Lecture 7 — Vector spaces

Dr. D. Bernhard

In this lecture: vector spaces — linear independence and bases — linear maps — application to finite fields

Learning outcomes. After this lecture and revision, you should be able to:

- Define vector spaces.
- Multiply and (where possible) invert matrices, especially over finite fields.
- Interpret polynomial spaces over fields and finite fields as vector spaces and operations on them as linear maps.

7 Vector spaces

After the last lecture's very intense computation in finite fields, we close this section of the course with a slightly easier topic that we've actually been using implicitly already, namely vector spaces.

7.1 Definitions

The new thing about vector spaces is that we start with a given structure, a field, and build a vector space over this field. Vector spaces are groups in which you can add vectors but also "scale" vectors by multiplying them with field elements; in general you cannot multiply vectors with each other. The multiplication operation that takes a field element and a vector is sometimes called scalar multiplication.

> **Definition 7.1.** Start with any field $\mathbb{F}$. A vector space over $\mathbb{F}$ is a structure $(V, +, \cdot)$ where $+ : V \times V \to V$ and $\cdot : \mathbb{F} \times V \to V$ satisfying the following laws:
>
> 1. $(V, +)$ is an Abelian group.
> 2. Field and scalar multiplication associate: for any field elements f, g and any vector $\overline{a}$ we have $(fg) \cdot \overline{a} = f \cdot (g \cdot \overline{a})$. Here fg is multiplication in $\mathbb{F}$.

3. Field multiplication distributes over vector addition: for any vectors $\overline{a}, \overline{b} \in V$ and any field elements $f, g \in \mathbb{F}$ we have $f \cdot (\overline{a} + \overline{b}) = f \cdot \overline{a} + f \cdot \overline{b}$ and $(f +_{\mathbb{F}} g) \cdot \overline{a} = f \cdot \overline{a} + g \cdot \overline{a}$. We marked the field addition with $+_{\mathbb{F}}$ to distinguish it from vector addition here.

We use the convention that we write vectors with a line over them, e.g. $\overline{a}$ to distinguish them from field elements.

Examples. We have encountered many vector spaces already without mentioning it.

- Any field is automatically a vector space over itself; vector addition and scalar multiplication are just field addition and multiplication.
- The most common notion of a vector is a tuple or sequence of elements. For any field $\mathbb{F}$, the vector space $\mathbb{F}^n$ consists of vectors of length n with componentwise addition and the scalar multiplication $f \cdot (v_1, \ldots, v_n) := (f v_1, \ldots, f v_n)$.
- By the same logic, the polynomials over a field form a vector space, as do the polynomials over a field modulo some fixed polynomial. It is thus possible to interpret $GF(p^n)$ as the n-dimensional vector space $GF(p)^n$, “forgetting” about the multiplication of polynomials.

Linear (in)dependence and bases. Where a vector space is, a basis (plural: bases) is not far away. A basis plays a similar role to a set of generators of a group, but the additional field multiplication that turns a group into a vector space gives us much more to work with. Specifically, linear combinations:

Definition 7.2 (linear combination). For a finite set $\{\overline{v}_i\}_i$ of vectors, a linear combination is a sum $\sum_i c_i \cdot \overline{v}_i$ with coefficients c_i in the field $\mathbb{F}$.

If the index set is something like $I = \{1, 2, \ldots, n\}$ then we can write a linear combination as $c_1 \cdot \overline{v}_1 + \ldots + c_n \cdot \overline{v}_n$.

◇ The basic definitions of linear algebra (linear independence, basis etc.) can also be defined for infinite sets, but the exact definition is a bit subtle. We will not need to worry about this too much in this course. A linear combination for an infinite set V of vectors is a sum where only a finite number of coefficients are non-zero.

A linear combination of vectors where all coefficients are zero (the neutral element of field addition) is automatically the zero vector (the neutral element of vector addition). A set of vectors is linearly independent if this is the only linear combination that is zero; another way of saying this is that no vector in the set can be written as a linear combination of the others.

Definition 7.3 (linear (in)dependence). A set $\{\overline{v}_i\}_i$ of vectors is linearly independent if no linear combination of the vectors $\sum_i c_i \cdot \overline{v}_i$ with coefficients in $\mathbb{F}$ gives the zero vector (neutral element of vector addition), unless all coefficients are already zero (the neutral element of the field's addition). A set of vectors that is not linearly independent is called linearly dependent.

◇ An infinite set V is linearly independent if no finite sum of elements in V with coefficients in $\mathbb{F}$ gives the zero vector, unless all coefficients are zero. This is equivalent to saying that every finite subset of V is linearly independent.

