3.1 Linear Algebra

Start with a field F (this will be the field of **scalars**).

Definition: A **vector space over** F is a set V with a vector addition and scalar multiplication ("scalars" in F times "vectors" in V) so that:

(a) Vector addition is associative and commutative.

(b) There is an additive identity vector, denoted 0, or sometimes $\vec{0}$.

(c) Every vector $\vec{v}$ has an additive inverse vector $-\vec{v}$.

(d) Scalar multiplication distributes with vector addition.

(e) If $c, k \in F$ are scalars and $\vec{v} \in V$ is a vector, then $c(k\vec{v}) = (ck)\vec{v}$.

(f) If $1 \in F$ is the multiplicative identity, then $1\vec{v} = \vec{v}$ for all $\vec{v}$.

Examples: (a) F^n is the standard finite-dimensional vector space of n-tuples of elements of F. Vectors $\vec{v} \in F^n$ will be written vertically:

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad k \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_n \end{bmatrix}$$

(b) If $F \subset D$ and D is a commutative ring with 1, then D is a vector space over F. The scalar multiplication is ordinary multiplication in D, and property (e) is the associative law for multiplication in D. Thus, for example, vector spaces over $\mathbb{Q}$ include $\mathbb{R}, \mathbb{C}, \mathbb{Q}[x]$ and $\mathbb{Q}(x)$.

Definition: A **basis** of a vector space V is a set of vectors $\{\vec{v}_i\}$ that:

(i) **Span**. Every vector is a linear combination of the $\vec{v}_i$:

$$\vec{v} = k_1\vec{v}_1 + ... + k_n\vec{v}_n$$

and

(ii) **Are Linearly Independent.** The only way:

$$k_1\vec{v}_1 + ... + k_n\vec{v}_n = 0$$

is if all the scalars $k_1, ..., k_n$ are zero.

Proposition 3.1.1. *If $\{\vec{v}_1, ..., \vec{v}_n\}$ is a basis of V, then every vector $\vec{v} \in V$ is a* **unique** *scalar linear combination of the basis vectors:*

$$\vec{v} = k_1\vec{v}_1 + ... + k_n\vec{v}_n$$

and any other basis $\{\vec{w}_i\}$ of V must also consist of a set of n vectors. The number n is called the **dimension** *of the vector space V over F.*

Proof: Since the $\{\vec{v}_i\}$ span, each vector $\vec{v}$ has at least one expression as a linear combination of the $\vec{v}_i$, and if there are two:

$$\vec{v} = k_1\vec{v}_1 + ... + k_n\vec{v}_n \text{ and } \vec{v} = l_1\vec{v}_1 + ... + l_n\vec{v}_n$$

then subtracting them gives: $0 = (k_1 - l_1)\vec{v}_1 + ... + (k_n - l_n)\vec{v}_n$. But then each $k_i = l_i$ because the $\{\vec{v}_i\}$ are linearly independent, and thus the two linear combinations are the same. This gives uniqueness.

Now take another basis $\{\vec{w}_i\}$ and solve: $\vec{w}_1 = b_1\vec{v}_1 + ... + b_n\vec{v}_n$. We can assume (reordering the $\vec{v}_i$ if necessary) that $b_1 \neq 0$. Then:

$$\vec{v}_1 = \frac{1}{b_1}\vec{w}_1 - \frac{b_2}{b_1}\vec{v}_2 - ... - \frac{b_n}{b_1}\vec{v}_n$$

and then $\{\vec{w}_1, \vec{v}_2, ..., \vec{v}_n\}$ is another basis of V because every

$$\vec{v} = k_1\vec{v}_1 + ... + k_n\vec{v}_n = k_1(\frac{1}{b_1}\vec{w}_1 - \frac{b_2}{b_1}\vec{v}_2 - ... - \frac{b_n}{b_1}\vec{v}_n) + k_2\vec{v}_2 + ... + k_n\vec{v}_n$$

so the vectors span V, and the only way:

$$0 = k_1\vec{w}_1 + ... + k_n\vec{v}_n = k_1(b_1\vec{v}_1 + ... + b_n\vec{v}_n) + k_2\vec{v}_2 + ... + k_n\vec{v}_n$$

is if $k_1b_1 = 0$ (so $k_1 = 0$) and each $k_1b_i + k_i = 0$ (so each $k_i = 0$, too!)

