space V_1 is contained in a vector space V_2 and they have the same dimension k, they are equal.*

Proof. A basis for V_1 is a set of k linearly independent vectors in V_2. Therefore, every vector in V_2 is also in V_1. Q.E.D.

An *inner product* or *dot product* of two n-tuples is a scalar and is defined as follows:

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = a_1 b_1 + \cdots + a_n b_n.$$

It is easily verified that $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ and that $\mathbf{w} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{w} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{v}$. If the inner product of two vectors is zero, they are said to be *orthogonal*.

2.6 Matrices

The purpose of this section is to outline the parts of matrix theory that apply to the codes studied in the next three chapters. For the most part, proofs are given, but this can hardly serve as more than a review of the necessary parts of matrix theory.

An $n \times m$ matrix is an ordered set of nm elements in a rectangular array of n rows and m columns:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} = [a_{ij}].$$

The elements of a matrix may in general be elements of any ring, but in this book only matrices with elements in a field find application. The n rows may be thought of as n m-tuples or vectors, and similarly, the m columns may be thought of as vectors. The set of elements a_{ii} for which the column number and row number are equal is called the main diagonal.

The *row space* of an $n \times m$ matrix **M** is the set of all linear combinations of row vectors of **M**. They form a subspace of the vector space of m-tuples. The dimension of the row space is called the *row rank*. Similarly, the set of all linear combinations of column vectors of the matrix forms the *column space*, whose dimension is called the *column rank*. It can be shown that row rank equals column rank; this value is referred to as the *rank* of the matrix.

There is a set of *elementary row operations* defined for matrices:

1. Interchange of any two rows.
2. Multiplication of any row by a nonzero field element.
3. Addition of any multiple of one row to another.

The inverse of each elementary row operation is clearly an elementary row operation of the same kind.

THEOREM 2.10. *If one matrix is obtained from another by a succession of elementary operations, both matrices have the same row space.*

Proof. If the theorem is true for each elementary row operation, it will clearly be true for a succession. It is obviously true of row operations 1 and 2. Suppose that the matrix **M**′ is obtained from the matrix **M** by a type 3 elementary row operation. Then, since the altered row of **M**′ is a linear combination of two rows of **M**, any linear combination of rows of $\mathbf{M}'$ is also a linear combination of rows of $\mathbf{M}$, so the row space of $\mathbf{M}'$ is contained in the row space of $\mathbf{M}$. But $\mathbf{M}$ can be obtained from $\mathbf{M}'$ by the inverse operation, which is again an operation of type 3, so the row space of $\mathbf{M}$ must be contained in the row space of $\mathbf{M}'$. Therefore they are equal. Q.E.D.

Elementary row operations can be used to simplify a matrix and put it in a standard form. The form, called *echelon canonical form*, is as follows:

1. Every leading term of a nonzero row is 1.
2. Every column containing such a leading term has all its other entries zero.
3. The leading term of any row is to the right of the leading term in every preceding row. All zero rows are below all nonzero rows.

The procedure is essentially the same as that used in solving linear equations by elimination of one variable at a time. It is best illustrated by an example. Consider the following matrix with real numbers as elements:

$$\begin{bmatrix} 0&0&2&2&0&2\\ 2&2&6&8&4&8\\ 1&1&5&6&2&5\\ 1&1&3&4&2&7 \end{bmatrix}$$

To simplify the matrix, the first step would be to locate the first column with a nonzero element, interchange rows if necessary to place a nonzero element in the first row, and multiply the row by the inverse of that element to give a leading 1. Interchanging rows 1 and 2 and dividing by 2 give

$$\begin{bmatrix} 1&1&3&4&2&4\\ 0&0&2&2&0&2\\ 1&1&5&6&2&5\\ 1&1&3&4&2&7 \end{bmatrix}$$

The next step is to subtract a multiple of the first row from each other row to make the rest of the column corresponding to the leading element in the first row 0:

$$\begin{bmatrix} 1&1&3&4&2&4\\ 0&0&2&2&0&2\\ 0&0&2&2&0&1\\ 0&0&0&0&0&3 \end{bmatrix}$$

Then, disregarding the first row, again the first column with a nonzero element is located, and rows are interchanged if necessary to place a nonzero element in this column in the second row. The row is next multiplied by the inverse of its leading element to give a leading 1. This is accomplished in the above matrix by dividing the second row by 2. Then the appropriate multiple of this row is subtracted from each other row to make *all* the other entries 0 in the column of the leading element of the second row. This yields

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

One more step in the process yields

$$\begin{bmatrix} 1\,1\,0\,1\,2\,0 \\ 0\,0\,1\,1\,0\,0 \\ 0\,0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0\,0 \end{bmatrix}$$

This process will always result in a matrix in echelon canonical form.

