LINEAR ALGEBRA: INVARIANT SUBSPACES

PETE L. CLARK

Contents

Let $F[t]$ be the ring of polynomials in one indeterminate, with coefficients in F.

Introduction

We give a treatment of the theory of invariant subspaces for an endomorphism of a vector space, up to and including the rational and Jordan canonical forms. Our approach should be suitable for students of mathematics at the advanced undergraduate level or beyond, although those who are sufficiently far beyond will find certain aspects of our treatment a bit pedestrian. Here are some features:

• We do not impose the condition of finite-dimensionality at the very beginning, but only towards the end of §1. Frankly this is pedagogically dubious – with the single possible exception of Theorem 1.2, we have no significant results to offer in the infinite-dimensional case – but we were unable to resist the lure of developing certain basic definitions in their "natural generality".

• We make use of quotient vector spaces. This places our exposition beyond the level of most first courses in linear algebra. But the gains in efficiency and simplicity from making use of this technique are considerable. At one point we (momentarily) performed the exercise of taking a proof using quotient spaces and rewriting it without using them: it became longer and – worse! – considerably more obscure.

• We work over an arbitrary ground field F and develop most of the theory without making reference to the algebraic closure of F. Until §8 we take the perspective that a linear endomorphism need not have any eigenvalues at all. Thus for us it is important that the minimal polynomial need not split into linear factors. However, the split case is slightly easier and will be more familiar to many readers, so we give many of the main results in the split case first before pursuing the general case, even though it would be more efficient to develop the general case first and then deduce the results in the split case (e.g. triangularization) as rather degenerate cases of more general theorems.

• At the heart of this exposition lie **prime vectors** and **primary vectors**. These are the appropriate analogues of eigenvectors and generalized eigenvectors over an arbitrary ground field.

• We introduce early on the **local minimal polynomials** P_v along with the "global" minimal polynomial P.

• We make a distinction between the characteristic polynomial and the Cayley-Hamilton polynomial (though they turn out to be equal). Here we were inspired by Axler's text [A], which takes the philosophy that defining the eigenvalues in terms of $\det(t - T)$ introduces unnecessary opacity. Rather, our notion of the characteristic polynomial is essentially that of the characteristic ideal of a finite-length module over an integral domain, namely the product of all the maximal ideals $\mathfrak{m}_i$ such that the Jordan-Hölder factors are the $R/\mathfrak{m}_i$. We introduce what is essentially the Jordan-Hölder uniqueness theorem in this context, but with a proof which is simpler than the one needed for a finite-length module over an arbitrary ring.

• We *do not* take an explicitly module-theoretic perspective: e.g. we do not switch from V to a different, but isomorphic, $F[t]$-module, even what that would simplify matters: e.g. we do not identify the underlying space of a cyclic endomorphism with $F[t]/(p(t))$. On the other hand, the techniques used here could be used, with only minor modifications, to classify finite length modules over any PID.

• At the end of the notes we include a discussion of the induced endomorphism T^* on the dual space V^* and show that it is similar to T.

• There are a number of exercises posed for the reader. They are meant to be closely related to the development of the material in the various sections, which is why they appear intertextually and not collected at the end. However almost all of the results stated here are given complete proofs, as I feel that this burden should fall on the writer/lecturer/instructor rather than the student.

1. Invariant Subspaces

Let V be a nonzero F-vector space. Let $T \in \operatorname{End} V$ be a linear endomorphism of V. A **T-invariant subspace** of V is a subspace $W \subset V$ such that $T(W) \subset W$. Actually though we will just say "invariant subspace": throughout these notes we work with only one endomorphism at a time, so the dependence on T in the terminology and notation will usually be suppressed.

Remark 1.1. *The invariant subspaces are precisely the subspaces W of V for which it makes sense to* restrict *T to an endomorphism of W. This already gives some insight into their importance.*

$\{0\}$ and V are invariant subspaces. We call them **trivial** and look for others.

We say V is **simple** if it has no nontrivial invariant subspaces. We say V is **semisimple** if it is a direct sum of simple invariant subspaces. We say V is **diagonalizable** if there is a basis $\{e_i\}_{i \in I}$ such that for all $i \in I$, $Te_i \in \langle e_i \rangle$: equivalently, V is a direct sum of one-dimensional invariant subspaces. Thus diagonalizable implies semisimple.

Theorem 1.2. *The following are equivalent:*
(i) V is semisimple.
(ii) If $W \subset V$ is an invariant subspace, it has an invariant complement: i.e., there is an invariant subspace W' such that $V = W \oplus W'$.
(iii) V is spanned by its simple invariant subspaces.

Proof. Three times in the following argument we assert the existence of invariant subspaces of V which are *maximal* with respect to a certain property. When V is finite-dimensional it doesn't matter what this property is: one cannot have an infinite, strictly ascending chain of subspaces of a finite-dimensional vector space. In the general case the claimed maximality follows from **Zorn's Lemma**, which we will not rehearse but rather trust that readers sufficiently sophisticated to care about the infinite-dimensional case will know what this is and how to use it here.
(i) $\implies$ (ii): Suppose $V = \bigoplus_{i \in I} S_i$, with each S_i a simple invariant. For each $J \subset I$, put $V_J = \bigoplus_{i \in J} S_i$. Now let W be an invariant subspace of V. There is a maximal subset $J \subset I$ such that $W \cap V_J = 0$. For $i \notin J$ we have $(V_J \oplus S_i) \cap W \neq 0$, so choose $0 \neq x = y + z$, $x \in W$, $y \in V_J$, $z \in S_i$. Then $z = x - y \in (V_j + W) \cap S_i$, and if $z = 0$, then $x = y \in W \cap V_j = 0$, contradiction. So $(V_J \oplus W) \cap S_i \neq 0$. Since S_i is simple, this forces $S_i \subset V_J \oplus W$. It follows that $V = V_J \oplus W$.
(ii) $\implies$ (i): The hypothesis on V passes to all invariant subspaces of V. We CLAIM that every nonzero invariant subspace $C \subset V$ contains a simple invariant subspace. PROOF OF CLAIM: Choose $0 \neq c \in C$, and let D be an invariant subspace of C that is maximal with respect to not containing c. By the observation of the previous paragraph, we may write $C = D \oplus E$. Then E is simple. Indeed, suppose not and let $0 \subsetneq F \subsetneq E$. Then $E = F \oplus G$ so $C = D \oplus F \oplus G$. If both $D \oplus F$ and $D \oplus G$ contained c, then $c \in (D \oplus F) \cap (D \oplus G) = D$, contradiction. So either $D \oplus F$ or $D \oplus G$ is a strictly larger invariant subspace of C than D which does not contain c, contradiction. So E is simple, establishing our claim. Now let $W \subset V$ be maximal with respect to being a direct sum of simple invariant subspaces, and write $V = W \oplus C$. If $C \neq 0$, then by the claim C contains a nonzero simple submodule, contradicting the maximality of W. Thus $C = 0$ and V is a direct sum of simple

invariant subspaces.
(i) $\implies$ (iii) is immediate.
(iii) $\implies$ (i): There is an invariant subspace W of V that is maximal with respect to being a direct sum of simple invariant subspaces. We must show $W = V$. If not, since V is assumed to be generated by its simple invariant subspaces, there exists a simple invariant subspace $S \subset V$ that is not contained in W. Since S is simple we have $S \cap W = 0$ and thus $W + S = W \oplus S$ is a strictly larger direct sum of simple invariant subspaces than W: contradiction. $\square$

What general methods do we have for producing invariant subspaces?

Proposition 1.3. *Both the kernel and image of T are invariant subspaces.*

Proof. If $v \in \operatorname{Ker} T$ then $Tv = 0$, and then $T(Tv) = T0 = 0$, so $Tv \in \operatorname{Ker} T$. As for the image $T(V)$ we have
$$T(T(V)) \subset T(V).$$
$\square$

As a modest generalization of the invariance of $T(V)$, we observe that if $W \subset V$ is an invariant subspace, then $T(W) \subset W$ so $T(T(W)) \subset T(W)$ and thus $T(W)$ is also an invariant subspace. It follows that $T^2(W) = T(T(W))$ is an invariant subspace, and so forth: we get a descending sequence of invariant subspaces:
$$W \supset T(W) \supset T^2(W) \supset \ldots \supset T^n(W) \supset \ldots.$$
If W is finite-dimensional, this sequence eventually stabilizes. In general it need not.

Similarly $\operatorname{Ker} T^k$ is an invariant subspace for all $k \in \mathbb{N}$, as is easily checked. This yields an increasing sequence of invariant subspaces
$$0 \subset \operatorname{Ker} T \subset \operatorname{Ker} T^2 \subset \ldots$$
A toothier generalization is the following.

Proposition 1.4. *Let $f(T) \in F[t]$. Then* $\operatorname{Ker} f(T)$ *and* $\operatorname{Image} f(T)$ *are invariant subspaces of V.*

Proof. Suppose $v, w \in \operatorname{Ker} f(T)$. Then for all $\alpha \in F$ we have
$$f(T)(\alpha v + w) = \alpha f(T)v + f(T)w = 0,$$
so $\operatorname{Ker} f(T)$ is a subspace. If $v \in \operatorname{Ker} f(T)$, then
$$f(T)(Tv) = T(f(T)v) = T0 = 0,$$
so $Tv \in \operatorname{Ker} f(T)$. The argument for $\operatorname{Image} f(T)$ is left to the reader. $\square$

This construction need not give all invariant subspaces. For instance, suppose $T = 1_V$ is the identity map. Then every $f(T)$ is a scalar map: if it is nonzero its kernel is $\{0\}$ and image is V; and if it is zero its kernel is V and its image is $\{0\}$. On the other hand in this case *every* subspace of V is invariant!

Lemma 1.5. *Let $W_1, W_2 \subset V$ be invariant subspaces. Then $W_1 + W_2$ and $W_1 \cap W_2$ are invariant subspaces.*

Exercise 1.1. *Prove Lemma 1.5.*

Let $v \in V$. The **orbit of T on V** is the set $\{T^k v\}_{k=0}^{\infty}$ (this is standard terminology whenever we have a mapping from a set to itself); the **linear orbit** $[v]$ of v is the subspace spanned by the orbit of T on v.

Proposition 1.6. *For any $v \in V$, the linear orbit $[v]$ of vis an invariant subspace of V. Moreover it is the minimal invariant subspace containing v: if $W \subset V$ is an invariant subspace and $v \in W$, then $[v] \subset W$.*

Exercise 1.2. *Prove Proposition 1.6.*

Exercise 1.3. *Let $S \subset V$ be any subset. Define the orbit of T on S as the union of the orbits of T on s for all $s \in S$. Define the linear orbit $[S]$ of T on S to be the span of the orbit of T on S. Show that $[S]$ is an invariant subspace of V and is minimal in the sense of Proposition 1.6 above.*

For $v \in V$, there is a natural linear map given by evaluation at v:

$$E_v : \operatorname{End} V \to V,\ A \mapsto Av.$$

For $T \in \operatorname{End} V$, there is a natural linear map given by evaluation at T:

$$E_T : F[t] \to \operatorname{End} V,\ p(t) \mapsto p(T).$$

Consider the composition of these maps:

$$\mathcal{E} = \mathcal{E}_{T,v} \coloneqq E_v \circ E_T : F[t] \to V,\ p \mapsto p(T)v.$$

Lemma 1.7. *a) The image $\mathcal{E}(F[t])$ of the map $\mathcal{E}$ is $[v]$, the linear orbit of v.*
b) The kernel $\mathcal{K}$ of $\mathcal{E}$ is an ideal of $F[t]$.

Proof. a) This follows immediately upon unwinding the definitions.
b) Since $\mathcal{E}$ is an F-linear map, its kernel is an F-subspace of $F[t]$. However $\mathcal{E}$ is not a ring homomorphism (e.g. because V is not a ring!) so we do need to check that $\mathcal{K}$ is an ideal. But no problem: suppose $p \in \mathcal{K}$ and $q \in F[t]$. Then

$$\mathcal{E}(qp) = (q(T)p(T))v = q(T)(p(T)v) = q(T)0 = 0. \qquad \square$$

Recall that every ideal I of $F[t]$ is principal: this is clear if $I = (0)$; otherwise I contains a monic polynomial $a(t)$ of least degree. Let $b(t) \in I$. By polynomial division, there are $q(t), r(t) \in F[t]$ with $\deg r < \deg a$ such that $b(t) = q(t)a(t) + r(t)$. But $r(t) = b(t) - q(t)a(t) \in I$. If $r(t) \neq 0$, then multiplying by the inverse of its leading coefficient, we would get a monic polynmial in I of degree smaller than that of $a(t)$, contradicting the definition of $a(t)$. So $r(t) = 0$ and $I = (a(t))$.

Consider the ideal $\mathcal{K}$ of $F[t]$ defined in Lemma 1.7. There is a clear dichotomy:

• Case 1: $\mathcal{K} = 0$. In this case $\mathcal{E} : F[t] \cong [v]$, so every invariant subspace of V containing v is infinite-dimensional. We put $P_v(t) = 0$ (the zero polynomial, which generates $\mathcal{K}$). In this case we say that the vector v is **transcendental**.

• Case 2: $\mathcal{K} = (P_v(t))$ for some nonzero monic $P_v(t)$. Then $\mathcal{E} : F[t]/\mathcal{K} \xrightarrow{\sim} [v]$, so $\deg P_v = \dim[v]$ are both finite. We say that the vector v is **algebraic**.

In either case we call P_v the **local minimal polynomial of T at v**, and we will now study it in detail.

Exercise 1.4. *Show that for $v \in V$, we have $v = 0 \iff P_v(t) = 1$.*

Exercise 1.5. *Let $v \in V$.*
a) Show that for all $w \in [v]$, we have $P_v(T)w = 0$.
b) Deduce that for all $w \in [v]$, we have $P_w(t) \mid P_v(t)$.

Lemma 1.8. *Let $v \in V$ be a transcendental vector.*
a) For every monic polynomial $f(t) \in F[t]$, $f(T)[v]$ is an invariant subspace of V.
b) For distinct monic polynomials f_1, f_2, the invariant subspaces $f_1(T)[v]$, $f_2(T)[v]$ are distinct. Thus $[v]$ has infinitely many invariant subspaces and is not simple.
c) Every nonzero invariant subspace of $[v]$ is of the form $f(T)[v]$ for some monic polynomial $f(t) \in F[t]$.