Similarly we can replace each $\vec{v}_i$ with a $\vec{w}_i$ to get a sequence of bases: $\{\vec{w}_1, \vec{w}_2, \vec{v}_3, ..., \vec{v}_n\}, \{\vec{w}_1, \vec{w}_2, \vec{w}_3, \vec{v}_4, ..., \vec{v}_n\}$, etc. If there were **fewer** of the $\vec{w}_i$ basis vectors than $\vec{v}_i$ basis vectors we would finish with a basis:

$$\{\vec{w}_1, ..., \vec{w}_m, \vec{v}_{m+1}, ..., \vec{v}_n\}$$

which is impossible, since $\{\vec{w}_1, ..., \vec{w}_m\}$ is already a basis! Similarly, reversing the roles of the $\vec{v}_i$'s and $\vec{w}_i$'s, we see that there cannot be fewer $\vec{v}_i$'s than $\vec{w}_i$'s. So there must be the same number of $\vec{w}_i$'s as $\vec{v}_i$'s!

Examples:

(a) F^n has n "standard" basis vectors:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, ..., \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

(b) $\mathbb{R}^1$ is the line, $\mathbb{R}^2$ is the plane, and $\mathbb{R}^3$ is space.

(c) $\mathbb{C}$ has basis $\{1, i\}$ as a vector space over $\mathbb{R}$.

(d) $\mathbb{Q}[x]$ has infinite basis $\{1, x, x^2, x^3, ...\}$ as a vector space over $\mathbb{Q}$.

(e) It is hard to even imagine a basis for $\mathbb{R}$ as a vector space over $\mathbb{Q}$.

(f) Likewise it is hard to imagine a basis for $\mathbb{Q}(x)$ over $\mathbb{Q}$.

We can create vector spaces with **polynomial clock arithmetic**. Given

$$f(x) = x^d + a_{d-1}x^{d-1} + ... + a_0 \in F[x]$$

we first define the "mod $f(x)$" equivalence relation by setting

$$g(x) \equiv h(x) \pmod{f(x)}$$

if $g(x) - h(x)$ is divisible by $f(x)$, and then the "polynomial clock":

$$F[x]_{f(x)} = \{[g(x)]\}$$

is the set of "mod $f(x)$" equivalence classes.

Proposition 3.1.2. *The polynomial clock $F[x]_{f(x)}$ is a commutative ring with* 1 **and** *a vector space over F with basis:*

$$\{[1], [x], ..., [x^{d-1}]\}$$

and if $f(x)$ is a **prime** *polynomial, then the polynomial clock is a field.*

Proof: Division with remainders tells us that in every equivalence class there is a "remainder" polynomial $r(x)$ of degree $< d$. This tells us that the vectors:

$$[1], [x], [x^2], ..., [x^{d-1}] \in F[x]_{f(x)}$$

span the polynomial clock. They are linearly independent since if:

$$b_{d-1}[x^{d-1}] + ... + b_0[1] = 0$$

then $r(x) = b_{d-1}x^{d-1} + ... + b_0$ is divisible by $f(x)$, which is impossible (unless $r(x) = 0$) because $f(x)$ has larger degree than $r(x)$.

The addition and multiplication are defined as in the ordinary clock arithmetic (and are shown to be well-defined in the same way, see §8). As in the ordinary (integer) clock arithmetic, if $[r(x)]$ is a non-zero remainder polynomial and $f(x)$ is **prime**, then 1 is a gcd of $f(x)$ and $r(x)$, and we can solve:

$$1 = r(x)u(x) + f(x)v(x)$$

and then $[u(x)]$ is the multiplicative inverse of $[r(x)]$.