The nonzero rows of a matrix in echelon canonical form are linearly independent, and thus the number of nonzero rows is the dimension of the row space. It can be shown that there is only one matrix in echelon canonical form for any given row space.

If all the rows of an $n \times n$ matrix are linearly independent, the matrix is said to be nonsingular. When such a matrix is put in echelon canonical form, there must still be n linearly independent rows, and thus every row must contain a 1. This can occur only if it has 1's on the main diagonal and 0's elsewhere. Such a matrix is called an *identity matrix* and denoted **I**. Thus any nonsingular matrix can be transformed into an identity matrix by elementary row operations.

The *transpose* of an $n \times m$ matrix $\mathbf{M}$ is an $m \times n$ matrix, denoted $\mathbf{M}^T$, whose rows are the columns of $\mathbf{M}$, and thus whose columns are the rows of $\mathbf{M}$. The transpose of $[a_{ij}]$ is $[a_{ji}]$.

Two $n \times m$ matrices can be added, element by element:

$$[a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}].$$

With this definition it is easily verified that matrices form an Abelian group under addition.

An $n \times k$ matrix $[a_{ij}]$ and a $k \times m$ matrix $[b_{ij}]$ can be multiplied to give an $n \times m$ product matrix $[c_{ij}]$ by the rule

$$c_{ij} = \sum_{l=1}^{k} a_{il} b_{lj}.$$

It can be verified by direct calculation that with this definition matrix multiplication satisfies the associative law, and multiplication and addition satisfy the distributive law.

The element c_{ij} of the product is the inner product of the ith row of $[a_{ij}]$ by the jth column of $[b_{ij}]$. Also the ith row vector of the product $[c_{ij}]$ is a linear combination of the row vectors of $[b_{ij}]$ with the coefficient a_{il} on the lth row. Similarly the columns of the product are linear combinations of the column vectors of $[a_{ij}]$.

Multiplying an $n \times m$ matrix $\mathbf{M}$ on the left by an $n \times n$ matrix $\mathbf{P}$ that has one 1 in each row and each column and all the rest of the elements 0 simply permutes the rows of the matrix $\mathbf{M}$, and any permutation of rows can be accomplished in this way. Thus the first elementary row operation can be accomplished by multiplying on the left by a permutation matrix. The second elementary row operation, multiplying the jth row of $\mathbf{M}$ by c, can be accomplished by multiplying $\mathbf{M}$ on the left by a matrix that has 0's off the main diagonal, c on the main diagonal in the jth row, and 1's on the rest of the main diagonal. Finally, the third elementary operation, adding c times the jth row to the kth row, can be accomplished by multiplying on the left by a matrix that has 1's on the main diagonal, c in the position that is in the jth column and kth row, and 0's elsewhere. These matrices are called *elementary matrices.*

THEOREM 2.11. *Every nonsingular matrix has a left inverse that is a product of elementary matrices.*

Proof. If a nonsingular $n \times n$ matrix $\mathbf{M}$ is transformed into echelon canonical form, it becomes an identity matrix. Since $\mathbf{M}$ can be put in echelon canonical form by elementary row operations, there is some set of elementary matrices $\mathbf{E}_1, \ldots, \mathbf{E}_k$ whose product with M is the identity matrix:

$$\mathbf{E}_k \mathbf{E}_{k-1} \cdots \mathbf{E}_1 \mathbf{M} = \mathbf{I}.$$

Then $\mathbf{E}_k \cdots \mathbf{E}_1$ is the left inverse of $\mathbf{M}$. Q.E.D.

It can be shown that the left inverse of a matrix is also a right inverse.

THEOREM 2.12. *If* $\mathbf{M}$ *is an* $n \times m$ *matrix and* $\mathbf{S}$ *is a nonsingular* $n \times n$ *matrix, then the product of* $\mathbf{S}$ *and* $\mathbf{M}$ *has the same row space as* $\mathbf{M}$ *has.*

Proof. The rows of $\mathbf{SM}$ are linear combinations of the rows of $\mathbf{M}$, and therefore the row space of $\mathbf{SM}$ is contained in the row space of $\mathbf{M}$. But $\mathbf{S}$ has a left inverse $\mathbf{S}^{-1}$, and the rows of $\mathbf{S}^{-1}\mathbf{SM} = \mathbf{M}$ are linear combinations of the rows of $\mathbf{SM}$, and hence the row space of $\mathbf{M}$ is contained in the row space of $\mathbf{SM}$. Therefore, they must be equal.