Proof. a) Apply Proposition 1.4 with $[v]$ in place of V.
b) We claim $f_1(T)v \in f_2(T)[v]$ if and only if $f_2(t) \mid f_1(t)$; if so, $f_1(T)[v] = f_2(T)[v]$ implies $f_1 \mid f_2$ and $f_2 \mid f_1$ so $f_1 = f_2$. If $f_2 \mid f_1$, write $f_1 = gf_2$ and then $f_1(T)v = g(T)f_2(T)v \in f_2(T)[v]$. Conversely, if $f_1(T)v \in f_2(T)[v]$, then there is a polynomial $g(t) \in F[t]$ such that $f_1(T)v = f_2(T)g(T)v$, so $(f_1(T) - f_2(T)g(T))v = 0$, and thus the local minimal polynomial of v divides $f_1(t) - f_2(t)g(t)$. But since v is transcendental, its local minimal polynomial is zero and thus $0 = f_1 - f_2 g$ and thus $f_2 \mid f_1$. The second sentence of part b) follows immediately.
c) Let $W \subset [v]$ be a nonzero invariant subspace. It therefore contains a nonzero vector, which may be written as $f(T)v$ for a monic polynomial $f \in F[t]$. Among all nonzero vectors choose one which may be written in this way with $f(t)$ of least degree: we claim $W = f(T)[v]$. Indeed, consider any nonzero $w = g(T)v \in W$. By polynomial division there are $q(t), r(t) \in F[t]$ with $\deg r < \deg f$ such that $g(t) = q(t)f(t) + r(t)$, and thus $w = q(T)f(T)v + r(T)v$. Then $r(T)v = w - q(T)f(T)v \in W$; since $\deg r < \deg q$ we get a contradiction unless $r = 0$, in which case $w = q(T)f(T)v = f(T)(q(T)v) \in f(T)[v]$. $\square$

Exercise 1.6. *Consider the linear map $F[t] \to \operatorname{End} V$ given by $p(t) \mapsto P(T)$.*
a) Show that its kernel $\mathcal{M}$ is an ideal of $F[t]$, and thus of the form $(P(t))$ where $P(t)$ is either monic or zero. It is called the ***minimal polynomial*** *of T on V.*
b) Show that for all $v \in V$, $P_v(t) \mid P(t)$.
c) Show that $P(t)$ is the least common multiple of $\{P_v(t)\}_{v \in V}$.
d) Suppose that $V = [v]$ for some $v \in v$. Show that $P(t) = P_v(t)$.

Exercise 1.7. *This exercise requires some familiarity with functional analysis.*
a) Suppose that $\dim V$ *is uncountable. Show that V has a nontrivial invariant subspace.*
b) Let V be an infinite-dimensional real Banach space. Show that V has a nontrivial invariant subspace.
c) Suppose that V is a real Banach space which is not separable. Show that V admits a nontrivial closed invariant subspace.
d) Prove or disprove: every bounded linear operator on a separable complex Hilbert space V of dimension greater than one has a nontrivial closed invariant subspace.
(Comment: This is one of the more famous open problems in all of mathematics. At least we will be able to handle the finite-dimensional case later on!)

We say that V is **locally algebraic** if each vector $v \in V$ is algebraic, i.e., that for all $v \in V$, the local minimal polynomial $P_v(t)$ is nonzero. We say that V is **algebraic** if the minimal polynomial P is nonzero.

Proposition 1.9. *a) If V is finite-dimensional, it is algebraic: $P_V \neq 0$.*
b) If V is algebraic, it is locally algebraic.

Proof. a) Let P be the minimal polynomial of V. By Exercise 1.6 we have an injection $F[t]/(P) \hookrightarrow \operatorname{End} V$. Since V is finite-dimensional, so is $\operatorname{End} V$, hence so

is $F[t]/(P)$, which implies $P \neq 0$.
b) This is immediate from the fact that $P_v \mid P$ for all $v \in V$. $\square$

Exercise 1.8. *a) Show that $P_{T(V)} \mid P_V$.*
b) Suppose that $T(V)$ is finite-dimensional. Show that P is algebraic.
c) Exhibit an algebraic endomorphism T with infinite-dimensional image.

Exercise 1.9. *Suppose $B = \{b_i\}_{i \in I}$ is a basis for V. For each $i \in I$, let $\lambda_i \in F$. Define T by $T(b_i) = \lambda_i b_i$ for all i.*
a) Compute the local minimal polynomials of all $v \in V$.
b) Deduce from part a) that V is locally algebraic.
c) Show that V is algebraic iff $\{\lambda_i\}_{i \in I}$ is finite.

Exercise 1.10. *Let $V = F[t]$; recall $\{t^n \mid n \in \mathbb{N}\}$ is a basis.*
Let $D \in \operatorname{End} V$ be polynomial differentiation. Concretely,

$$D(1) = 0;\ \forall n \in \mathbb{Z}^+, D : t^n \mapsto nt^{n-1}.$$

a) Compute the local minimal polynomials of all $v \in V$.
b) Show that V is locally algebraic but not algebraic.
c) Find all finite-dimensional invariant subspaces of V.

Exercise 1.11. *Let $V = F[t]$. Let $I \in \operatorname{End} V$ be polynomial integration with zero constant term. Concretely,*

$$\forall n \in \mathbb{N},\ I(t^n) = \frac{t^{n+1}}{n+1}.$$

a) Show that every nonzero $v \in V$ is a transcendental vector.
b) Deduce that V is not algebraic or locally algebraic.
c) Find all finite-dimensional invariant subspaces of V.

Exercise 1.12. *Suppose V is locally algebraic.*
a) Show that for all finite-dimensional subspaces $W \subset V$, $[W]$ is finite-dimensional.
b) For a subspace $W \subset V$, show that the following are equivalent:
(i) W is an invariant subspace.
(ii) W is union of finite-dimensional invariant subspaces.

Thus to find all invariant subspaces of a locally finite endomorphism it suffices to study finite-dimensional invariant subspaces.

From now on we assume that V is nonzero and finite-dimensional.

Proposition 1.10. *The degree of the minimal polynomial is at most* $\dim V$.

Proof. (Burrow [Bu73]) We go by induction on $\dim V$, the case $\dim V = 1$ being handled, for instance, by the bound $\dim P \leq (\dim V)^2$. Now let $\dim V = d$ and suppose the result holds in smaller dimension. Choose a nonzero $v \in V$, and let P_v be the local minimal polynomial, so $\deg P_v > 0$. Let $W = \operatorname{Ker} P_v(T)$, so that W is a nonzero invariant subspace. If $W = V$ then $P_v = P$ and we're done. Otherwise we consider the induced action of T on the quotient space V/W. Let P_W and $P_{V/W}$ be the minimal polynomials of T on W and V/W. By induction, $\deg P_W \leq \dim W$ and $\deg P_{V/W} \leq \dim V/W$, so $\deg P_W P_{V/W} = \deg P_W + \deg P_{V/W} \leq \dim W + \dim V/W = \dim V$. Finally, we claim that $P_W(T)P_{V/W}(T)V = 0$. Indeed, for all $v \in V$, $P_{V/W}(T)v \in W$ so $P_W(T)P_{V/W}(T)v = 0$. $\square$

For an invariant subspace $W \subset V$, we let P_W be the minimal polynomial of $P|_W$.

Exercise 1.13. *Let $W \subset V$ be an invariant subspace.*
a) Show that $P_W = \operatorname{lcm}_{w \in W} P_w$.
b) Show that for all $v \in V$, $P_{\langle v \rangle} = P_v$.

Proposition 1.11. *Let W_1, W_2 be invariant subspaces of V; put $W = W_1 + W_2$. Then*

$$P_W = \operatorname{lcm} P_{W_1}, P_{W_2}.$$

Proof. Put

$$P = P_{W_1+W_2},$$
$$P_1 = P_{W_1} = p_1^{a_1} \cdots p_r^{a_r}$$
$$P_2 = P_{W_2} = p_1^{b_1} \cdots p_r^{b_r}$$

with $a_i, b_i \in \mathbb{N}$, and then

$$P_3 = \operatorname{lcm} P_1, P_2 = p_1^{\max a_1, b_1} \cdots p_r^{\max a_r, b_r}.$$

We may write $P_3 = f_1 P_1 = f_2 P_2$. Then every vector $w \in W_1 + W_2$ is of the form $w = w_1 + w_2$ for $w_1 \in W_1$, $w_2 \in W_2$ and

$$P_3(T)w = P_3(T)w_1 + P_3(T)w_2 = f_1(T)P_1(T)w_1 + f_2(T)P_2(T)w_2 = 0,$$

so $P \mid P_3$. To show that $P_3 \mid P$, since $P = \operatorname{lcm}_{v \in W_1+W_2} P_v$ and $\operatorname{lcm} p_1^{c_1}, \ldots, p_r^{c_r} = p_1^{c_1} \cdots p_r^{c_r}$, it is enough to find for each $1 \leq i \leq r$ a vector $v_i \in W_1 + W_2$ such that $p_i^{\max a_i, b_i} \mid P_{v_i}$. But since $p_i^{a_i} \mid P_1$, there is $w_{i,1} \in W_1$ with $p_i^{a_i} \mid w_{i,1}$ and $w_{i,2} \in W_2$ with $p_i^{b_i} \mid w_{i,2}$. One of these vectors does the job. $\square$

For any polynomial $f \in F[t]$, put

$$V_f \coloneqq \{v \in V \mid f(T)v = 0\}.$$

Proposition 1.12. *Let W be an invariant subspace of V, and let $f \mid P_V$. Then*

$$W_f = W \cap V_f.$$

Proof. Although this is a distinctly useful result, its proof is absolutely trivial:

$$W_f = \{v \in W \mid f(T)v = 0\} = W \cap \{v \in V \mid f(T)v = 0\}. \qquad \square$$

Note that $V_0 = V$. Henceforth we restrict attention to nonzero polynomials.

Proposition 1.13. *Let $f \in F[t]^\bullet$.*
a) V_f is an invariant subspace of V.
b) V_f is the set of vectors v such that $P_v \mid f$.
c) If $f \mid g$, then $V_f \subset V_g$.
d) For $\alpha \in F^\times$, we have $V_{\alpha f} = V_f$.
e) We have $V_f = V_{\gcd P, f}$, where P is the minimal polynomial.

Proof. a) It is immediate to check that V_f is linear subspace of V. Further, if $f(T)v = 0$, then $f(T)(Tv) = T(f(T)v) = T0 = 0$.
b) This follows from the fact that P_v is the generator of the ideal of all polynomials g with $g(T)v = 0$.
c) If $f \mid g$ then $g = h(t)f(t)$, so if $f(T)v = 0$ then $g(T)v = (h(T)f(T))v = h(T)(f(T)v) = h(T)0 = 0$.
d) For any $v \in V$, $f(T)v = 0 \iff \alpha f(T)v = 0$.
e) Since $\gcd(P, f) \mid f$, $V_{\gcd(P,f)} \subset V_f$. Conversely, let $v \in V_f$. Then $P_v \mid P$ and $P_v \mid f$, so $P_v \mid \gcd(P, f)$, so $\gcd(P, f)(T)(v) = 0$ and $v \in V_{\gcd P, f}$. $\square$

In view of Proposition 1.13e) there are only finitely many distinct spaces V_f, since there are only finitely many monic polynomials dividing P.

If there is a vector $v \in V$ with $P = P_v$, we say that the minimal polynomial P is **locally attained**. Since it was immediate from the definition that $\deg P_v \leq \dim V$, if the minimal polynomial is locally attained then we get another, better, proof that $\deg P \leq \dim V$. The next exercise gives many cases in which the minimal polynomial is locally attained.

Exercise 1.14. *a) Show that for each proper divisor f of P, $V_f \subsetneq V$.*
b) Suppose F is infinite. Show that there is $v \in V$ with $P_v = P$: we say that ***the minimal polynomial is locally attained****. (Hint: no nonzero vector space over an infinite field is the union of finitely many of its proper subsapces: c.f.* [Cl12]*.)*
c) Use the main result of [Cl12] *to show that if F is finite but $\#F$ is sufficiently large with respect to $\dim V$, then the minimal polynomial is locally attained.*

Proposition 1.14. *For $n \geq 2$, let $f_1, \ldots, f_n \in F[t]$ be pairwise coprime. Then the subspaces $V_{f_1}, \ldots, V_{f_n}$ are independent and we have $\bigoplus_{i=1}^n V_{f_i} = V_{f_1 \cdots f_n}$.*

Proof. We go by induction on n.
Base Case ($n = 2$): let $v \in V_{f_1} \cap V_{f_2}$. Since f_1 and f_2 are coprime, there are $a(t), b(t) \in F[t]$ such that $af_1 + bf_2 = 1$, and then

$$v = 1v = (a(T)f_1(T) + b(T)f_2(T))v = a(T)(f_1(T)v) + b(T)(f_2(T)v)) = 0,$$

which shows that $W := V_{f_1} + V_{f_2} = V_{f_1} \oplus V_{f_2}$. It is easy to see that $W \subset V_{f_1 f_2}$: every $w \in W$ is a sum of a vector w_1 killed by $f_1(T)$ and a vector w_2 killed by $f_2(T)$, so $f_1(T)f_2(T)w = 0$. Conversely, let $v \in V$ be such that $f_1(T)f_2(T)v = 0$. As above, we have $v = a(T)f_1(T)v + b(T)f_2(T)v$. Then $a(T)f_1(T)v \in V_{f_2}$ and $b(T)f_2(T) \in V_{f_1}$, so $v \in V_{f_1} \oplus V_{f_2}$.
Induction Step: Suppose $n \geq 3$ and that the result holds for any $n-1$ pairwise coprime polynomials. Put $W = V_{f_1} + \ldots + V_{f_{n-1}}$. By induction,

$$W = \bigoplus_{i=1}^{n-1} V_{f_i} = V_{f_1 \cdots f_{n-1}}.$$

The polynomials $f_1 \cdots f_{n-1}$ and f_n are coprime, so applying the base case we get

$$W + V_{f_n} = \bigoplus_{i=1}^{n-1} V_{f_i} \oplus V_{f_n} = \bigoplus_{i=1}^{n} V_{f_i} = V_{f_1 \cdots f_n}.$$ □