Example: We saw that $x^2 + x + 1 \in F_2[x]$ is prime. From this, we get $\{[1], [x]\}$ as the basis of the polynomial clock defined by x^2+x+1, which is a vector space over F_2 of dimension 2 **and** a field with 4 elements (removing the cumbersome brackets):

$$0, 1, x, x + 1$$

Let's write down the multiplication and addition laws for this field. Notice that this is **not** $\mathbb{Z}_4$ ($\mathbb{Z}_4$ isn't a field!). We'll call this field F_4:

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\times$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	x

Next recall that an algebraic number α is a complex root of a prime polynomial:

$$f(x) = x^d + a_{d-1}x^{d-1} + ... + a_d \in \mathbb{Q}[x]$$

We claim next that via α, the polynomial $f(x)$-clock can be regarded as a **subfield** of the field $\mathbb{C}$ of complex numbers. In fact:

Proposition 3.1.3. *Suppose $F \subset \mathbb{C}$ is a subfield and $\alpha \in \mathbb{C}$ is a root of a prime polynomial:*

$$f(x) = x^d + a_{d-1}x^{d-1} + ... + a_0 \in F[x]$$

Then the $f(x)$-clock becomes a subfield of $\mathbb{C}$ when we set $[x] = \alpha$. This subfield is always denoted by $F(\alpha)$, and it sits between F and $\mathbb{C}$:

$$F \subset F(\alpha) \subset \mathbb{C}$$

Proof: The $f(x)$-clock is set up so that:

$$[x]^d + a_{d-1}[x]^{d-1} + \cdots a_0 = 0$$

But if $\alpha \in \mathbb{C}$ is a root of $f(x)$, then it is also true that

$$\alpha^d + a_{d-1}\alpha^{d-1} + \cdots a_0 = 0$$

so setting $[x] = \alpha$ is a well-defined substitution, and because $f(x)$ is prime, it follows that the clock becomes a subfield of $\mathbb{C}$.

Examples: We can give multiplication tables for clocks by just telling how to multiply the basis elements of the vector spaces:

(a) $F = \mathbb{R}$ and $f(x) = x^2 + 1$. The $x^2 + 1$-clock has table:

$\times$	1	x
1	1	x
x	x	-1

On the other hand, $\mathbb{R}(i)$ and $\mathbb{R}(-i)$ have multiplciation tables:

$\times$	1	i
1	1	i
i	i	-1

and

$\times$	1	$-i$
1	1	$-i$
$-i$	$-i$	-1

Both $\mathbb{R}(i)$ and $\mathbb{R}(-i)$ are, in fact, **equal** to $\mathbb{C}$. The only difference is in the basis as a vector space over $\mathbb{R}$. One basis uses i and the other uses its complex conjugate $-i$.

(b) If $F = \mathbb{Q}$ and $f(x) = x^3 - 2$, the clock has multiplication table:

$\times$	1	x	x^2
1	1	x	x^2
x	x	x^2	2
x^2	x^2	2	$2x$

and $\mathbb{Q}(\sqrt[3]{2})$ (necessarily) has the same multiplication table:

$\times$	1	$\sqrt[3]{2}$	$\sqrt[3]{4}$
1	1	$\sqrt[3]{2}$	$\sqrt[3]{4}$
$\sqrt[3]{2}$	$\sqrt[3]{2}$	$\sqrt[3]{4}$	$\sqrt[3]{8} = 2$
$\sqrt[3]{4}$	$\sqrt[3]{4}$	$\sqrt[3]{8} = 2$	$\sqrt[3]{16} = 2\sqrt[3]{2}$