Q.E.D.

THEOREM 2.13. *The set of all n-tuples orthogonal to a subspace* V_1 *of n-tuples forms a subspace* V_2 *of n-tuples.*

Proof. Let V_1 be a subspace of the vector space of all n-tuples over a field. Let V_2 be the set of all vectors orthogonal to every vector in V_1. Let $\mathbf{v}$ be any vector in V_1 and $\mathbf{u}_1$ and $\mathbf{u}_2$ any vectors in V_2. Then $\mathbf{v} \cdot \mathbf{u}_1 = \mathbf{v} \cdot \mathbf{u}_2 = 0$, and $\mathbf{v} \cdot \mathbf{u}_1 + \mathbf{v} \cdot \mathbf{u}_2 = 0 = \mathbf{v} \cdot (\mathbf{u}_1 + \mathbf{u}_2)$. Therefore $\mathbf{u}_1 + \mathbf{u}_2$ is in V_2. Also $\mathbf{v} \cdot (c\mathbf{u}_1) = c(\mathbf{v} \cdot \mathbf{u}_1) = 0$, so $c\mathbf{u}_1$ is in V_2. Thus V_2 must be a subspace. Q.E.D.

The subspace V_2 in Theorem 2.13 is called the *null space* of V_1.

THEOREM 2.14. *If a vector is orthogonal to every vector of a set which spans* V_1, *it is in the null space of* V_1.

Proof. If $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span V_1, then every vector in V_1 can be expressed in the form $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$. Then

$$\mathbf{v} \cdot \mathbf{u} = (c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{u} = c_1\mathbf{v}_1 \cdot \mathbf{u} + \cdots + c_k\mathbf{v}_k \cdot \mathbf{u}$$

and if $\mathbf{u}$ is orthogonal to each $\mathbf{v}_i$, it is orthogonal to $\mathbf{v}$. Q.E.D.

The null space of the row space of a matrix is called the null space of the matrix. A vector is in the null space of a matrix if it is orthogonal to each row of the matrix. If the n-tuple $\mathbf{v}$ is considered to be a $1 \times n$ matrix, $\mathbf{v}$ is in the null space of an $m \times n$ matrix $\mathbf{M}$ if and only if $\mathbf{vM}^T = \mathbf{0}$.

THEOREM 2.15. *If the dimension of a subspace of n-tuples is k, the dimension of the null space is* $n - k$.

The proof of this theorem will be omitted, because it requires some background otherwise unnecessary. One consequence of the theorem is

THEOREM 2.16. *If* V_2 *is a subspace of n-tuples and* V_1 *is the null space of* V_2, *then* V_2 *is the null space of* V_1.

Proof. If V_2 has dimension k, then V_1 has dimension $n - k$, and the null space of V_1 has dimension k. Since V_2 is contained in the null space of V_1 and has the same dimension, they are equal. Q.E.D.

If $\mathbf{M}_1$ and $\mathbf{M}_2$ are two matrices that have n columns, and if $\mathbf{M}_1 \mathbf{M}_2^T$ is a matrix of all 0's, then the row space of $\mathbf{M}_2$ is contained in the null space of $\mathbf{M}_1$, and vice versa. If the row rank of $\mathbf{M}_1$ and the row rank of $\mathbf{M}_2$ add to n, then the row space of $\mathbf{M}_2$ is the null space of $\mathbf{M}_1$, and vice versa.

Let $U \cap V$ be vector spaces and let $U \cap V$ denote the set of vectors that are contained in both U and V. It is easy to verify that $U \cap V$ is a subspace. Let $U \oplus V$ denote the subspace consisting of all linear combinations $a\mathbf{u} + b\mathbf{v}$, where $\mathbf{u} \in U$, $\mathbf{v} \in V$, and a and b are scalars.

THEOREM 2.17. *The sum of the dimensions of $U \cap V$ and $U \oplus V$ equals the sum of the dimensions of U and V.*

Proof. Let k_1 denote the dimension of U, k_2 the dimension of V, and k_0 the dimension of $U \cap V$. Then there exists a basis of k_0 vectors for $U \cap V$. It will be possible to find a basis for U consisting of these k_0 vectors and $k_1 - k_0$ others not in $U \cap V$, and a basis for V consisting of the basis of $U \cap V$ and $k_2 - k_0$ others. Then together, the k_0 vectors in the basis of $U \cap V$, the $k_1 - k_0$ additional vectors in the basis of U and the $k_2 - k_0$ vectors in the basis of V form a basis of $U \oplus V$. Therefore, the dimension of $U \oplus V$ is $k_0 + (k_1 - k_0) + (k_2 - k_0)$. Q.E.D