Lemma 1.15. *Let $v \in V^\bullet$. For any monic polynomial $f \mid P_v$, we have*

$$P_{f(T)v} = \frac{P_v}{f}.$$

Proof. Write $P_v = fg$. Since $\frac{P_v}{f}(T)f(T)v = P_v(T)v = 0$, we have $P_{f(T)v} \mid \frac{P_v}{f}$. If there were a proper divisor h of g such that $h(T)(f(T)v) = 0$, then $hf(T)v = 0$. That is, hf kills v but has smaller degree than $gf = P_v$, contradiction. □

Exercise 1.15. *Show that for any $f \in F[t]$, we have $P_{f(T)v} = \frac{P_v}{\gcd P_v, f}$.*

Theorem 1.16 (Local Attainment Theorem)**.** *Every monic divisor f of the minimal polynomial is a local minimal polynomial: $f = P_v$ for some $v \in V$.*

Proof. Step 1: Let $P = p_1^{a_1} \cdots p_r^{a_r}$. Since P is the lcm of the local minimal polynomials, there is $w_i \in V$ such that the exponent of p_i in P_{w_i} is a_i. Let $v_i = \frac{P}{p_i^{a_i}}(T)w_i$. By Lemma 1.15, $P_{v_i} = p_i^{a_i}$.
Step 2: Put $v = v_1 + \ldots + v_r$. We claim that $P_v = P$. Indeed, since $p_1^{a_1}, \ldots, p_r^{a_r}$ are pairwise coprime, the spaces $V_{p_1^{a_1}}, \ldots, V_{p_r^{a_r}}$ are independent invariant subspaces. It follows that for all $f \in F[t]$, the vectors $f(T)v_1, \ldots, f(T)v_r$ are linearly independent. In particular, if $0 = f(T)v = f(T)v_1 + \ldots + f(T)v_r$, then $f(T)v_1 = \ldots = f(T)v_r = 0$. This last condition occurs iff $p_i^{a_i} \mid f$ for all i, and again by coprimality this gives $P = p_1^{a_1} \cdots p_r^{a_r} \mid f$.
Step 3: Now suppose that we have monic polynomials f, g with $fg = P$. By Step 2, there is $v \in V$ with $P_v = P$. By Lemma 1.15, $P_{g(T)v} = \frac{P}{g} = f$. $\square$

Let $W \subset V$ be an invariant subspace. Then T induces a linear endomorphism on the quotient space V/W given by $T(v + W) = T(v) + W$. Let's check that this is well-defined, i.e., that if $v' + W = v + W$, then $T(v) + W = T(v') + W$. There is $w \in W$ such that $v' = v + w$, so

$$T(v') + W = T(v + w) + W = T(v) + T(w) + W = T(v) + W,$$

since $T(W) \subset W$. We call V/W an **invariant quotient**.

Exercise 1.16. *Let $W \subset V$ be an invariant subspace. For $f(t) \in F[t]$, show that the following are equivalent:*
(i) $f(T)V \subset W$.
(ii) $f(T)V/W = 0$.
(iii) $P_{V/W} \mid f$.

Exercise 1.17. *Let W_1, W_2 be subspaces of V, with W_2 invariant. Define*

$$I_{W_1,W_2} = \{f \in F[t] \mid f(W_1) \subset W_2\}.$$

a) Show that I_{W_1,W_2} is an ideal of $F[t]$.
b) Show that if $W_1 \subset W_1'$ then $I_{W_1,W_2} \supset I_{W_1',W_2}$.
c) Show that if $W_2 \subset W_2'$ are invariant subspaces, then $I_{W_1,W_2} \subset I_{W_1,W_2'}$.
d) Deduce that $I_{W_1,W_2} \supset I_{V,0} = (P)$. In particular I_{W_1,W_2} is nonzero so has a unique monic generator P_{W_1,W_2}, the ***conductor polynomial*** *of W_1 into W_2.*
e) Show that these conductor polynomials recover as special cases: the minimal polynomial, the local minimal polynomials, the minimal polynomial of an invariant quotient, and the local minimal polynomials of an invariant quotient.

Proposition 1.17. *Let $W \subset V$ be an invariant subspace.*
a) For $v \in V$, let $\overline{v}$ be its image in V/W. Then $P_{\overline{v}} \mid P_v$.
b) For every $\overline{v} \in V/W$, there is $v' \in V$ such that $P_{\overline{v}} = P_{v'}$.
c) $P_{V/W} \mid P_V$.

Proof. a) Since $P(T)v = 0$, also $P(T)v \in W$; the latter means $P(T)\overline{v} = 0$.
b) Let v be any lift of $\overline{v}$ to V. By part a) we may write $P_v(t) = f(t)P_{\overline{v}}(t)$ for some polynomial f. By Lemma 1.15, $P_{f(T)v} = P_{\overline{v}}$.
c) Since $P_V(T)$ kills every vector of V, it sends every vector of V into W. (One could also use the characterizations of the global minimal polynomial as the lcm of the local minimal polynomials together with part b).) $\square$

Exercise 1.18. *This exercise is for those who are familiar with antitone Galois connections as in e.g. CITE. We consider the Galois connection between $2^{F[t]}$ and*

2^V induced by the relation $(f,v) \in R \subset F[t] \times V$ iff $f(T)v = 0$.
a) For a subset $S \subset F[t]$, put

$$W(S) := \{v \in V \mid \forall f \in S,\ f(T)v = 0\}.$$

Let $\langle S\rangle$ be the ideal generated by S. Since $F[t]$ is a PID, this ideal is principal, say generated by the monic polynomial f_S. Show:

$$W(S) = V_{f_S} = \{v \in V \mid f_S(T)v = 0\}.$$

Deduce from the formalism of Galois connections that for polynomials $f, g \in F[t]$, if $f \mid g$ then $V_f \subset V_g$.
b) For a subset $U \subset V$, put

$$I(U) := \{f \in F[t] \mid \forall v \in U, f(T)v = 0\}.$$

Let $\langle U\rangle$ be the subspace generated by U. Show:

$$I(U) = \langle P_{\langle U\rangle}\rangle.$$

Deduce from the formalism of Galois connections that for subspaces W_1, W_2 of V, if $W_1 \subset W_2$ then $P_{W_1} \mid P_{W_2}$.
c) For $S \subset F[t]$, show that the closure of S with respect to the Galois connection – i.e., $I(W(S))$ – is the ideal generated by $\gcd(f_S, P_V)$.
d) For $U \subset V$, the closure of U with respect to the Galois connection – i.e., $W(I(U))$ – is the the kernel of $P_{\langle U\rangle}(T)$. Is there a better description?

2. Eigenvectors, Eigenvalues and Eigenspaces

A nonzero vector $v \in V$ is an **eigenvector** for T if $Tv = \lambda v$ for some $\lambda \in F$, and we say that λ is the corresponding **eigenvalue**. A scalar $\lambda \in F$ is an **eigenvalue** of T if there is some nonzero vector $v \in v$ such that $Tv = v$.

In fact this is a special case of a concept studied in the last section. Namely, for $v \in V$, we have $Tv = \lambda v$ iff $(T - \lambda)v = 0$ iff $v \in V_{t-\lambda}$. Thus v is an eigenvector iff the local minimal polynomial P_v is linear.

Exercise 2.1. *Prove it.*

Proposition 2.1. *The following are equivalent:*
(i) 0 is an eigenvalue.
(ii) T is not *invertible.*

Exercise 2.2. *Prove it.*

Proposition 2.2. *Let $P(t)$ be the minimal polynomial for T on V.*
a) For $\lambda \in F$, the following are equivalent:
(i) λ is an eigenvalue of T.
(ii) $P(\lambda) = 0$.
b) It follows that T has at most $\dim V$ eigenvalues.

Proof. a) By Proposition 1.13e), we have $V_{t-\lambda} = V_{\gcd(t-\lambda,P)}$. It follows that if $P(\lambda) \neq 0$ then $t - \lambda \nmid P(t)$ and thus $V_{t-\lambda} = V_1 = \{0\}$, so λ is not an eigenvalue. If λ is an eigenvalue, there is $v \in V$ with $P_v = t - \lambda$. Since $P_v = t - \lambda \mid P$, so $P(\lambda) = 0$. b) By Proposition 1.9, $P \neq 0$, and by Proposition 1.11, $\deg P \leq \dim V$, so P has at most $\dim V$ roots. $\square$

Exercise 2.3. *Show that the following are equivalent:*
(i) Every subspace of V is invariant.
(ii) Every nonzero vector of V is an eigenvector.
(iii) The minimal polynomial P has degree 1 and for all $v \in V^\bullet$, $P_v = P$.
(iv) There is $\alpha \in F$ such that $T(v) = \alpha v$ for all $v \in V$.
(v) The matrix representation of T (with respect to any basis!) is a scalar matrix.

Corollary 2.3. *If F is algebraically closed, then there is an eigenvector for T.*

Proof. Indeed P is a polynomial of positive degree, so by the very definition of being algebraically closed, there is $\alpha \in F$ with $P(\alpha) = 0$. Apply Proposition 2.2. $\square$

For $\lambda \in V$, the **λ-eigenspace of V** is

$$V_\lambda = \{v \in V \mid Tv = \lambda v\}.$$

In fact the eigenspace V_λ is precisely the subspace $V_{t-\lambda}$. In particular it is an invariant subspace.

Exercise 2.4. *Let $\lambda \in F$. Show that λ is an eigenvalue iff $V_\lambda \neq \{0\}$.*

Proposition 2.4. *Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues for T on V. Then the eigenspaces $V_{\lambda_1}, \ldots, V_{\lambda_n}$ are independent.*

Proof. Because $V_{\lambda_i} = V_{t-\lambda_i}$, this is a special case of Proposition 1.14. $\square$

Corollary 2.5. *The following are equivalent:*
(i) T is diagonalizable.
(ii) V is the sum of its nonzero eigenspaces V_λ.
(iii) V is the direct sum of its nonzero eigenspaces V_λ.

Exercise 2.5. *Prove Corollary 2.5.*

Corollary 2.6. *a) If T is diagonalizable, it has an eigenvalue.*
b) If T has exactly one eigenvalue α, the following are equivalent:
(i) T is diagonalizable.
(ii) $Tv = \alpha v$ for all $v \in V$.
(iii) The minimal polynomial of T is $(t - \alpha)$.
(iv) The matrix of T with respect to any basis is the scalar matrix αI.
c) If T has $\dim V$ distinct eigenvalues, it is diagonalizable.

Exercise 2.6. *Prove Corollary 2.6.*

Proposition 2.7. *If T is diagonalizable with distinct eigenvalues $\lambda_1, \ldots, \lambda_d$ then the minimal polynomial P is $\prod_{i=1}^d (t - \lambda_i)$. In particular P is squarefree and split.*

Proof. If T is diagonalizable then there is a basis $\{e_i\}_{i=1}^n$ consisting of eigenvectors. Then if $p(t) := \prod_{i=1}^d (t - \lambda_i)$, we have that $p(T)e_i = 0$ for all i and thus $p(T) = 0$. It follows that $P(t) \mid p(t)$. We have already seen that every eigenvalue is a root of P, so also $p(t) \mid P(T)$ and thus $P(t) = p(t)$. $\square$

It is natural to ask about the converse of Proposition 2.7. In fact it is true: if P is squarefree and split then T is diagonalizable – this is one of the main results of the theory. But the proof must lie deeper than anything we've done so far. To see why, suppose $\dim V = 3$ and $P(t) = (t - \lambda_1)(t - \lambda_2)$ for $\lambda_1 \neq \lambda_2$. Since P is squarefree and split, we'd like to show that T is diagonalizable. We know that λ_1 and λ_2 are the only eigenvalues and that $\dim V_{\lambda_1} \geq 1$ and $\dim V_{\lambda_2} \geq 1$. So to have a basis of eigenvectors it must be the case that either $\dim V_{\lambda_1} = 2$ or $\dim V_{\lambda_2} = 2$. But since $\dim V = 3$, it can't be the case that *both* eigenspaces are two-dimensional: it must be one or the other. Clearly we won't figure out which by staring at $P(t) = (t - \lambda_1)(t - \lambda_2)$: we are not yet seeing the full picture.

A basis $e_1, \ldots, e_n$ for V is **triangular** if for all $1 \leq i \leq n$, $Te_i \in \langle e_1, \ldots, e_i \rangle$. Equivalently, for all $1 \leq i \leq n$, the subspace $V_i = \langle e_1, \ldots, e_i \rangle$ is invariant. We say that T is **triangularizable** if it admits a triangular basis.

Exercise 2.7. *Show that the matrix M of T with respect to a triangular basis is upper triangular: $m_{ij} = 0$ for all $i > j$.*

The careful reader will have noticed that we have not used the observation that the *image* of any $f(T)$ is an invariant subspace. We do so now: it is one of the key ideas in the following proof.

Theorem 2.8. *Suppose V admits a triangular basis $e_1, \ldots, e_n$ with respect to T. Let $m_{11}, \ldots, m_{nn}$ be the diagonal entries of the corresponding matrix.*
a) T is invertible iff $m_{ii} \neq 0$ for all $1 \leq i \leq n$.
b) Each m_{ii} is an eigenvalue; each eigenvalue λ is equal to m_{ii} for at least one i.

Proof. a) If each diagonal entry is nonzero, it is easy to see by back substitution that the only solution of the linear system $Mv = 0$ is $v = 0$, so T is invertible. Conversely, if some $m_{ii} = 0$, then $T : \langle e_1, \ldots, e_i \rangle \to \langle e_1, \ldots, e_{i-1} \rangle$, so by the Dimension Theorem T has a nonzero kernel: T is not invertible.
b) For $\lambda \in F$, the matrix $M - \lambda I_d$ represents the linear transformation $T - \lambda$ with respect to the basis $e_1, \ldots, e_d$. Thus λ is an eigenvalue iff $M - \lambda I_d$ is not invertible iff some diagonal entry $m_{ii} - \lambda = 0$ iff $\lambda = m_{ii}$ for at least one i. □

Theorem 2.9. *The following are equivalent:*
(i) The minimal polynomial P of T is split.
(ii) T is triangularizable.