To find, for example, the inverse of $x^2 + 1$ in the clock, we solve:

$$1 = (x^2 + 1)u(x) + (x^3 - 2)v(x)$$

which we do, as usual, using Euclid's algorithm:

$$\begin{array}{lclcl} x^3 - 2 & = & (x^2 + 1)x & + & (-x - 2) \\ x^2 + 1 & = & (-x - 2)(-x + 2) & + & 5 \end{array}$$

so, solving back up Euclid's algorithm:

$$\begin{array}{lclcl} 5 & = & (x^2 + 1) & - & (-x - 2)(-x + 2) \\ & = & (x^2 + 1) & - & \big((x^3 - 2) - (x^2 + 1)x\big)(-x + 2) \\ & = & (x^2 + 1)(-x^2 + 2x + 1) & + & (x^3 - 2)(x - 2) \end{array}$$

giving us the inverse in the $x^3 - 2$-clock:

$$(x^2 + 1)^{-1} = \frac{1}{5}(-x^2 + 2x + 1)$$

which we can substitute $x = \sqrt[3]{2}$ to get the inverse in $\mathbb{Q}(\sqrt[3]{2})$:

$$(\sqrt[3]{4} + 1)^{-1} = \frac{1}{5}(-\sqrt[3]{4} + 2\sqrt[3]{2} + 1)$$

Definition: A **linear transformation** of a vector space is a function:

$$T : V \to V$$

such that:

$$T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w}) \quad \text{and} \quad T(k\vec{v}) = kT(\vec{v})$$

for all vectors $\vec{v}, \vec{w}$ and all scalars k. The linear transformation is **invertible** if there is an inverse function $T^{-1} : V \to V$, which is then automatically **also** a linear transformation!

Definition: Given a vector space V of dimension n with a basis $\{\vec{v}_i\}$ and a linear transformation $T : V \to V$, the associated $n \times n$ **matrix**

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

is defined by:

$$T(\vec{v}_j) = a_{1j}\vec{v}_1 + a_{2j}\vec{v}_2 + ... + a_{nj}\vec{v}_n = \sum_{i=1}^{n} a_{ij}\vec{v}_i$$

Examples: (a) **Rotations in the $\mathbb{R}^2$ plane.** We start with the basis:

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and we want the matrix for $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by counterclockwise rotation by an angle of θ. For the matrix, use:

$$T(\vec{e}_1) = \cos(\theta)\vec{e}_1 + \sin(\theta)\vec{e}_2$$

by the definition of sin and cos. Since $\vec{e}_2$ can be thought of as $\vec{e}_1$ already rotated by $\frac{\pi}{2}$, we can think of $T(\vec{e}_2)$ as the rotation of $\vec{e}_1$ by $\frac{\pi}{2} + \theta$ so:

$$T(\vec{e}_2) = \cos(\frac{\pi}{2} + \theta)\vec{e}_1 + \sin(\frac{\pi}{2} + \theta)\vec{e}_2$$

and then the matrix for counterclockwise rotation by θ is:

$$A = \begin{bmatrix} \cos(\theta) & \cos(\frac{\pi}{2} + \theta) \\ \sin(\theta) & \sin(\frac{\pi}{2} + \theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

(using the identities: $\cos(\frac{\pi}{2} + \theta) = -\sin(\theta)$ and $\sin(\frac{\pi}{2} + \theta) = \cos(\theta)$)

(b) **Multiplication by a scalar.** If $k \in F$, let $T(\vec{v}) = k\vec{v}$, so:

$$T(\vec{v}_1) = k\vec{v}_1, ..., T(\vec{v}_n) = k\vec{v}_n$$

for any basis, and then:

$$A = \begin{bmatrix} k & 0 & \cdots & 0 \\ 0 & k & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & k \end{bmatrix}$$

In particular, the negation transformation is the case $k = -1$.