And finally, a basis is a finite set of linearly independent vectors (in a particular order) that generates the entire space.

Definition 7.4 (basis). A basis of a vector space V is a finite list $(\overline{v}_1, \ldots, \overline{v}_n)$ of vectors that is linearly independent and generates V, i.e. $V = \langle \overline{v}_1, \ldots, \overline{v}_n \rangle$.

In other words, every vector $\overline{w}$ in the space can be written as a linear combination of the basis vectors: $w = w_1\overline{v}_1 + \ldots + w_n\overline{v}_n$. In fact, if a vector space has a basis then any two bases of the space have the same number of elements (which we call the dimension of the space) and for any vector v and any basis in a fixed order, there is exactly one way (one tuple of coefficients) to write W as a linear combination of the basis vectors.

◇ An infinite set W of vectors is a basis of a vector space V if (1) it is linearly independent — that is, every finite subset of W is linearly independent in the usual sense and (2) every element in $v \in V$ can be written as a *finite* linear combination of elements in W. This definition is required to make the theorem "every vector space has a basis" true even in the infinite-dimensional case, assuming the Axiom of Choice.

Proposition 7.5. Any two bases of a vector space have the same number of elements. If a vector space has a basis with n elements, we say that the space has dimension n.

And finally, every vector space has a basis. This proposition is only really mathematically interesting to discuss in the infinite case but it is the start of most constructions in linear algebra: given any vector space V, we can simply assume that a basis B is given as well.

Proposition 7.6. Every vector space has a basis.

Linear maps. A vector space homomorphism is a function $f : V \to W$ between two vector spaces over the same field $\mathbb{F}$ that preserves vector addition and scalar multiplication. We call such a function a linear map.

Definition 7.7 (linear). If V and W are two vector spaces over a field $\mathbb{F}$, we call a function $f : V \to W$ linear if for any $\overline{x}, \overline{y}$ in V and any $a \in \mathbb{F}$ we have

- $f(\overline{v} + \overline{w}) = f(\overline{v}) + f(\overline{w})$.
- $f(a \cdot \overline{v}) = a \cdot f(\overline{v})$.

The reader should understand by now which operation symbols refer to V-operations and which ones refer to W-operations.

If V is a vector space with basis $B = (\overline{b}_1, \ldots, \overline{b}_n)$ then a linear map $f : V \to W$ can be computed on any vector from its values on the basis alone. Namely, if you know $f(\overline{b}_1), \ldots, f(\overline{b}_n)$ and want to compute $f(\overline{v})$ then you can write $\overline{v}$ in exactly one way as $v = c_1 \cdot \overline{b}_1 + \ldots + c_n \cdot \overline{b}_n$, giving $f(\overline{v}) = c_1 \cdot f(\overline{b}_1) + \ldots + c_n \cdot f(\overline{b}_n)$.

If we have a basis $P = (\overline{p}_1, \ldots, \overline{p}_m)$ of W as well, you can compute the coefficients of the images of the basis elements under f: there are unique coefficients $(a_{1,1}, \ldots, a_{1,m})$ such that $f(\overline{b}_1) = a_{1,1} \cdot \overline{p}_1 + \ldots + a_{1,m} \cdot \overline{p}_m$ and the same for the other basis elements. In other words, a linear map between a n-dimensional vector space V and a m-dimensional vector space W can be specified as a $n \cdot m$ rectangle of coefficients in the field $\mathbb{F}$:

$$f : V \to W \quad \leftrightarrow \quad \begin{pmatrix} a_{1,1} & \ldots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \ldots & a_{m,n} \end{pmatrix}$$

We call such a rectangle of coefficients a matrix. Matrices form a vector space which we write $\mathbb{F}^{m \times n}$ for the space of matrices with m rows and n columns as in the example above. Matrix addition is component-wise; scalar multiplication with a field element just multiplies all matrix components with the field element.

Applying a linear map to a vector becomes matrix-vector multiplication. If we consider linear maps from a space to itself, we can compose them: for $f, g : V \to V$ we can from the map fg that takes $\overline{v}$ to $f(g(\overline{v}))$. If both f and g are linear, so is fg (exercise). We can use this to define a multiplication operation on square matrices, turning $\mathbb{F}^{n \times n}$ into a ring (for every positive integer n) — this is just the usual matrix multiplication. Of course we can define matrix multiplication between compatible non-square matrices too but we don't get a ring that way.