Proof. (i) $\implies$ (ii): We go by induction on $\dim V$. The case of $\dim V = 1$ is clear, so suppose $\dim V = d$ and the result holds in dimension less than d. Since P is split, T has an eigenvalue α. Thus $\mathrm{Ker}(T - \alpha)$ is nontrivial, so $W = \mathrm{Image}(T - \alpha)$ is a proper invariant subspace, say of dimension $d' < d$. The minimal polynomial P_W of $T|_W$ is the lcm of the local minimal polynomials P_w for $w \in W$, so it divides P and is thus also split. By induction, $T|_W$ is triangularizable: let $e_1, \ldots, e_{d'}$ be a triangular basis. Extend this to a basis $e_1, \ldots, e_d$ of V in any way. We claim this basis is upper triangular, i.e., that each $V_i = \langle e_1, \ldots, e_i \rangle$ is an invariant subspace. We already know this if $i \leq d'$, so suppose $d' < i \leq d$ and $v \in V_i$. Then

$$Tv = (T - \alpha)v + \alpha v \in V_{d'} + V_i = V_i.$$

(ii) $\implies$ (i): Let $b_1, \ldots, b_d$ be a triangular basis for V; put $V_0 = \{0\}$, and for $1 \leq i \leq d$, put $V_i = \langle b_1, \ldots, b_i \rangle$. Then V_i/V_{i-1} is one-dimensional, so the minimal polynomial of T on it is $T - \lambda_i$ for some $\lambda_i \in F$. Thus for $1 \leq i \leq d$, $(T - \lambda_i)b_i \in V_{i-1}$. It follows that for all $1 \leq i \leq d$, $(T - \lambda_1) \cdots (T - \lambda_i)V_i = 0$. In particular, taking $i = d$ and putting $Q(t) = \prod_{i=1}^{d}(t - \lambda_i)$ we find that $Q(T)V = 0$. Thus the minimal polynomial P divides the split polynomial Q so P is itself split. □

Corollary 2.10. *If F is algebraically closed, then every linear transformation on a finite-dimensional F-vector space is triangularizable.*

3. Cyclic Spaces

The local minimal polynomial of the zero vector is 1, and the local minimal polynomial of an eigenvector is $t - \alpha$. It's less clear what to say about local minimal polynomials of larger degree, and in fact a natural question is which monic polynomials can arise as minimal polynomials of a linear transformation T on a finite-dimensional F-vector space. The answer may be surprising: all of them!

Example 3.1. *Let $p(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0 \in F[t]$. Let $V = F^n$ with standard basis $e_1, \ldots, e_n$, and consider the following endomorphism T_p of V:*

$$\forall 1 \leq i \leq n-1,\ T_p(e_i) = e_{i+1},$$
$$T_p(e_n) = -a_0 e_1 - \ldots - a_{n-1} e_n.$$

Then

$$p(T_p)e_1 = (T_p^n + a_{n-1}T_p^{n-1} + \ldots + a_1 T + a_0)e_1$$
$$= T_p e_n + a_{n-1}e_n + a_{n-2}e_{n-1} + \ldots + a_1 e_2 + a_0 e_1 = 0.$$

Also $p(T_p)e_2 = p(T_p)T_p e_1 = T_p p(T_p)e_1 = 0$, and similarly $p(T_p)e_i = 0$ for all $3 \leq i \leq n$. Thus the minimal polynomial P of T_p divides p. On the other hand we have $V = [e_1]$, so the degree of the local minimal polynomial P_{e_1} is $\dim V = n$. Since $P_{e_1} \mid P$, we conclude that $\deg P \geq n$ and thus $P_{e_1} = P = p$. The matrix of T_p with respect to the basis $e_1, \ldots, e_n$ is

$$\begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ 0 & 0 & 1 & \ldots & 0 & -a_3 \\ \vdots & & & & & \\ 0 & 0 & 0 & \ldots & 1 & -a_{n-1} \end{bmatrix}.$$

We call this the ***companion matrix*** *$C(p)$ of the monic polynomial p.*

We say that V is **cyclic** if $V = [v]$ for some $v \in V$. Consider the sequence of vectors

$$v, Tv, T^2 v, \ldots.$$

By definition of cyclic, they span the n-dimensional vector space V. Moreover, if $T^i v$ lies in the span of $v, Tv, \ldots, T^{i-1}v$, then $T^{i+1}v$ lies in the span of $Tv, \ldots, T^i v$, hence also in the span of $v, Tv, \ldots, T^{i-1}v$, and similarly for $T^j v$ for all $j > i$. It follows that $v, Tv, T^2 v, \ldots, T^{n-1}v$ is a basis for V, and we may write

$$T^n v = -a_0 v - a_1 T v - \ldots - a_{n-1}T^{n-1}v.$$

Then $p(t) = t^n + a_{n-1}t^{n-1} + \ldots + a_1 t + a_0$ is the minimal polynomial for T and the matrix for T with respect to our basis is the companion matrix $C(p)$. Thus the above "example" is precisely the general case of a cyclic linear transformation.

We now quote from Paul Halmos's *Automathography*:

"Given a linear transformation, you just choose a vector to start a basis with, and keep working on it with the transformation. When you stop getting something new, you stop, and, lo and behold, you have a typical companion matrix. I didn't know what he [David Netzorg] meant then, but I do now, and it's an insight, it's worth knowing."

Theorem 3.2 (Cyclicity Theorem).
Suppose $V = [v]$ is cyclic, with minimal polynomial $P = p_1^{a_1} \cdots p_r^{a_r}$.
a) For each monic polynomial divisor f of P, we have

$$V_f = [(P/f)(T)v]$$

and

$$P_{V_f} = f.$$

b) Every invariant subspace of V is of the form V_f for some monic polynomial $f \mid P$. In particular:
c) There are precisely $\prod_{i=1}^r (a_i + 1)$ invariant subspaces of V.
d) Every invariant subspace of V is cyclic.
e) Every quotient of V by an invariant subspace is cyclic.

Proof. As seen above, we have $P = P_v$.
a) Step 1: Write $fg = P$. Since $f(T)g(T)v = 0$, we have $[g(T)v] \subset V_f$. By Lemma 1.15,

$$\dim[g(T)v] = \deg P_{g(T)v} = \deg f$$

and similarly

$$\dim[f(T)V] = \deg g.$$

Thus

$$\dim V_f = \dim \operatorname{Ker} f(T) = \dim V - \dim f(T)V = \deg P - \dim[f(T)v]$$

$$= \deg P - \deg g = \deg f = \dim[g(T)v],$$

so $[g(T)v] = V_f$.
Step 2: By Step 1 and Lemma 1.15 we have

$$P_{V_f} = P_{(P/f)(T)v} = \frac{P_v}{\frac{P}{f}} = f.$$

b) Let W be an invariant subspace. By the Local Attainment Theorem there is $w \in W$ with $P_w = P_W$. Write $P = fP_W$. Using part a) we have

$$[w] \subset W \subset V_{P_W} V_{P_w} = [f(T)v].$$

By Lemma 1.15, $P_{f(T)v} = P_W$, so $\dim[f(T)v] = \deg P_{f(T)v} = \deg P_W = \deg P_w = [w]$. It follows that $W = [f(T)v]$ is cyclic.
Parts c) and d) follow immediately.
e) If W is an invariant subspace of $V = [v]$, let $\overline{v} = v + W$. Since every element of V is of the form $f(T)v$ for some f, every element of V/W is of the form $f(T)\overline{v}$ for some f: $V/W = [\overline{v}]$. □

Exercise 3.1. *Show that the following are equivalent:*
(i) V has infinitely many invariant subspaces.
(ii) F is infinite and V is not *cyclic.*

Theorem 3.3. *For a field F, the following are equivalent:*
(i) F is algebraically closed.
(ii) Every endomorphism of a nonzero finite-dimensional F-vector space is triangularizable.
(iii) Every endomorphism of a nonzero finite-dimensional F-vector space has an eigenvalue.

Proof. (i) $\implies$ (ii) is Corollary 2.10.
(ii) $\implies$ (iii): the first vector in a triangular basis is an eigenvector.
(iii) $\implies$ (i): We show the contrapositive: suppose F is not algebraically closed, so there is a polynomial $p(t) \in F[t]$ of positive degree with no root in F. Then $p(t)$ is the minimal polynomial P of the linear transformation T_p of Example 3.1, so by Proposition 2.2, T_p has no eigenvalue. □

Theorem 3.3 raises the prospect of proving the Fundamental Theorem of Algebra – that $\mathbb{C}$ is algebraically closed – by showing that every endomorphism of a nonzero finite-dimensional $\mathbb{C}$-vector space has an eigenvalue. This has indeed been done by H. Derksen [De03]; see also [Co] for a moderately simplified exposition. The argument is actually a bit more general: it shows that if R is a field in which every odd degree polynomial has a root, and for all $x \in R^\bullet$ exactly one of $\pm x$ is a square, then $R(\sqrt{-1})$ is algebraically closed.

4. Prime and Primary Vectors

Proposition 1.14 leads us to factor the minimal polynomial $P = p_1^{a_1} \cdots p_r^{a_r}$ and study the invariant subspaces V_{p_i} and $V_{p_i^{a_i}}$ more closely. We do so now.

A vector $v \in V$ is **prime** if its local minimal polynomial $P_v(t)$ is irreducible. A vector $v \in V$ is **primary** if $P_v(t)$ is a power of an irreducible polynomial. For a prime (resp. primary) vector v, the local minimal polynomial p (resp. p^a) is called the **prime value** (resp. **primary value**) of v.

Exercise 4.1. *Show that the zero vector is* not *primary.*

Lemma 4.1. *Let $p(t)$ a prime polynomial. If $v \in V_p^\bullet$, then $P_v = p$: v is a prime vector with prime value p.*

Exercise 4.2. *Prove Lemma 4.1.*

As the following exercises explore, prime vectors is that they are a suitable analogue of eigenvectors when the ground field is not algebraically closed.

Exercise 4.3. *Show that an eigenvector is a prime vector.*

Exercise 4.4. *Show that for a field F, the following are equivalent:*
(i) For every linear transformation on a finite-dimensional F-vector space, every prime vector is an eigenvector.
(ii) F is algebraically closed.

Proposition 4.2.
a) Let $p(t)$be a prime factor of the minimal polynomial T. Then $V_p \neq \{0\}$.
b) In particular every linear transformation admits a prime vector.

Proof. This is immediate from the Local Attainment Theorem (Theorem 1.16). □

Proposition 4.3. *a) It V is simple, then its minimal polynomial is prime.*
b) If $v \in V$ is a prime vector, then the invariant subspace $[v]$ is simple.

Proof. a) If $P(T)$ were not prime, then it would have monic a monic divisor $f \notin \{1, P\}$. By the Local Attainment Theorem (Theorem 1.16) there is $w \in V$ such that $P_w = f$ and thus $\dim[w] = \deg f \notin \{0, \deg P\}$, so $[w]$ is a proper, nonzero invariant subspace.
b) This is immediate from the Cyclicity Theorem (Theorem 3.2). □

For a prime polynomial p, we call V_p the **p-isotypic subspace of V**.

Proposition 4.4. *For every prime p, the p-isotypic subspace V_p is semisimple.*

Proof. Every nonzero vector $v \in V_p$ lies in the simple invariant subspace $[v]$, so V_p is spanned by its simple invariant subspaces. Apply Theorem 1.2. □

We define the **socle** $\mathfrak{s}(V)$ to be the subspace spanned by the prime vectors of V.

Theorem 4.5 (Prime Decomposition Theorem)**.**
a) We have $\mathfrak{s}(V) = \bigoplus V_p$, the sum extending over all prime factors of P.
b) The socle $\mathfrak{s}(V)$ is the largest semisimple invariant subspace of V.
c) In particular, the following are equivalent:
(i) The space V is semisimple.
(ii) The space V is its own socle: $V = \mathfrak{s}(V)$.
(iii) The space V admits a basis of prime vectors. (iv) The minimal polynomial P is squarefree.

Proof. a) The independence of the spaces V_p follows from Proposition 1.14. Moreover we know that $V_p = 0$ unless $p \mid P$. The result follows.
b) Since each V_p is semisimple, so is $\bigoplus V_p = \mathfrak{s}(V)$. Suppose W is a semisimple invariant subspace properly containing $\mathfrak{s}(V)$. Since W is spanned by its simple invariant subspaces and strictly contains $\mathfrak{s}(V)$, there must be a simple invariant subspace S of W and not contained in $\mathfrak{s}(V)$. Since S is simple, this gives $S \cap \mathfrak{s}(V) = 0$. But by the Local Attainment Theorem (Theorem 1.16), S admits a prime vector: contradiction.
c) This follows immediately. □

Proposition 4.6.
a) The socle $\mathfrak{s}(V)$ consists of all vectors with squarefree local minimal polynomial.
b) The space V is semisimple iff its minimal polynomial is squarefree.

Proof. a) Let $p_1, \ldots, p_r$ be the distinct prime divisors of P, so $\mathfrak{s}(V) = \bigoplus_{i=1}^r V_{p_i} = V_{p_1 \cdots p_r}$ by Proposition 1.14. Since a divisor f of P is squarefree iff it divides $p_1 \cdots p_r$, this proves part a).
b) This is immediate from part a) and Theorem 4.5c). □

We can now give a description of all semisimple invariant subspaces.

Proposition 4.7 (Classification of Semisimple Invariant Subspaces)**.** *Let $W \subset V$ be a semisimple invariant subspace. Then:*
a) For every prime divisor p of P, $W_p = V_p \cap W$.
b) We have $W = \bigoplus W_p$.
c) If V is split semisimple, then we get an invariant subspace selecting for each eigenvalue λ of T any subspace W_λ of the λ-eigenspace V_λ and putting $W = \bigoplus_\lambda W_\lambda$, and every invariant subspace arises in this way.