(c) **Multiplication by α.** If α has characteristic polynomial:

$$x^d + a_{d-1}x^{d-1} + ... + a_0 \in \mathbb{Q}[x]$$

then multiplication by α on the vector space $\mathbb{Q}(\alpha)$ is defined by:

$$T(1) = \alpha, T(\alpha) = \alpha^2, ..., T(\alpha^{d-1}) = \alpha^d = -a_0 - ... - a_{d-1}\alpha^{d-1}$$

giving us the matrix:

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ & & \vdots & & \\ 0 & 0 & \cdots & 1 & -a_{d-1} \end{bmatrix}$$

The fact that multiplication by α is a linear transformation comes from:

Proposition 3.1.4. *Multiplication by* **any** $\beta \in \mathbb{Q}(\alpha)$ *is linear.*

Proof: We need to show that $\beta(\vec{v} + \vec{w}) = \beta\vec{v} + \beta\vec{w}$ and $\beta(k\vec{v}) = k(\beta\vec{v})$. But in this vector space, all the vectors are **complex numbers!** For convenience set $\vec{v} = s$ and $\vec{w} = t$ to help us remember that they are numbers. Then:

$$\beta(s + t) = \beta s + \beta t$$

is the distributive law! And:

$$\beta(ks) = (\beta k)s = (k\beta)s = k(\beta s)$$

are the associative and commutative laws for multiplication.

Matrix multiplication (of matrices $A = (a_{ij})$ and $B = (b_{jk})$) is given by the prescription:

$$AB = C \text{ for } c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + ... + a_{in}b_{nk} = \sum_j a_{ij}b_{jk}$$

Fix a basis $\{\vec{v}_i\}$ for V. If the matrices A and B are associated to the linear transformations S and T, respectively, and if $U = S \circ T$, then:

$$U(\vec{v}_k) = S(T(\vec{v}_k)) = S(\sum_j b_{jk}\vec{v}_j) = \sum_{i,j} a_{ij}b_{jk}\vec{v}_i = \sum_i c_{ik}\vec{v}_i$$

is the kth column of C. So the product of two matrices is the matrix of the composition of the linear transformations.

We see from this that **matrix multiplication is associative:**

$$(AB)C = A(BC)$$

since composition of functions is associative:

$$(R \circ S) \circ T = R \circ S \circ T = R \circ (S \circ T)$$

Composition of linear transformations often isn't commutative, so matrix multiplication often isn't commutative (but sometimes it is!).

The identity transformation corresponds to the **identity matrix**:

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

which is a (multiplicative) identity, since $I_n A = A = AI_n$ for all A. So I_n commutes with all matrices! In fact, multiplication by any scalar commutes with all matrices, by definition of a linear transformation.

If T is an **invertible** linear transformation with matrix A, then the matrix A^{-1} associated to T^{-1} is the (two-sided) **inverse matrix** because the inverse function is always a two-sided inverse! In other words, the inverse matrix satisfies:

$$AA^{-1} = I_n = A^{-1}A$$

(so A commutes with its inverse matrix, whenever an inverse exists!)

Examples: (a) The matrices for rotations by θ and ψ are:

$$A_\theta = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \text{ and } A_\psi = \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix}$$

The product of the two matrices is:

$$A_\theta A_\psi = \begin{bmatrix} \cos(\theta)\cos(\psi) - \sin(\theta)\sin(\psi) & -\cos(\theta)\sin(\psi) - \sin(\theta)\cos(\psi) \\ \cos(\theta)\cos(\psi) - \sin(\theta)\sin(\psi) & -\sin(\theta)\sin(\psi) + \cos(\theta)\cos(\psi) \end{bmatrix}$$

and by the angle sum formula from trig (see also §4) this is $A_{\theta+\psi}$, which is, as it must be, the matrix associated to the rotation by $\theta + \psi$. Notice that here, too, the matrix multiplication **is** commutative, since $\theta + \psi = \psi + \theta$!