7.2 Polynomial spaces as vector spaces

We look at the space $V = GF(p)^n$ constructed by taking a finite field $GF(p^n) = GF(p)[X]/q(X)$ modulo an irreducible polynomial q of degree $n \geq 1$ and interpreting it as a vector space. The elements of this space are of the form $(c_0, c_1, \ldots, c_{n-1})$ which could be written as $c_0 + c_1X + \ldots + c_{n-1}X^{n-1}$. Interpreting it as a vector space means forgetting about multiplication of $GF(p^n)$ elements but adding multiplication with $GF(p)$ elements: $a \cdot (c_0, \ldots, c_{n-1}) := (a \cdot c_0, \ldots, a \cdot c_{n-1})$.

One basis of this vector space consists of vectors $\overline{b}_i$ for $i = 0$ to $n-1$ where $\overline{b}_i$ is 1 at position r and 0 elsewhere. Written as polynomials, the i-th basis vector $\overline{b}_i$ is the monomial X^i.

The map $GF(p) \to GF(p^n), a \mapsto (a, 0, \ldots, 0)$ is a field homomorphism. It is sometimes called the embedding of the base field $GF(p)$ into the extension field $GF(p^n)$. This map commutes with field multiplication in the following way: for any element a of $GF(p)$ and any element $\overline{v}$ of $GF(p^n)$, you get the same if you perform the scalar multiplication $a \cdot \overline{v}$ or if you embed a in $GF(p^n)$ and then do field multiplication there. We can express this in a diagram.

$$\begin{tikzcd}
\begin{array}{c} a \in GF(p) \\ \overline{v} \in GF(p^n) \end{array} \arrow[r, "\text{embedding}"] \arrow[d, "\text{scalar mult.}"'] & \begin{array}{c} (a, 0, \ldots, 0) \in GF(p^n) \\ \overline{v} = (v_0, \ldots, v_{n-1}) \in GF(p^n) \end{array} \arrow[dl, "GF(p^n)\ \text{mult.}"] \\
(av_0, \ldots av_{n-1}) &
\end{tikzcd}$$

7.3 Automorphisms revisited

An automorphism f of $GF(p^n)$ can be represented as a $n \times n$ matrix, since such a f must be linear over the field $\mathbb{F}$: if $a \in \mathbb{F}$ and $\overline{v} \in V$ then $f(a\overline{v}) = a \cdot f(\overline{v})$. However, we know that field automorphisms cannot change degree-0 polynomials: $f(1, 0, \ldots, 0) = (1, 0, \ldots, 0)$. Writing f out as a matrix, we see

$$\begin{pmatrix} f_{0,0} & f_{0,1} & \cdots & f_{0,n-1} \\ f_{1,0} & f_{1,1} & \cdots & f_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n-1,0} & f_{n-1,1} & \cdots & f_{n-1,n-1} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} f_{0,0} \\ f_{1,0} \\ \vdots \\ f_{n-1,0} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

so the zeroth column of the matrix of a must start with a 1 and be zero everywhere else. Since we also know how to compute $f(X^2)$ from $f(X)$, this means we know how to compute $f(0,0,1,0,\dots,0)$ from $f(0,1,0,\dots,0)$ and so on — so all the information about f is contained in the first column of the matrix of f and we can always compute the other columns from it.

We look at two examples. The first is $GF(7^2)$ where for $p(X) = X^2+X+6$. we found two automorphisms id, ϕ with $id(X) = X$ and $\phi(X) = 6+6X$ (the Frobenius map). As matrices, these are

$$id = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \phi = \begin{pmatrix} 1 & 6 \\ 0 & 6 \end{pmatrix}$$

It is now obvious how to calculate ϕ on an arbitrary field element $(a+bX)$:

$$\phi(a+bX) = \begin{pmatrix} 1 & 6 \\ 0 & 6 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+6b \\ 6b \end{pmatrix}$$

If we introduce another irreducible polynomial $q(Y) = Y^2+1$, the isomorphisms we found between these representations last time were $f_1(a,b) = (a+3b, 2b)$ and $f_2(a,b) = (a+3b, 5b)$. If we know f_1, we can calculate f_2 by multiplying from the right with the Frobenius map:

$$f_1\phi = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 6 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 0 & 5 \end{pmatrix} = f_2$$

Matrices give us an easy way to find the inverses of isomorphisms, that is the isomorphisms of $GF(7^2)$ from the representation modulo $q(Y)$ back to the representation modulo $p(X)$. All we need to do is invert the matrices of f_1, f_2:

$$\begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}^{(-1)} = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 3 \\ 0 & 5 \end{pmatrix}^{(-1)} = \begin{pmatrix} 1 & 5 \\ 0 & 3 \end{pmatrix}$$

Giving $f_1^{(-1)}(a+bX) = (a+2b)+4bX$ and $f_2^{(-1)}(a+bX) = (a+5b)+3bX$.