Exercise 4.5. *Prove Proposition 4.7.*

Now we go deeper by looking not just at prime vectors but primary vectors. Recall that we have factored our minimal polynomial as

$$P = p_1^{a_1} \cdots p_r^{a_r}.$$

The **p-primary subspace** of V is

$$V^p = V_{p_i^{a_i}}.$$

Theorem 4.8 (Primary Decomposition Theorem)**.** *We have $V = \bigoplus V^p$, the sum extending over the distinct prime divisors of P.*

Proof. Let the minimal polynomial be given by $P = p_1^{a_1} \cdots p_r^{a_r}$. Since $p_1^{a_1}, \ldots, p_r^{a_r}$ are pairwise coprime, by Proposition 1.14 we have

$$V_{p_1^{a_1}} + \ldots + V_{p_r^{a_r}} = \bigoplus_{i=1}^{n} V_{p_i^{a_i}} = V_{p_1^{a_1} \cdots p_r^{a_r}} = V_P = V. \qquad \square$$

Proposition 4.9. *Let $p(t)$ be a prime, and let $W \subset V$ be invariant. Then:*
a) $(V/W)^p = (V^p + W)/W$.
b) $\dim V^p = \dim W^p + \dim(V/W)^p$.

Proof. a)Let $\overline{v} \in V/W$, and let v be any lift of $\overline{v}$ to V. Then $\overline{v}$ is p-primary iff there is some $a \in \mathbb{Z}^+$ such that $p(T)^a v \in W$ iff $v \in V^p + W$.
b) We have

$$\dim W^p + \dim(V/W)^p = \dim W^p + \dim(V^p + W)/W$$

$$= \dim W^p + \dim(V^p + W) - \dim W = \dim W \cap V^p + \dim(V^p + W) - \dim W$$

$$\dim W \cap V^p + \dim V^p + \dim W - \dim W \cap V^p - \dim W = \dim V^p. \qquad \square$$

5. The Characteristic Polynomial

Theorem 5.1. *For each prime divisor p of P, we have $\deg(p) \mid \dim V^p$.*

Proof. We may assume $V = V^p$ and go by induction on $\dim V$. Since V is nonzero, it has a prime vector $[v]$ and thus $\dim V \geq \dim[v] = \deg(p)$. We have $\dim V = \dim[v] + \dim V/[v]$. Since V is p-primary, so is $V/[v]$, and since $\dim V/[v] < \dim V$, by induction we have $\dim V/[v] = k \deg(p)$ for some $k \in \mathbb{N}$, and thus

$$\dim V = \dim[v] + \dim V/[v] = \deg(p) + k \deg(p) = (k + 1) \deg(p). \qquad \square$$

Corollary 5.2. *If $F = \mathbb{R}$ and $\dim V$ is odd, then there is an eigenvalue.*

Exercise 5.1. *Prove Corollary 5.2.*

In light of Theorem 5.1, for any prime divisor p of P we may define

$$\chi_p(t) = p(t)^{\frac{\dim V^p}{\deg p}}$$

and the **characteristic polynomial**

$$\chi(t) = \prod_{p \mid P} \chi_p(t).$$

Proposition 5.3. *Let P and χ be the minimal and characteristic polynomials.*
a) The polynomials P and χ have the same prime divisors.
b) We have $\deg \chi = \dim V$.
c) We have $P \mid \chi$: equivalently, $\chi(T) = 0$.
d) We have $\deg P \leq \dim V$.

Proof. a) This is built into our definition of χ.
b) We have

$$\deg \chi = \sum_p \deg \chi_p = \sum_p \deg p^{\frac{\dim V_{p^a}}{\deg p}} = \sum_p \dim V_{p^a} = \dim \bigoplus V_{p^a} = \dim V.$$

c) Let $P = p_1^{a_1} \cdots p_r^{a_r}$. We must show that for all $1 \leq i \leq r$, $p_i^{a_i} \mid \chi_{p_i}$; equivalently, $\deg \chi_{p_i} \geq a_i$. For each $1 \leq i \leq r$, there is $v_i \in V$ with $P_{v_i} = p_i^{a_i}$. Since $V^{p_i} \supset [v_i]$, we have

$$\deg \chi_{p_i} = \dim V^{p_i} \geq \dim [v_i] = \deg p_i^{a_i}.$$

d) This is immediate from b) and c). □

Theorem 5.4. *Let $W \subset V$ be invariant. Let χ' be the characteristic polynomial of W and χ'' be the characteristic polynomial of V/W. Then*

$$\chi(t) = \chi'(t)\chi''(t).$$

Proof. Let $P = p_1^{a_1} \cdots p_r^{a_r}$ be the minimal polynomial of T on V. It is enough to show that $\chi_p(t) = \chi'_p(t)\chi''_p(t)$ for all p |P. By Proposition 4.9,

$$\dim V^p = \dim W^p + \dim (V/W)^p$$

and thus

$$\chi_p(t) = p(t)^{\frac{\dim V^p}{\deg p}} = p(t)^{\frac{\dim W^p + \dim (V/W)^p}{\deg p}} = p(t)^{\frac{\dim W^p}{\deg p}} p(t)^{\frac{\dim (V/W)^p}{\deg p}} = \chi'_p(t)\chi''_p(t).$$

□

We can now give an important interpretation of the characteristic polynomial. A **composition series** in V is a maximal chain of invariant subspaces:

$$0 \subset V_0 \subsetneq \ldots \subsetneq V_n \subset V;$$

that is, each V_i is an invariant subspace, and for all $0 \leq i \leq n-1$, $V_i \subset V_{i+1}$ and there is no invariant subspace properly in between them. We say that the composition series has **length** n.

Example 5.5. *If $b_1, \ldots, b_d$ is a triangular basis, then $V_0 = \{0\}$, $V_i = \langle b_1, \ldots, b_i \rangle$ is a composition series. Conversely, given a composition series with $\dim V_i = i$ for all i, then taking $b_i \in V_i \setminus V_{i-1}$ gives a triangular basis.*

However, triangular bases exist only in the split case. A composition series is a suitable analogue in the general case.

Observe that the statement that there is no invariant subspace properly in between V_i and V_{i+1} is equivalent to the quotient V_{i+1}/V_i being simple. Thus V_{i+1}/V_i is cyclic and prime and has minimal polynomial equal to its characteristic polynomial equal to a prime polynomial p_i: we call p_i's the **composition factors** of the composition series. By induction on Theorem 5.4 we find that

$$\chi(t) = p_1(t) \cdots p_n(t).$$

We have proved an important result.

Theorem 5.6 (Jordan-Hölder)**.**
a) Any two composition series have the same composition factors up to order.
b) In particular any two composition series have the same length.
c) The product of the composition factors is equal to the characteristic polynomial.

A basis $b_1, \ldots, b_d$ is **adapted** to a composition series $\{V_i\}_{i=0}^n$ if for all i there are $\dim V_i - \dim V_{i-1}$ basis vectors lying in $V_i \setminus V_{i-1}$.

Exercise 5.2. *a) Show that any composition series admits a basis adapted to it.*
*b) Show that the matrix of T with respect to a basis adapted to a composition series is in **block upper triangular form**: if $\chi(t) = p_1 \cdots p_n$, then such a matrix consists of $\deg p_i \times \deg p_i$ square matrices along the diagonal and all zeros below and to the left of these blocks.*

Theorem 5.7. *Let T be an endomorphism of a nonzero finite dimensional vector space V, with minimal polynomial $P = \prod_{i=1}^r p_i^{a_i}$.*
a) (Semisimplicity Theorem) V is semisimple iff P is squarefree.
b) (Diagonalizability Theorem) V is diagonalizable iff P is squarefree and split.
c) (Simplicity Theorem) V simple iff χ is prime.

Proof. a) We have proved this already; we just repeat it for comparison.
b) By Proposition 2.7, if V is diagonalizable then P is squarefree and split. Conversely, suppose P is squarefree and split. By part a), $V = \bigoplus V_p$ is the direct sum of its p-isotypic subspaces, and since P is split, each p has degree one and thus $V_p = V_{t-\lambda} = V_\lambda$ is an eigenspace. So V has a basis of eigenvectors and is diagonalizable.
c) If V is simple then it is semisimple, so $P = p_1 \cdots p_r$ is squarefree and $V = \bigoplus V_{p_i}$. Since V is simple $r = 1$, so $P = p_1$ and thus $\chi = p_1^a$. But $\dim V = \deg p_1 = \deg \chi$ so $a = 1$ and $\chi = p_1$. Convesely, if $\chi = p$ is prime, this forces $P = p$ to be squarefree and thus V to be semisimple of dimension $\deg p$, hence simple. □

Exercise 5.3. *a) Let $F = \mathbb{C}$, and suppose that the matrix M of T with respect to some basis of V is "a root of unity": $M^n = I$ for some $n \in \mathbb{Z}^+$. Show that T is diagonalizable.*
b) Show that if instead $F = \mathbb{R}$ the result no longer holds.
c) Show that if instead F has positive characteristic the result no longer holds.

The following result gives a generalization.

Theorem 5.8. *Let T be an endomorphism of a nonzero finite dimensional vector space V, and let $N \in \mathbb{Z}^+$.*
- a) *If T is diagonalizable, then so is T^N.*
- b) *Suppose moreover that F is algebraically closed of characteristic not dividing N (e.g. $F = \mathbb{C}$) and that T is invertible. If T^N is diagonalizable, so is T.*

Proof. a) If in a basis $e_1, \ldots, e_n$ the transformation T is diagonal with diagonal entries $\lambda_1, \ldots, \lambda_n$, then in the same basis the transformation T^N is diagonal with diagonal entries $\lambda_1^N, \ldots, \lambda_n^N$.
b) If T^N is diagonalizable, then by the Diagonalizability Theorem (Theorem 5.7b)) there are distinct $\lambda_1, \ldots, \lambda_k \in F$ such that

$$\prod_{i=1}^k (T^N - \lambda_i) = 0.$$

For each such i, let $\lambda_i^{1/N}$ be any element of F such that $(\lambda_i^{1/N})^N = \lambda_i$. (Since F is algebraically closed, such an element exists.) Moreover, since the characteristic of F does not divide N, there is a primitive Nth root of unity ζ_N in F and thus the roots of $t^n - \lambda_i$ in F are precisely $\zeta_N^j \lambda_i^{1/N}$ for $0 \leq j \leq N-1$. This means that the polynomial $P(t) \coloneqq \prod_{i=1}^K (t^n - \lambda_i) \in F[t]$ is squarefree and split, and since $P(T) = 0$, the minimal polynomial of T divides $P(t)$ and thus is itself squarefree and split. Applying the Diagonalizability Theorem again shows that T is diagonalizable. $\square$

Exercise 5.4. *a) Show that for all $n, N \geq 2$ there is $M \in M_{n,n}(\mathbb{C})$ such that M^N is diagonalizable but M is not.*
b) Suppose F is algebraically closed of characteristic $p > 0$. Show that for all positive integers N that are divisible by p, there is a finite-dimensional F-vector space V and an invertible endomorphism T of V such that T^N is diagonalizable but T is not.

Exercise 5.5. *Let p be a prime divisor of the minimal polynomial P: suppose $P = p(t)^a Q(t)$ with $p \nmid Q$. We define the* ***algebraic multiplicity*** *of p to be the exponent of p in χ, i.e., the largest e such that $p^e \mid \chi$. We define the* ***geometric multiplicity*** *of p to be $\frac{\dim V_p}{\deg p}$, i.e., the number of linearly independent p-simple invariant subspaces.*
a) Show that the algebraic multiplicity of P is equal to $\frac{\dim V_{p^a}}{\deg p}$.
b) Show that the algebraic multiplicity of p is less than or equal to the geometric multiplicity of p.
c) Show that V is semisimple iff the algebraic multiplicity of p is equal to the geometric multiplicity of p for all primes $p \mid P$.

6. The Cyclic Decomposition Theorem

Lemma 6.1. *Let $p \in F[t]$ be a prime polynomial. Suppose that V is p-primary and V_p is cyclic. Then V is cyclic.*

Proof. By induction on $\dim V$, the case of $\dim V = 1$ being immediate. Let $n = \dim V$ and suppose that the result holds for spaces of dimension less than n. The result is clear if $V = V_p$ so we may suppose $V_p \subsetneq V$. The quotient space V/V_p is p-primary of dimension smaller than n. Further, $V/V_p \cong p(T)V$, so

$$(V/V_p)_p \cong (p(T)V)_p = p(T)V \cap V_p$$

is a submodule of a cyclic module, hence cylic. By induction there is $\overline{v} \in V/V_p$ with $V/V_p = [\overline{v}]$. Lift $\overline{v}$ to $v \in V$; then $V = \langle V_p, [v] \rangle$. Finally, since the p-isotypic space V_p is cyclic, it is simple. Since $[v] \neq 0$ is p-primary, $V_p \cap [v] = [v]_p \neq 0$, and thus $V_p \subset [v]$. We conclude $V = \langle V_p, [v] \rangle = [v]$. $\square$

An invariant subspace $W \subset V$ is **replete** if the minimal polynomial P_W is the minimal polynomial of T on V.

Theorem 6.2 (Cyclic Decomposition Theorem)**.**
a) Suppose V is primary, and let $W \subset V$ be a replete, cyclic invariant subspace. Then there is an invariant subspace $W' \subset V$ such that $V = W \oplus W'$.
b) Any V can be written as a direct sum of cyclic subspaces W_i in which each P_{W_i} is a prime power.
c) Suppose $V = \bigoplus_{i=1}^m W_i = \bigoplus_{j=1}^n W_j'$ are two decompositions into direct sums of

cyclic primary subspaces. Then $m = n$, and there is a permutation σ of $\{1, \ldots, n\}$ such that for all i, $P_{W_i} = P_{W'_{\sigma(i)}}$.