(b) We saw in an earlier example that in $\mathbb{Q}(\sqrt[3]{2})$, there is an equality:

$$(\sqrt[3]{4} + 1)(-\sqrt[3]{4} + 2\sqrt[3]{2} + 1) = 5$$

Let's check this out with matrix multiplication. Start with:

$$A = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 1 & 0 & 0 \end{bmatrix}$$

(the matrices for multiplication by $\sqrt[3]{2}$ and $\sqrt[3]{4}$, respectively)

The matrices for multiplication by $\sqrt[3]{4}+1$ and $-\sqrt[3]{4}+2\sqrt[3]{2}+1$ are:

$$A^2+I_3 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}, \quad -A^2+2A+I_3 = \begin{bmatrix} 1 & -2 & 4 \\ 2 & 1 & -2 \\ -1 & 2 & 1 \end{bmatrix}$$

and then the matrix version of the equality above is:

$$(A^2+I_3)(-A^2+2A+I_3) = 5I_3$$

as you may directly check with matrix multiplication!

Recall some more basic concepts from linear algebra:

Similarity: Two $n \times n$ matrices A and A' are **similar** if

$$B^{-1}AB = A'$$

for some invertible matrix B. This is an **equivalence relation:**

(i) Reflexive: $I_n^{-1}AI_n = A$

(ii) Symmetric: If $B^{-1}AB = A'$, then $(B^{-1})^{-1}A'B^{-1} = A$.

(iii) Transitive: If $B^{-1}AB = A'$ and $C^{-1}A'C = A''$, then:

$$A'' = C^{-1}(B^{-1}AB)C = (BC)^{-1}A(BC)$$

Note: Similarity occurs when we change basis. If A is the matrix for a transformation T with basis $\{\vec{v}_i\}$ and if $\{\vec{w}_j\}$ is another basis with:

$$\vec{w}_j = b_{1j}\vec{v}_1 + b_{2j}\vec{v}_2 + ... + b_{nj}\vec{v}_n$$

then $A' = B^{-1}AB$ is the matrix for T with the basis $\{\vec{w}_j\}$.

Determinant: The determinant is the unique function:

$$\det : \text{square matrices} \to F$$

that satisfies the following properties:

(i) $\det(AB) = \det(A)\det(B)$ for square $n \times n$ matrices A and B.

(ii) $\det(A) = 0$ if and only if A is not invertible.

(iii) The determinants of the "basic" matrices satisfy:

(a) $\det(A) = -1$ when A transposes two basis vectors $\vec{v}_i$ and $\vec{v}_j$:

$$T(\vec{v}_i) = \vec{v}_j, T(\vec{v}_j) = \vec{v}_i, \quad \text{otherwise } T(\vec{v}_l) = \vec{v}_l$$

(b) $\det(A) = 1$ when A adds a multiple of one basis vector to another:

$$T(\vec{v}_j) = \vec{v}_j + k\vec{v}_i, \quad \text{otherwise } T(\vec{v}_l) = \vec{v}_l$$

(c) $\det(A) = k$ when A multiplies one basis vector by k:

$$T(\vec{v}_i) = k\vec{v}_i \text{ and otherwise } T(\vec{v}_l) = \vec{v}_l$$

Example: The basic 2×2 matrices are:

$$\det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1$$

$$\det \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} = 1, \quad \det \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} = 1$$

$$\det \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} = k, \quad \det \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} = k$$

Since each matrix is a product of basic matrices (Gaussian elimination!) the determinant is completely determined by property (iii).

Note: $\det(B^{-1})\det(B) = \det(I_n) = 1$ when B is invertible, and

$$\det(A') = \det(B^{-1})\det(A)\det(B) = \det(B)^{-1}\det(A)\det(B) = \det(A)$$

when $A' = B^{-1}AB$, so the determinants of similar matrices are equal. Thus the determinant **doesn't care** about the choice of basis.