Our second example is $GF(2^3)$, this time with the irreducible polynomial $p(X) = X^3+X+1$. The Frobenius map sends $X \mapsto X^2$ and $X^2 \mapsto [X^4] = X^2+X$. As a matrix, we get

$$\phi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \qquad \phi \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a \\ c \\ b+c \end{pmatrix}$$

from which we read off $\phi(a+bX+cX^2) = a+cX+(b+c)X^2$. Finding the other automorphisms is easy too:

$$\phi^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

which maps $X \mapsto X + X^2$ (middle column) and $X^2 \mapsto X$ (right column). Multiplying with the column vector $(a; b; c)$ we get $\phi^2(a + bX + cX^2) = a + (b + c)X + bX^2$. If we look at the third power

$$\phi \cdot \phi^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

we get the identity map back, as expected. If we take the isomorphism $f(X) = Y^2 + 1$ into the representation modulo $q(Y) = Y^3 + Y^2 + 1$, we find $f(X^2) = Y^2 + Y$. The other isomorphisms are

$$f_1\phi = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = f_2$$

$$f_1\phi^2 = f_3$$

from which we read off the columns $f_2(X) = Y^2 + Y$ and $f_2(X^2) = 1 + Y$, giving $f_2(a + bX + cX^2) = (a + c) + (b + c)Y + bY^2$. We could read this last formula off the rows of the matrix directly, since we get the last formula by multiplying the matrix of f_2 with the column vector $(a; b; c)$. To invert the isomorphisms, one again just needs to invert the matrices.

Exercise. *The rest of the example.*

- ($\star$) Compute $f_3 = f_1\phi^2$, then write out the expressions for $f_3(X)$, $f_3(X^2)$ and $f_3(a + bX + cX^2)$.
- ($\star\star$) Invert the matrix for f_1 to get the inverse isomorphism. Note: matrix inversion is a lot easier in $GF(2)$ as $1 + 1 = 0$ so you never need to multiply rows through to cancel constants!

Exercise. ($\star$) *More finite fields.* Consider the field $GF(3^3)$ with the irreducible polynomials $p(X) = X^3 + 2X + 1$ and $q(Y) = Y^3 + 2Y^2 + Y + 1$.

1. Find the Frobenius map ϕ as a 3×3 matrix modulo $p(X)$.
2. Find the powers of the Frobenius map.
3. One isomorphism f between the representations $p(X)$ and $q(Y)$ has $f(X) = 2X^2 + 2X$. Find the matrix of f.
4. Find the inverse of f by inverting the matrix of f.
5. How many isomorphisms are there between the two representations?

6. Find the other isomorphisms by matrix multiplication using the matrix of f and the Frobenius map ϕ.
7. ($\star\star$) Here is another way to find the Frobenius map in the representation modulo $q(Y)$. We have the following situation with $V = GF(3)[X]/p(X)$ and $W = GF(3)[Y]/q(Y)$:

$$\begin{tikzcd}
V \arrow[r, shift left, "f"] \arrow[d, "\phi"'] & W \arrow[l, shift left, "f^{(-1)}"] \arrow[d, "\hat{\phi}"] \\
V \arrow[r, shift left, "f"] & W \arrow[l, shift left, "f^{(-1)}"]
\end{tikzcd}$$

From this we see that the Frobenius map in W has matrix $\hat{\phi} = f \cdot \phi \cdot f^{(-1)}$. Compute $\hat{\phi}$ this way.

◇ The conditions for field automorphisms say that $f(\overline{v} + \overline{w}) = f(\overline{v}) + f(\overline{w})$, $f(\overline{v} \cdot \overline{w}) = f(\overline{v}) \cdot f(\overline{w})$ and $f(1, 0, \ldots, 0) = (1, 0, \ldots, 0)$ since this element is the one of the field. Since the scalar multiplication $a \cdot \overline{v}$ we get for V as a vector space is equivalent to the vector multiplication $(a, 0, \ldots, 0) \cdot \overline{v}$, a field automorphism must satisfy $f(a \cdot \overline{v}) = f((a, 0, \ldots, 0) \cdot \overline{v}) = f(a, 0, \ldots, 0) \cdot f(\overline{v}) = (a, 0, \ldots, 0) \cdot f(\overline{v}) = a \cdot f(\overline{v})$. This explains why in field multiplication we get $f(a \cdot \overline{v}) = a \cdot f(\overline{v})$ with the a appearing "outside f" whereas the automorphism rule says $f(\overline{v} \cdot \overline{w}) = f(\overline{v}) \cdot f(\overline{w})$.