Proof. a) Let $P = P_V = p^a$. We go by induction on $\dim V$, the case $\dim V = 1$ being clear. If V is cyclic, then by the Cyclicity Theorem its only replete invariant subspace is V itself, a trivial case. Thus we may assume V is not cyclic, hence by Lemma 6.1 that V_p is not cyclic: $\dim V_p > \deg p$. Since $\dim W_p = p$, there is a prime invariant subspace K of dimension $\dim p$ such that $W \cap K = 0$.
We CLAIM that $\overline{W} = (W + K)/K$ is cyclic and primary in the quotient space V/K.
PROOF OF CLAIM Let $W = [w]$ and $\overline{w} = w + K$. Then $\overline{W} = [\overline{w}]$ is cyclic. Further, *a priori* the minimal polynomial $\overline{P} = P_{V/K}$ divides $P = p^a$. We will show $P_{\overline{W}} = p^a$, which suffices. Indeed, if $W = [w]$, then $p^{a-1}(T)w$ is a nonzero element of W and hence does not lie in K since $K \cap W = \{0\}$, so $p^{a-1}(T)\overline{w} \neq 0$ and thus $P_{\overline{w}} = p^a$.
By induction, there is an invariant subspace $U \subset V/K$ such that

$$V/K = (W + K)/K \oplus U. \tag{1}$$

Let $W' = \{v \in V \mid v + K \subset U\}$ (in other words, W' is the subspace of V containing K which *corresponds* to U). Then $W' \supset K$ and $W'/K = U$. From (1) we get

$$V = W + K + W' = W + W'.$$

If $v \in W \cap W'$, then $w + K \in (W + K)/K \cap W'/K = 0$, so $w \in K$ and thus $w \in W \cap K = 0$. It follows that $V = W \oplus W'$.
b) Let $V = \bigoplus V^p$ be its primary decomposition. Since a direct sum of a direct sum of cyclic primary invariant subspaces is a direct sum of cyclic primary invariant subspaces, it is enough to treat the case $V = V^p$. This follows from part a).
c) As usual it suffices to consider the case that $V = V^p$ is primary. We go by induction on $\dim V$. We may suppose that

$$V = \bigoplus_{i=1}^{r} C(p^{a_i}) = \bigoplus_{j=1}^{s} C(p^{b_i})$$

with $a_1 \geq \ldots \geq a_r$ and $b_1 \geq \ldots \geq b_s$, and our task is to show that $r = s$ and $a_i = b_i$ for all i. We have

$$p(T)V = \bigoplus_{i=1}^{r} C(p^{a_i - 1}) = \bigoplus_{j=1}^{s} C(p^{b_{i-1}}). \tag{2}$$

Since $V_p \neq 0$, $\dim p(T)V < \dim V$ and thus the cyclic decomposition of $p(T)V$ is unique. We do need to be careful about one point: if $a_i = 1$, then $C(p^{a_i - 1})$ is the zero vector space so needs to be removed from the direct sum decomposition. To take care of this, let I be such that $a_I > 1$ but $a_i = 1$ for $i > I$; and similarly let J be such that $a_J > 1$ but $a_j = 1$ for $j > J$. Then induction gives $I = J$ and $a_i - 1 = b_i - 1$ for all $1 \leq i \leq I$, hence of course that $a_i = b_i$ for all $1 \leq i \leq I$. Finally we must show that $r - I = s - J$, but this follows by comparing dimensions:

$$r - I = \dim V - \sum_{i=1}^{I} \deg p_i^{a_i} = \dim V - \sum_{j=1}^{J} \deg p_i^{b_i} = s - J.$$

$\square$

Write $V = \bigoplus_{i=1}^{n} W_i$ with each W_i a cyclic invariant subspace with prime power minimal polynomial $f_i = P_{W_i}$. By Theorem 6.2c) the multiset of these polynomials

– i.e., like a set but each element carries with it a certain positive integer, the **multiplicity** – is invariant of the chosen decomposition. These polynomials are called **elementary divisors**.

Proposition 6.3. *a) The lcm of the elementary divisors is the minimal polynomial.* *b) The product of the elementary divisors is the characteristic polynomial.*

Exercise 5.1: Prove Proposition 6.3.

Exercise 5.2: a) Show that we may write $V = \bigoplus_{i=1}^n W_i$ such that: each W_i is cyclic and for all $1 \leq i \leq n-1$, the minimal polynomial f_{i+1} of W_{i+1} divides the minimal polynomial f_i of W_i.
b) Show that the sequence of monic polynomials $f_n \mid f_{n-1} \mid \ldots \mid f_1$ of part a) is unique. The polynomials in this sequence are called **invariant factors**.

7. Rational and Jordan Canonical Forms

Let T be a linear endomorphism of a finite-dimensional F-vector space V. By Cyclic Decomposition we may write $V = \bigoplus_{i=1}^n W_i$ with each W_i a primary cyclic invariant subspace. For each i, choose a vector $w_i \in W_i$ with $W_i = [w_i]$, let $p_i = P_{w_i} = P_{W_i}$, and let $b_{i1} = w_i, b_{i2} = Tw_i, \ldots, b_{i \deg p_i} = T^{\deg p_i - 1} w_i$. Then $\mathcal{B}_R = b_{11}, \ldots, b_{1 \deg p_1}, \ldots, b_{n \deg p_n}$ is an especially pleasant basis for V; the corresponding matrix for $\mathcal{B}_R$ is

$$M = \bigoplus_{i=1}^{n} C(p_i).$$

The matrix M is called the **Rational Canonical Form**, and it is uniquely associated to T up to a permutation of the diagonal blocks comprising the companion matrices.

An endomorphism T is **nilpotent** if there is some positive integer N such that $T^N = 0$. It follows that the minimal polynomial is t^a for some $a \leq d = \dim V$ – thus $T^d = 0$ – and T is primary with characteristic polynomial t^d. Further each elementary divisor is of the form t^b for some $1 \leq b \leq a$. Notice that the companion matrices $C(t^b)$ take an especially simple form: in particular they are all **strictly lower triangular** (and conversely are the only strictly lower triangular companion matrices) and indeed is identically zero except for having ones along the **subdiagonal**: the diagonal immediately below the main one.

Moreover, any *split* endomorphism has a canonical form which is almost as simple. The Cyclic Decomposition Theorem reduces us to the cyclic primary case, in which the minimal and characteristic polynomials are both of the form $(t - \lambda)^a$ for some eigenvalue $\lambda \in F$. This means precisely that $T - \lambda$ is nilpotent, so has a basis with respect to which its matrix is also zero except having 1's along the subdiagonal. Adding back the scalar matrix λI_n, we find that in this basis the matrix of T has λ's along the main diagonal, 1's along the subdiagonal, and is otherwise 0. Such a matrix is called a **Jordan block** $J(n, \lambda)$. A matrix which is a direct sum of Jordan blocks is said to be in **Jordan canonical form**.

Conversely, suppose that the matrix of T with respect to some basis $b_1, \ldots, b_n$ is the Jordan block $J(n, \lambda)$. Then $T - \lambda$ is a cyclic endomorphism, so every $v \in V$ is of the form $f(T - \lambda)v$ for some $f \in F[t]$. But $f(T - \lambda) = g(T)$ is a polynomial in

T, so V is also cyclic for the endomorphism T. Further, $(T-\lambda)^n$ kills each b_i hence is zero, so the minimal polynomial of T divides $(t-\lambda)^n$ and thus $\chi_T(t) = (t-\lambda)^n$. We conclude that T is cyclic and λ-primary.

Theorem 7.1. *For a linear endomorphism T, the following are equivalent:*
(i) T is split.
(ii) There is a basis of V such that the matrix of T is in Jordan canonical form.

Proof. (i) $\implies$ (ii): We argued for this just above.
(ii) $\implies$ (i): By the above, if $M = \bigoplus_{i=1}^n J(n_i, \lambda_i)$, then $\chi(t) = \prod_{i=1}^n (t-\lambda)^{n_i}$. $\square$

8. Similarity

After having done so much work with a single linear endomorphism T of a finite-dimensional F-vector space V, suppose now that we have two endomorphisms T_1 and T_2. We can ask when T_1 and T_2 are "essentially the same": roughly, they are the same transformation written in different linear coordinate systems. What does this mean precisely? Here are two ways to construe it:

• There is an invertible linear endomorphism $A \in \operatorname{GL} V$ which carries the action of T_1 into the action of T_2: $AT_1 = T_2 A$.

• There are two bases $\mathcal{B} = (b_1, \ldots, b_n)$ and $\mathcal{B}' = (b_1', \ldots, b_n')$ such that the matrix of T_1 with respect to $\mathcal{B}$ is equal to the matrix of T_2 with respect to $\mathcal{B}'$.

These conditions are in fact equivalent: to each other, and also to the condition that there is $A \in \operatorname{GL} V$ such that $AT_1A^{-1} = T_2$. That the first condition is equivalent to this is clear: just compose $AT_1 = T_2A$ on the right with A^{-1}. As for the second: we may view giving a basis $\mathcal{B}$ of V as giving a linear isomorphism $B : V \to F^n$, uniquely determined by sending $b_i \mapsto e_i$, the ith standard basis vector. Then to say that the matrix of T_1 with respect to $\mathcal{B}$ is M is to say that $BT_1B^{-1} = M$. Similarly, if $B' : V \to \mathbb{F}^n$ be the linear isomorphism corresponding to $\mathcal{B}'$, then we get also $B'T_2(B')^{-1} = M$, so

$$BT_1B^{-1} = B'T_2(B')^{-1}$$

and thus

$$T_2 = (B')^{-1}BT_1B^{-1}B' = \left((B')^{-1}B\right) T_1 \left((B')^{-1}B\right)^{-1}.$$

Lemma 8.1. *Suppose $AT_1A^{-1} = T_2$. Let $W \subset V$ be a T_1-invariant subspace. Then:*
a) AW is a T_2-invariant subspace.
b) If $W = [v]$, then $AW = [Av]$.
c) For any $f \in F[t]$, $Af(T_1)A^{-1} = f(T_2)$.
d) The minimal polynomial for T_1 acting on W is equal to the minimal polynomial or T_2 acting on AW.

Exercise 7.1: Prove Lemma 8.1.

Theorem 8.2. *For linear endomorphims T_1, T_2 on V, the following are equivalent:*
(i) T_1 and T_2 are similar.
(ii) T_1 and T_2 have the same elementary divisors.
(iii) T_1 and T_2 have the same invariant factors.

Proof. (i) $\implies$ (ii): Suppose $AT_1A^{-1} = T_2$, and let $V = \bigoplus W_i$ be a decomposition into primary cyclic invariant subspaces for T_1. By Lemma 8.1, $V = \bigoplus AW_i$ is a decomposition into primary cyclic invariant subspaces for T_2, and the elementary divisors are the same.
(ii) $\implies$ (i): If the elementary divisors are $\{f_i\}$, each of T_1 and T_2 have a basis with respect to which the matrix is $\bigoplus_i C(f_i)$.
(ii) $\iff$ (iii): The list of elementary divisors determines the list of invariant factors, and conversely, in a straightforward way. We leave the details to the reader. □

Exercise 7.2: Write out a detailed proof of (ii) $\iff$ (iii) in Theorem 8.2.

We can take things a step further: it is not necessary for T_1 and T_2 to be endomorphisms of the same vector space V. Let V_1 and V_2 be two finite-dimensional vector spaces, and let $T_1 \in \operatorname{End} V_1$, $T_2 \in \operatorname{End} V_2$. We say that T_1 and T_2 are **similar** if there is an isomorphism $A : V_1 \to V_2$ such that $AT_1A^{-1} = T_2$.

Exercise 7.3: Show that $T_1 \in \operatorname{End} V_1$ and $T_2 \in \operatorname{End} V_2$ are similar iff there is a basis B_1 for V_1 and V_2 for B_2 such that the matrix of T_1 with respect to B_1 is equal to the matrix of T_2 with respect to B_2.

Theorem 8.3. *Let V_1 and V_2 be finite-dimensional vector spaces, and let $T_1 \in \operatorname{End} V_1$, $T_2 \in \operatorname{End} V_2$. The following are equivalent:*
(i) T_1 and T_2 are similar.
(ii) T_1 and T_2 have the same elementary divisors.
(iii) T_1 and T_2 have the same invariant factors.

Exercise 7.4: Prove Theorem 8.3.

9. The Cayley-Hamilton Polynomial (Or: Up With Determinants?)

Given a linear transformation T, how does one actually compute the eigenvalues? We choose a basis $e_1, \ldots, e_n$ with corresponding matrix M. Then λ is an eigenvalue for T iff $\lambda I_n - M$ is not invertible iff $\det(\lambda I_n - M) = 0$. We don't just have to randomly check one λ after another: if t is an indeterminate then $\det(tI_n - M)$ is a polynomial in t and its roots in F are precisely the eigenvalues. This motivates the following definition.

Let $e_1, \ldots, e_n$ be a basis for V and let M be the associated matrix of T. The **Cayley-Hamilton polynomial of T** is

$$X(t) = \det(tI_n - M) \in F[t].$$

Lemma 9.1. *The Cayley-Hamilton polynomial is independent of the choice of basis.*

Proof. For a different basis $e_1', \ldots, e_n'$ of V, the associated matrix M' is of the form PMP^{-1} for some invertible matrix M. Thus

$$\det(tI_n - M') = \det(tI_n - PMP^{-1}) = \det(P(tI_n - M)P^{-1})$$

$$= (\det P)\det(tI_n - M)(\det P^{-1}) = (\det P)\det(tI_n - M)(\det P)^{-1} = \det(tI_n - M).$$

□

And now a miracle occurs!

Theorem 9.2. *For any linear endomorphism T we have $\chi(t) = X(t)$.*

Proof. Step 1: Suppose V is cyclic. Then V admits a basis with respect to which the matrix M is a companion matrix $C(p)$. The Cayley-Hamilton polynomial of $C(p)$ is the determinant of the matrix XXX. This matrix is not quite upper triangular, but it is very close: to bring it to upper triangular form we multiply the first row by $\frac{1}{t}$ and add it to the second row, then we mutiply the second row by $\frac{1}{t}$ and add it to the third row, and so forth. We get a diagonal matrix the first $n-1$ diagonal entries of which are each equal to t and for which the last diagonal entry is $\frac{a_0}{t^{n-1}} + \frac{a_1}{t^{n-2}} + \ldots + \lambda - a_{n-1}$. Thus

$$X(C(p)) = \det(t - C(p)) = t^{n-1}\left(\frac{a_0}{t^{n-1}} + \frac{a_1}{t^{n-2}} + \ldots + t + a_{n-1}\right) = p(t).$$

Step 2: By the Cyclic Decomposition Theorem, there is a basis with respect to which M is a direct sum of companion matrices $C(p_i)$. By Step 1 we have

$$X(t) = \det(tI - M) = \prod_{i=1}^{n}(tI - C(p_i)) = \prod_{i=1}^{n} p_i = \chi(t). \qquad \square$$

Theorem 9.3 (Cayley-Hamilton)**.** *We have $X(T) = 0$.*

Exercise 8.1: Prove Theorem 9.3.