THEOREM 2.18. *Let U_2 be the null space of U_1 and V_2 the null space of V_1. Then $U_2 \cap V_2$ is the null space of $U_1 \oplus V_1$.*

Proof. Since U_1 is contained in $U_1 \oplus V_1$, every vector in the null space of $U_1 \oplus V_1$ must be in U_2, the null space of U_1. Similarly, every vector in the null space of $U_1 \oplus V_1$ must be in V_2, the null space of V_1. Therefore, the null space of $U_1 \oplus V_1$ is contained in $U_2 \cap V_2$. Every vector in $U_1 \oplus V_1$ can be written in the form $a\mathbf{u}_1 + b\mathbf{v}_1$. If $\mathbf{w}$ is any element of $U_2 \cap V_2$, then $\mathbf{u}_1 \cdot \mathbf{w} = \mathbf{v}_1 \cdot \mathbf{w} = 0$ and hence $(a\mathbf{u}_1 + b\mathbf{v}_1) \cdot \mathbf{w} = 0$. Therefore, $U_2 \cap V_2$ is contained in the null space of $U_1 \oplus V_1$. It follows that the null space of $U_1 \oplus V_1$ equals $U_2 \cap V_2$. Q.E.D

There are many important concepts and theorems of matrix theory that have not been mentioned. It should be emphasized that, while the material presented here may be adequate for understanding what follows, it is certainly no substitute for books or courses on modern algebra, which can provide a well-rounded understanding of the subject.

Notes

There are many good textbooks on algebra and on matrices. Birkhoff and Mac Lane (1941) covers all the material of this chapter and much more. It is clearly written and is probably the most easily understood text on modern algebra. It also contains an extensive bibliography. Van der Waerden (1949) is also well written and highly regarded, and goes generally deeper into the subject.

Problems

2.1. Show that there is only one group of three elements. Show that there are only two distinct groups with four elements, and that both are Abelian.

2.2. Show that if the operation is taken as addition in the groups of Problem 2.1, multiplication can be defined to make them rings.

2.3. The set of all nonnegative integers is not a group with the operation addition. Why? It is also not a group with the operation multiplication. Why?

2.4. The set of all $n \times n$ matrices is not a skew field. Why? The set of all nonsingular $n \times n$ matrices and the all-zero matrix is also not a skew field. Why?

2.5. Show that the set of all integers with the operation subtraction does not satisfy the associative law.

2.6. Solve these simultaneous equations for x and y, assuming the coefficients to be in the field of 4 elements as given in Table 2.3:

$$ax + y = b,$$
$$x + ay = b.$$

(Answer: $x = y = 1$.)

2.7. Calculate the determinant of the following matrix. Put the matrix in echelon canonical form and show that the rank is 3. Express the inverse as a product of elementary matrices. Assume the field of three elements.

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

2.8. By row operations, find the echelon canonical form for the following matrix. Also calculate the determinant. Assume the field with two elements.

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

2.9. Show that in the vector space of n-tuples over the field of two elements, every subgroup (under addition) is a subspace. (Compare Problem 6.8.)

2.10. Define the Hamming weight $w(\mathbf{v})$ of an n-tuple $\mathbf{v}$ as its Hamming distance from the zero n-tuple. Show that

$$d(\mathbf{u}, \mathbf{v}) = w(\mathbf{u} - \mathbf{v}).$$

2.11. Let H be a subspace of n-tuples, and define the Hamming weight of a coset of H as the minimum Hamming weight of elements of the coset. Define the distance between two cosets as the weight of the difference of the two cosets, which is also a coset. Show that this distance function is a metric. (Compare Problem 1.2.)

2.12. If an $n \times n$ matrix has the form

$$\begin{bmatrix} \mathbf{I}_k & \mathbf{P} \\ \mathbf{0} & \mathbf{I}_{n-k} \end{bmatrix} = \mathbf{M}$$

where $\mathbf{I}_k$ is a $k \times k$ identity matrix, $\mathbf{I}_{n-k}$ is an $(n - k) \times (n - k)$ identity matrix, $\mathbf{0}$ is an $(n - k) \times k$ matrix of all 0's, and $\mathbf{P}$ is an arbitrary $(k \times n - k)$ matrix, show that the inverse of $\mathbf{M}$ has the same form with $\mathbf{P}$ replaced by $-\mathbf{P}$.

2.13. Prove that the set of all $n \times n$ square matrices that have 1's on the main diagonal and 0's below the main diagonal forms a group under multiplication.

2.14. Show that the integers modulo 4 form a commutative ring but not a field. Compare Table 2.2. and Problem 2.1.