Characteristic Polynomial: This is the function:

$$ch : \text{square matrices} \to F[x]$$

defined by: $ch(A) = \det(xI_n - A)$ (assuming A is an $n \times n$ matrix). And the characteristic polynomial is the same for similar matrices, too:

$$ch(A') = \det(xI_n - B^{-1}AB) = \det(B^{-1}(xI_n - A)B) = \det(xI_n - A) = ch(A)$$

Examples: (a) The characteristic polynomial of rotation by θ:

$$\det \begin{bmatrix} x - \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & x - \cos(\theta) \end{bmatrix} = x^2 - 2\cos(\theta)x + 1$$

and the roots of this polynomial are the two complex numbers:

$$e^{i\theta} = \cos(\theta) + \sin(\theta)i \quad \text{and} \quad e^{-i\theta} = \cos(\theta) - \sin(\theta)i$$

(b) The characteristic polynomial of multiplication by $\alpha \in \mathbb{Q}(\alpha)$ is:

$$\det \begin{bmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \cdots & 0 & a_1 \\ 0 & -1 & \cdots & 0 & a_2 \\ & & \vdots & & \\ 0 & 0 & \cdots & -1 & x + a_{d-1} \end{bmatrix} = x^d + a_{d-1}x^{d-1} + ... + a_0$$

which is exactly the **same** as the characteristic polynomial of α thought of as an algebraic number! This apparent coincidence is explained by the following:

Proposition 3.1.5. *Each $n \times n$ matrix A is a "root" of its characteristic polynomial. That is, if*

$$ch(A) = x^n + a_{n-1}x^{n-1} + ... + a_0$$

then

$$A^n + a_{n-1}A^{n-1} + ... + a_0 I_n = 0$$

(this isn't a root in our usual sense, because A is a matrix, not a scalar!)

Proof: The sum:

$$B = A^n + a_{n-1}A^{n-1} + ... + a_0 I_n$$

is a **matrix**, so to see that it is zero, we need to see that it is the zero linear transformation, which is to say that $B\vec{v} = 0$ for all vectors $\vec{v} \in V$. In fact, it is enough to see that $B\vec{v}_i = 0$ for all basis vectors, but in this case it isn't helpful to restrict our attention to basis vectors.

So given an arbitrary vector $\vec{v}$, we know that eventually the vectors:

$$\vec{v}, A\vec{v}, A^2\vec{v}, ..., A^m\vec{v}$$

are linearly dependent (though we may have to wait until $m = n$). For the first such m, the vector $A^m\vec{v}$ is a linear combination of the others (which are linearly independent):

$$b_0\vec{v} + b_1 A\vec{v} + ... + b_{m-1}A^{m-1}\vec{v} + A^m\vec{v} = 0$$

Now I claim that the polynomial $x^m + b_{m-1}x^{m-1} + ... + b_0$ divides $ch(A)$. To see this, we extend $\vec{v}, ..., A^{m-1}\vec{v}$ to a basis of the vector space V:

$$\vec{v}, A\vec{v}, ..., A^{m-1}\vec{v}, \vec{w}_{m+1}, ..., \vec{w}_n$$

with some extra vectors $\vec{w}_{m+1}, ..., \vec{w}_n$ that I don't care about. The characteristic polynomial doesn't care what basis we use, so let's use this one. The point is that some of this matrix we know:

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -b_0 & * & \cdots & * \\ 1 & 0 & \cdots & 0 & -b_1 & * & \cdots & * \\ 0 & 1 & \cdots & 0 & -b_2 & * & \cdots & * \\ & & \vdots & & & & \vdots & \\ 0 & 0 & \cdots & 1 & -b_{m-1} & * & \cdots & * \\ 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ & & \vdots & & & & \vdots & \\ 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \end{bmatrix}$$

where the "$*$" denote entries that we do not know, since they involve the $\vec{w}_i$ basis vectors. But this is enough. It follows as in Example (b) above that $x^m + b_{m-1}x^{m-1} + ... + b_0$ divides the determinant of $xI_n - A$!

But now that $ch(A)$ factors, we can write

$$ch(A) = (x^{n-m} + c_{n-m-1}x^{n-m-1} + ... + c_0)(x^m + b_{m-1}x^{m-1} + ... + b_0)$$

for some other polynomial with c coefficients, and then:

$$B\vec{v} = (A^{n-m} + c_{n-m-1}A^{n-m-1} + ... + c_0 I_n)(A^m + b_{m-1}A^{m-1} + ... + b_0 I_n)\vec{v} = 0$$

because $A^m\vec{v} = -b_0\vec{v} - \cdots - b_{m-1}A^{m-1}\vec{v}$. That's the proof!