10. Extending The Ground Field

In truth it is not so useful to maintain that when the ground field F is not algebraically closed, there are linear endomorphisms without eigenvalues. A better perspective is to define the eigenvalues by passage to the algebraic closure. In the standard approach this just means taking the roots of the characteristic polynomial χ over $\overline{F}$, but our setup is a bit richer. Namely, χ has an intrinsic meaning over F – it measures the dimensions of the p-primary subspaces – and when we pass from F to $\overline{F}$ then any prime polynomial p of degree greater than 1 will no longer be prime: it will split into linear factors. Thus *a priori* the characteristic polynomial $\chi_{\overline{F}}$ of the extension of T to $V_{/\overline{F}}$ has a totally different definition – it is built out of different local building blocks – and it is not obvious that $\chi_{\overline{F}} = \chi$. Fortunately it is true, and even more: χ is invariant upon any field extension. We will establish that in this section and then use this invariance to show that our characteristic polynomial agrees with the usual one defined via a determinant.

10.1. **Some Invariances Under Base Extension.**

Proposition 10.1. *Let K/F be a field extension, and let $T_{/K}$ be the extended linear endomorphism of $V_{/K}$. Then the minimal polynomial of $T_{/K}$ is equal to the minimal polynomial of T.*

Proof. Step 0: An F-linearly independent subset $S \subset V$ cannot become K-linearly dependent in $V_{/K}$: for instance, we can reduce to the case $V = F^d$; linear dependence can then be checked by placing the vectors as rows of a matrix and putting the matrix in reduced row echelon form (rref); since rref is *unique*, it does not change upon base extension.
Step 1: Let P be the minimal polynomial of T and let Q be the minimal polynomial of $T_{/K}$. Since $P(T)$ kills every basis element of $V_{/K}$ it kills $V_{/K}$: thus $Q \mid P$.
Step 2: By the Local Attainment Theorem we have $P = P_v$ for some $v \in V$. The local minimal polynomial does not change under base extension: arguing as in Step 1, the only way it could change would be to become a proper divisor; on the other

hand, by Step 0 the vectors $v, Tv, \ldots, T^{\deg P_v - 1}v$ remain linearly independent in K, so the degree of the local minimal polynomial for v viewed as an element of $V_{/K}$ must be $\deg P_v$. $\square$

Theorem 10.2. *Let K/F be a field extension, and let $T_{/K}$ be the extended linear endomorphism of $V_{/K}$. Then the characteristic polynomial of T, viewed as an element of $K[t]$, is equal to the characteristic polynomial of $T_{/K}$.*

Proof. First proof: This is clear for the Cayley-Hamilton polynomial $X(t)$, and by Theorem 9.2, $\chi(t) = X(t)$.
Second proof: We may reduce to the case that T is cyclic and primary, with characteristic polynomial $p(t)^a$ for p irreducible in $F[t]$. Let C be an algebraically closed field containing K, and let χ_C be the characteristic polynomial of $T_{/C}$. It is enough to show that $\chi = \chi_C$. Let $\lambda_1, \ldots, \lambda_r$ be the distinct roots of $p(t)$ in C, and let r and s be the separable and inseparable degrees of the field extension $F[t]/(p(t))/F$. Then $\chi(t)$ factors in $C[t]$ as $\prod_{i=1}^r (t - \lambda_i)^{as}$. The minimal polynomial does not change upon field extension (explain), so the characteristic polynomial $\chi_C(t) = \prod_{i=1}^r (t - \lambda_i)^{a_i}$. We want to show that $a_i = as$ for all i; since $\deg \chi_C = \dim V_{/C} = \dim V$, it is enough to show that all of the a_i's are equal. For this we use the fact that for all $1 \leq i \leq r$ there is $\sigma_i \in \mathrm{Aut}(C/F)$ with $\sigma_i(\lambda_1) = \lambda_i$: first, we have such an automorphism of the normal closure of $F[t]/(p(t))$, and second we may extend it to an automorphism of C using [FT, § 12.2]. Then $\sigma_i(V^{\lambda_1}) = V^{\lambda_i}$, and this gives the equality of the a_i's. $\square$

Theorem 10.3. *Suppose $P(t) = \prod_{i=1}^r (t - \lambda_i)^{a_i}$ is split, let $\mathcal{B} = b_1, \ldots, b_d$ be a triangular basis for T, and let M be the corresponding matrix. Then the diagonal entries of $M_{\mathcal{B}}$ are precisely the eigenvalues, and each eigenvalue λ_i appears precisely $\dim V_{(t-\lambda_i)^{a_i}}$ times.*

Proof. Since M is upper triangular, so is $tI - M$, and thus

$$\chi(t) = X(t) = \det(tI - M) = \prod_{i=1}^n (t - m_{ii}).$$

The number of times a given eigenvalue λ appears on the diagonal is thus equal to the largest integer a such that $(t-\lambda)^a$ divides $\chi(t)$, which in turn is equal to to $\deg \chi_{t-\lambda}(t) = \dim V_{(t-\lambda_i)^{a_i}}$. $\square$

Exercise 9.1: Let M be a block diagonal matrix with blocks $A_1, \ldots, A_n$, and such that each $d_i \times d_i$ block A_i has prime minimal polynomial p_i of degree d_i. Show that the characteristic polynomial is $\prod_{i=1}^n p_i$.

Remark: Axler gives in [A] a determinant-free proof of Theorem 10.3: it takes a little over two pages. I believe this is the second longest proof in [A]; the longest is the proof of Exercise 8.2 in the case $F = \mathbb{R}$.

Let $T_1, T_2 \in \mathrm{End}\, V$. We say T_1 and T_2 are **potentially similar** if there is some field extension K/F such that $(T_1)_{/K}$ and $(T_2)_{/K}$ are similar.

Similar implies potentially similar: we may take $K = F$. But more is true.

Exercise 9.2: Let $T_1, T_2 \in \mathrm{End}\, V$ be similar. Show that for any K/F, $(T_1)_{/K}$

is similar to $(T_2)_{/K}$. (Suggestion: consider associated matrices.)

And in fact much more is true: we have the following extremely useful result.

Theorem 10.4 (Potential Similarity Theorem)**.**
a) Let $T \in \operatorname{End} V$ have invariant factors $f_r \mid \ldots \mid f_1 = P$. Let K/F be any field extension. Then the invariant factors of $T_{/K}$ are still $f_r \mid \ldots \mid f_1 = P$.
b) If T_1 and T_2 are potentially similar, then they are similar.

Proof. a) By definition of the invariant factors, there is a basis $\{b_1, \ldots, b_n\}$ for V such that the associated matrix for T is $C(f_1) \oplus \ldots C(f_r)$. Of course this is still the matrix for $T_{/K}$ with respect to the same basis. It follows (using, of course, the uniqueness of invariant factors: see Exercise 5.2) that the invariant factors for $T_{/K}$ are still $f_r \mid \ldots \mid f_1 = P$.
b) This is immediate from part a) and Theorem 8.3. $\square$

Remark: Our presentation has emphasized primary vectors and primary decomposition, and thus until now we have preferred to work with elementary divisors rather than invariant factors. But unlike the invariant factors, the elementary divisors *can change* upon base extension, because they depend on the prime divisors of the minimal polynomial P, and while the minimal polynomial does not change after base extension, if it has a prime divisor p of degree greater than one, then in some extension (e.g. any splitting field) p will factor into polynomials of smaller degree. Analyzing how a prime polynomial factors in an arbitrary field extension requires some nontrivial field theory, especially when the ground field is not perfect. We invite the ambitious reader to try it out in the following exercise.

Exercise 9.3: Let $T \in \operatorname{End} V$. Let K/F be a field extension, with $K = \overline{K}$.
a) Give an explicit description of the invariant factors for $T_{/K}$ in terms of the invariant factors for T.
(Remark: to do this in the general case requires knowledge of separable versus inseparable field extensions. See the following subsection for some basic definitions.)
b) Show in particular that the mapping from sequences of invariant factors over F to sequences of invariant factors over K is injective.
c) Use part b) to give another proof of Theorem 10.4.

Corollary 10.5. *Let $P(t) \in F[t]$ be a polynomial which factors into* distinct *linear factors in an algebraic closure $\overline{F}$ of F.*[1] *If $T_1, T_2 \in \operatorname{End} V$ each have minimal polynomial P, then T_1 and T_2 are similar.*

Proof. By hypothesis, over $\overline{F}$ we have $P(t) = (t - \alpha_1) \cdots (t - \alpha_n)$ for distinct $\alpha_1, \ldots, \alpha_n$. It follows that T_1 and T_2 are both diagonalizable with diagonal entries $\alpha_1, \ldots, \alpha_n$, so they are similar over $\overline{F}$. By Theorem 10.4, T_1 and T_2 are similar. $\square$

10.2. **Semisimplicity Versus Potential Diagonalizability.**

Let V/F be a finite-dimensional vector space and $T \in \operatorname{End} V$. A field extension K/F is a **splitting field** for T if the characteristic polynomial $\chi(t)$ splits into linear factors over K. (There is a notion of splitting field of a polynomial in field theory. They are related but not the same: a splitting field K for T is a field extension of

[1] Such polynomials are called "separable" and are analyzed in the next subsection.

F *containing* a splitting field of $\chi(t)$.) Every algebraically closed extension of F is a splitting field for T. In particular, if F is a subfield of $\mathbb{C}$ then $\mathbb{C}$ is a splitting field for every $T \in \operatorname{End} V$, and this is a popular choice in many circles.

We say that T is **potentially diagonalizable** if there is a field extension K/F such that $T_{/K} \in \operatorname{End} V_{/K}$ is diagonalizable.

Exercise 9.4: We say $T \in EndV$ is **potentially triangularizable** if there is a field extension K/F such that $T_{/K}$ is triangularizable.
a) Show that in fact every T is potentially triangularizable.
b) Show that the field extensions K/F over which K is triangularizable are precisely the splitting fields K/F.
c) Deduce that if $T_{/K}$ is diagonalizable, then K is a splitting field for T.

Exercise 9.5: Let $P(t) \in F[t]$ be a nonzero polynomial. Show TFAE:
(i) For every field extension K/F, $P \in K[t]$ is squarefree.
(ii) $P \in \overline{F}[t]$ is squarefree.
(iii) In some splitting field K/F, $P \in K[t]$ is squarefree.
(iv) In every splitting field K/F, $P \in K[t]$ is squarefree.
(v) $\gcd(P, P') = 1$.
A polynomial satisfying these equivalent conditions is **separable**.

A field F is **perfect** if every prime polynomial $f \in F[t]$ is separable.

Proposition 10.6.
a) Every field of characteristic 0 *is perfect.*
Henceforth we suppose that F *has characteristic* $p > 0$.
b) F *is perfect iff: for all* $x \in F$, *there is* $y \in F$ *with* $y^p = x$.
c) Every algebraically closed field is perfect.
d) Every finite field is perfect.
e) If k *has characteristic* $p > 0$, *then the rational function field* $k(t)$ *is not perfect.*

Proof. Suppose $f \in F[t]$ is a prime polynomial. Then $\gcd(f, f') \neq 1 \iff f \mid f'$. Since $\deg f' < \deg f$, this can only happen if $f' = 0$.
a) Since $(t^n)' = nt^{n-1}$, in characteristic 0 the derivative of a polynomial of degree $n > 0$ has degree $n - 1 \geq 0$, so the derivative of a prime polynomial cannot be zero.
b) Ithere is $x \in F$ such that $y^p \neq x$ for all $y \in F$, then the polynomial $t^p - x$ is prime: let $\gamma \in \overline{F}$ be such that $\gamma^p = x$, so in $\overline{F}$, $t^p - x = (t - \gamma)^p$. Therefore any nontrivial prime factor of $t^p - x$ must be of the form $(t - \gamma)^i$ for some $0 < i < p$. But the coefficient of t^{i-1} in $(t - \gamma)^i$ is $-i\gamma$, which does not lie in F since γ does not lie in F.

Conversely, suppose that every $x \in F$ is of the form y^p for some $y \in F$, and let $f \in f(t)$ be a prime polynomial. As above, the only way for f not to be separable is $f' = 0$; since $(t^n)' = 0 \iff p \mid n$, we find that $f' = 0$ iff $f(t) = g(t^p)$ is a polynomial *in* t^p. If $g(t) = a_d t^d + \ldots + a_1 t + a_0$, then since every element of F is a power of p, we may write $a_i = b_i^p$ for all i and then $f(t) = g(t^p) = (b_d t^d + \ldots + b_1 t + b_0)^p$, so f is not a prime polynomial.
c) If F is algebraically closed then for all $a \in F$, $t^p - a$ has a root.
d) For any field of characteristic p, the map $x \mapsto x^p$ is a field endomorphism of F (this is because of the "schoolboy binomial theorem: $(x + y)^p = x^p + y^p$ in

characteristic p, since $p \mid \binom{p}{i}$ for all $0 < i < p$), hence injective. An injective map from a finite set to itself is surjective.
e) There is no rational function $r = \frac{f(t)}{g(t)}$ with $r^p = t$: e.g. because the degree of the numerator minus the degree of the denominator would have to be $\frac{1}{p}$! $\square$

Proposition 10.7. *Let F be a field.*
a) Every nonzero separable polynomial is squarefree.
b) The following are equivalent:
(i) F is perfect.
(ii) Every nonzero squarefree polynomial is separable.

Proof. a) If for some prime polynomial p, $p^2 \mid f$, then an easy application of the product rule shows $p \mid \gcd(f, f')$.
b) (i) $\implies$ (ii): Suppose F is perfect and $f = p_1 \cdots p_r$ is a product of distinct primes. Then $f' = p_1' p_2 \cdots p_r + \ldots + p_1 \cdots p_{r-1} p_r'$. So for each $1 \leq i \leq r$, $p_i \mid f' \iff p_i \mid p_i' \iff p_i' = 0$. Since f is perfect, these conditions don't hold.
$\neg$ (i) $\implies$ $\neg$ (ii): If F is not perfect, there is a prime polynomial which is not separable. Prime polynomials are squarefree. $\square$

Theorem 10.8. *a) For any $T \in \operatorname{End} V$, the following are equivalent:*
(i) T is diagonalizable over every splitting field K.
(ii) T is diagonazliable over some splitting field K.
(iii) T is potentially diagonalizable.
(iv) The minimal polynomial $P(t)$ is separable.
b) For any $T \in \operatorname{End} V$, if T is potentially diagonalizable then it is semisimple.
c) For a field F, the following are equivalent:
(i) F is perfect.
(ii) Every semisimple linear endomorphism on a finite-dimensional F-vector space is potentially diagonalizable.