Final Remarks: Given an $n \times n$ matrix A, then any vector satsifying:

$$A\vec{v} = \lambda\vec{v}$$

is an **eigenvector** of the linear transformation and λ is its **eigenvalue**. If $\vec{v}$ is a nonzero eigenvector, then

$$(\lambda I_n - A)\vec{v} = 0$$

so in particular, $\lambda I_n - A$ is **not** an invertible matrix, and so:

$$\det(\lambda I_n - A) = 0$$

In other words, an eigenvalue is a **root** of the characteristic polynomial, and conversely, each root is an eigenvalue for some eigenvector. Notice that if the vector space happens to have a **basis** $\{\vec{v}_i\}$ of eigenvectors with eigenvalues $\{\lambda_i\}$, then by changing to this basis, we get a matrix A' similar to A with:

$$A' = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \vdots & \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

In this case A is said to be **diagonalizable**.

Example: Rotation by θ is not diagonalizable if $\mathbb{R}$ is our scalar field, since the eigenvalues for rotation are the complex numbers $e^{i\theta}$ and $e^{-i\theta}$. However, if we broaden our horizons and allow $\mathbb{C}$ to be the scalar field, then:

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} \cos(\theta) + i\sin(\theta) \\ \sin(\theta) - i\cos(\theta) \end{bmatrix} = e^{i\theta} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

and

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} \cos(\theta) - i\cos(\theta) \\ \sin(\theta) + i\cos(\theta) \end{bmatrix} = e^{-i\theta} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

so we have our basis of eigenvectors and in that basis, rotation is given by the matrix:

$$\begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix}$$

3.1.1 Linear Algebra Exercises

10-1 Recall that the polynomial $f(x) = x^3 + x + 1 \in F_2[x]$ is prime. This means that the $f(x)$-clock is a field with 8 elements. Complete the following addition and multiplication tables for this field:

$+$	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
0								
1								
x								
$x+1$								
x^2								
x^2+1								
x^2+x								
x^2+x+1								

$\times$	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
0								
1								
x								
$x+1$								
x^2								
x^2+1								
x^2+x								
x^2+x+1								

10-2 Repeat 10-1 for the prime polynomial $f(x) = x^2 + 1 \in F_3[x]$. Hint: This time you'll get a field with 9 elements!

10-3 In the field $\mathbb{Q}(\sqrt{2})$ do the following:

(a) Find the multiplicative inverse of $1 + \sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$.

(b) Write the 2×2 matrix for multiplication by $1 + \sqrt{2}$ in $\mathbb{Q}(\sqrt{2})$.

(c) Find the characteristic polynomial for the matrix in (b).

(d) Find the (complex!) eigenvalues of the matrix in (b).

(e) Find the 2×2 matrix for multiplication by $(1 + \sqrt{2})^{-1}$ in $\mathbb{Q}(\sqrt{2})$.

(f) Multiply the matrices (for $1 + \sqrt{2}$ and for $(1 + \sqrt{2})^{-1}$) to see that they are really inverses of each other.

10-4 Let $\alpha = \cos(\frac{2\pi}{5}) + i\sin(\frac{2\pi}{5})$. In the field $\mathbb{Q}(\alpha)$ do the following:

(a) Find the characteristic polynomial of the algebraic number α. (Hint: It is a polynomial of degree 4).

(b) Fill out the following multiplication table for $\mathbb{Q}(\alpha)$:

$\times$	1	α	α^2	α^3
1				
α				
α^2				
α^3				

(c) Find the multiplicative inverse of α^2 in $\mathbb{Q}(\alpha)$.

(d) Write the 4×4 matrix for multiplication by α^2.

10-5 Find the characteristic polynomials and eigenvalues of the following:

(a)

$$\begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}$$

(b)

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

(c)

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