Proof. a) (i) $\implies$ (ii) $\implies$ (iii) is immediate.
(iii) $\implies$ (iv): Suppose K/F is a field extension such that $T_{/K}$ is diagonalizable. By Proposition 10.1, the minimal polynomial of T over K is simply P viewed as a polynomial with coefficients in K, so by the Diagonalizability Theorem, P splits into distinct linear factors in K. By Exercise 9.5, P is separable.
(iv) $\implies$ (i): Let K/F be a splitting field for P. Since P is separable, by Exercie 9.3, P splits into distinct linear factors in K and then $T_{/K}$ is diagonalizable by the Diagonalizability Theorem.
b) Since T is potentially diagonalizable iff its minimal polynomial is separable and semisimple iff its minimal polynomial is squarefree, this follows immediately from Proposition 10.7a).
c) (i) $\implies$ (ii): By the Semisimplicity Theorem, $T \in \operatorname{End} V$ is semisimple iff it has squarefree minimal polynomial, whereas by part a), T is potentially diagonalizable iff it has separable minimal polynomial. By Proposition 9.5, if F is perfect then squarefree = separable, so semisimple = potentially diagonalizable. Conversely, if F is not perfect there is a prime polynomial p which is not separable, and by Example 3.1 there is an endomorphim T with minimal polynomial p, hence a semisimple but not potentially diagonalizable endomorphism. $\square$

11. The Dual Endomorphism

11.1. **Review of Dual Spaces.**

For any set S, let F^S be the set of all functions $f : S \to F$. For $f, g \in F^S$ and $\alpha \in F$, we define

$$\alpha f + g : s \in S \mapsto \alpha f(s) + g(s).$$

Exercise 10.1: a) Show that F^S is an F-vector space.
b) For $s \in S$, let $\delta_s \in F^S$ be the function which maps s to 1 and every other element of S to 0. Show that $\Delta = \{\delta_s\}_{s \in S}$ is a linearly independent set.
c) Deduce that if S is infinite, F^S is infinite-dimensional.
d) Show that Δ is a basis for F^S iff S is finite.

Now let V be an F-vector space, for the moment not assumed to be finite-dimensional. Inside F^V we have the subset V^* of F-linear maps $f : V \to F$. Such maps are also called **linear functionals** on V.

Exercise 10.2: Show that V^* is a linear subspace of F^V, called the **dual space**.

For $\ell \in V^*$ and $v \in V$, we denote write $\langle \ell, v \rangle$ for $\ell(v)$.

By V^{**} we mean $(V^*)^*$, i.e., the space of linear functionals *on* the space of linear functionals on V. There is a canonical map

$$\iota : V \to V^{**}, v \mapsto (\ell \mapsto \langle \ell, v \rangle).$$

Lemma 11.1. *The map $\iota : V \to V^{**}$ is an injection.*

Proof. It is enough to show that for $v \in V^\bullet$, $\iota(v) \neq 0$: explicitly, there is a linear functional $\ell : V \to F$ such that $\ell(v) \neq 0$. But since $v \neq 0$ there is a basis $\mathcal{B}$ of V containing v, and then we can define ℓ by putting $\ell(v) = 1$ and defined arbitrarily on every other basis element (i.e., it does not matter how it is defined). $\square$

Suppose now that V is finite-dimensional, and let $e_1, \ldots, e_d$ be a basis. For $1 \leq i \leq d$ there is a unique linear functional $e_i^* : V \to F$ which maps e_i to 1 and every other basis element to 0. Suppose $\alpha_1, \ldots, \alpha_n \in F$ are such that $\alpha_1 e_1^* + \ldots + \alpha_n e_n^* = 0$. Evaluating at e_i we get $\alpha_i = 0$, so $e_1^*, \ldots, e_n^*$ are linearly independent. If $\ell \in V^*$ is any linear functional, then ℓ and $\ell(e_1)e_1^* + \ldots + \ell(e_n)e_n^*$ agree when evaluated at each basis element e_i so are equal. Thus $\langle e_1^*, \ldots, e_n^* \rangle = V^*$, so $e_1^*, \ldots, e_n^*$ is a basis for V^*, called the **dual basis** to $e_1, \ldots, e_n$. From this we deduce:

Corollary 11.2. *Let V be a finite-dimensional vector space. Then:*
a) $V^* \cong V$.
b) The map $\iota : V \to V^{**}$ is an isomorphism.

Proof. a) The above analysis shows that if V has finite dimension d then so does V^*, and any two vector spaces of the same dimension are isomorphic.
b) By part a), $\dim V^{**} = \dim (V^*)^* = \dim V^* = \dim V$. By Lemma 10.1, ι is an injective linear map between two vector spaces of the same finite dimension, hence is an isomorphism by the Dimension Theorem. $\square$

Remark 11.3. *It turns out that for every infinite-dimensional vector space V,*

$$\dim V^* > \dim V; \tag{3}$$

this is an inequality of infinite cardinal numbers. Thus $\dim V^{**} > \dim V^* > \dim V$*. In particular the canonical injection ι is not an isomorphism. However, establishing (3) would take us too far from what we want to discuss here.*

11.2. **The Dual Endomorphism.**

Let $T \in \operatorname{End} V$. There is a **dual endomorphism** T^* of $\operatorname{End} V^*$: if $\lambda : V \to F$, we put $T^*\lambda = \lambda \circ T : V \to F$. In other words, for all $\lambda \in V^*$ and $v \in V$,

$$\langle T^*\lambda, v\rangle = \langle \lambda, T(v)\rangle.$$

Exercise 10.3: More generally, let $L : V \to W$ be any F-linear map. Then there is a map $L^* : W^* \to V^*$, given by

$$(\ell : W \to F) \mapsto L^*\ell = \ell \circ L : V \to F.$$

a) Show that L^* is F-linear.
b) Show: if L is injective, L^* is surjective.
c) Show: if L is surjective, L^* is injective.

Lemma 11.4. *a) For all $f \in F[t]$, $\lambda \in V^*$ and $v \in V$, we have*

$$\langle f(T^*)\lambda, v\rangle = \langle \lambda, f(T)v\rangle.$$

b) The minimal polynomial P^* of T^* on V^* is equal to the minimal polynomial P of T on V.

Exercise 10.4: Prove Lemma 11.4.

For a subspace $W \subset V$ and $\lambda \in V^*$, we will write $\langle \lambda, W\rangle = 0$ as an abbreviation for "For all $w \in W$, $\langle \lambda, w\rangle = 0$. For $W' \subset V^*$ and $v \in V$ the notation $\langle W', v\rangle = 0$ is defined similarly.

For a subspace $W \subset V$ we put

$$W^\perp = \{\lambda \in V^* \mid \langle \lambda, W\rangle = 0\}.$$

Exercise 10.5: a) Show that $W^\perp$ is a linear subspace of V^*.
b) If W is T-invariant, show that $W^\perp$ is T^*-invariant.

Suppose we have a direct sum decomposition $V = W_1 \oplus W_2$. Let $\pi_1 : V \to W_1$ be given by $(w_1, w_2) \mapsto w_1$. Consider the dual map $\pi_1^* : W_1^* \to V^*$. Since π_1 is surjective, by Exercise 10.3c), π_1^* is injective. We claim $\pi_1^*(W_1^*) = W_2^\perp$. Indeed:

Let $w_2 \in W_2$ and $\lambda : W_1 \to F$. Then

$$\langle \pi_1^*(\lambda), (0, w_2)\rangle = (\lambda \circ \pi_\cdot)(0, w_2) = \lambda(0) = 0.$$

Conversely, let $\lambda : V \to F$ be such that $\lambda(W_2) = 0$. Let $\tilde{\lambda} = \lambda|_{W_1}$. Then for all $v = (w_1, w_2) \in V$,

$$\pi_1^*(\tilde{\lambda})(w_1, w_2) = \tilde{\lambda}(\pi_1(w_1, w_2)) = \tilde{\lambda}(w_1) = \lambda(w_1)$$

so $\pi_1^*(\tilde{\lambda}) = \lambda$.

Theorem 11.5. *Suppose $V = W_1 \oplus W_2$.*
a) There are canonical isomorphisms

$$\Phi = \pi_1^* : W_1^* \xrightarrow{\sim} W_2^\perp$$

and

$$\Psi = \pi_2^* : W_2^* \xrightarrow{\sim} W_1^\perp.$$

b) Suppose W_1 and W_2 are invariant subspaces. Then

$$T^*|_{W_2^\perp} \circ \Phi = \Phi \circ (T|_{W_1})^*$$

and

$$T^*|_{W_1^\perp} \circ \Psi = \Psi \circ (T|_{W_2})^*$$

c) Thus $T^* \in \operatorname{End} W_2^\perp$ and $T \in \operatorname{End} W_1$ are similar endomorphisms, as are $T^* \in \operatorname{End} W_1^\perp$ and $T \in \operatorname{End} W_2$.

Proof. In each part it suffices to prove the first assertion; the second one follows from interchanging the roles of W_1 and W_2.
a) This has been shown above.
b) Let $\lambda : W_1 \to F$ and $v = w_1 + w_2 \in W_1 \oplus W_2 = V$. Then

$$\langle (T^*|_{W_2^\perp} \circ \Phi)(\lambda), w_1 + w_2 \rangle = \langle \pi_1^* \lambda, T(w_1) + T(w_2) \rangle$$

$$= \langle \lambda \circ \pi_1, T(w_1) + T(w_2) \rangle = \lambda(T(w_1))$$

$$= \langle \lambda, T|_{W_1} w_1 \rangle = \langle (T|_{W_1})^* \lambda, w_1 \rangle$$

$$= \langle (T|_{W_1})^* \lambda, \pi_1(w_1 + w_2) \rangle = \langle \pi_1^* (T_{W_1})^* \lambda, w_1 + w_2 \rangle = \langle \Phi (T|_{W_1})^* \lambda, w_1 + w_2 \rangle.$$

c) Part b) asserts that $T^* \in \operatorname{End} W_2^\perp$ and $T \in \operatorname{End} W_1$ are similar, so the result follows from Theorem 8.3. □

Corollary 11.6. *Let $W \subset V$ be a subspace. Then:*
a) $\dim W^\perp = \dim V - \dim W$*.*
b) $W^{\perp\perp} = W$*.*

Proof. a) This follows from Theorem 11.5a).
b) $W^{\perp\perp}$ is the set of $v \in V$ such that for all $\ell \in V^*$ such that $\ell(W) = 0$, $\ell(v) = 0$; clearly this contains W! By part a), W and $W^{\perp\perp}$ have the same finite dimension, so the inclusion $W \subset W^{\perp\perp}$ must be an equality. □

Corollary 11.7. *If $V = W_1 \oplus W_2$ with W_1 and W_2 invariant subspaces, then $V^* = W_2^\perp \oplus W_1^\perp$.*

Proof. Step 1: Suppose $\ell \in W_1^\perp \cap W_2^\perp$. Let $v \in V$ and write $v = w_1 + w_2$ with $w_i \in W_i$. Then $\ell(v) = \ell(w_1) + \ell(w_2) = 0$. So $\ell = 0$. It follows that $W_2^\perp + W_1^\perp = W_2^\perp \oplus W_1^\perp$.
Step 2: By Corollary 11.6 we have

$$\dim W_2^\perp + \dim W_1^\perp = \dim V - \dim W_2 + \dim V - \dim W_1 = \dim V = \dim V^*.$$

Since V is finite-dimensional, we conclude $V^* = W_2^\perp \oplus W_1^\perp$. □

Theorem 11.8. *The elementary divisors of T on V are those of T^* on V^*.*

Proof. Let $V = \bigoplus_{i=1}^n C_i$ be a decomposition into primary cyclic subspaces. We go by induction on n.
Base Case ($n = 1$): Suppose V is cyclic and primary, so the minimal and characteristic polynomials are both equal to $p(t)^a$ for some prime polynomial $p(t)$ and $a \in \mathbb{Z}^+$, and $p(t)^a$ is the only elementary divisor. The minimal polynomial of V^* is also $p(t)^a$ and its degree is $\dim V = \dim V^*$, so V^* is also cyclic and primary with $p(t)^a$ as its only invariant factor.
Induction Step: Let $n \geq 2$ and assume the result when V has fewer than n elementary divisors. Put $W_2 = \bigoplus_{i=2}^n C_i$, so $V = C_1 \oplus W_2$: here C_1 is cyclic with minimal polynomial $p_1^{a_1}$. By Corollary 11.7, $V^* = W_2^\perp \oplus C_1^\perp$. By Theorem 11.5b), $W_2^\perp$ is cyclic with minimal polynomial $p_1^{a_1}$, so $p_1^{a_1}$ is an elementary divisor of V^*. By Theorem 11.5b), the endomorphism T on W_2 is similar to the endomorphism T^* on $C_1^\perp$, hence they have the same elementary divisors. We are done by induction. $\square$

11.3. **Jacob's Proof of the Existence of a Cyclic Decomposition.**

Above we gave M.D. Burrow's nice "early" proof that the minimal polynomial of any linear transformation on a finite dimensional vector space V has degree at most $\dim V$. Burrow's article [Bu73] also give nice inductive proofs of the facts that if V is p-primary, $\deg p \mid \dim V$ and that that Cayley-Hamilton polynomial is equal to the characteristic polynomial. It happens that the following article [Ja73] in the December 1973 issue of the Monthly, which begins on the same page as Burrow's article ends, contains a nice proof of the existence part of the Cyclic Decomposition Theorem using dual spaces. We reproduce Jacob's proof here.

It is enough to show: for any $v \in v$ with $P_v = P$, there is an invariant subspace $W' \subset V$ with $V = [v] \oplus W'$.

To show this: let $k = \deg P$, $v_1 = v, v_2 = Tv, \ldots, v_k = T^{k-1}v_1$, so $v_1, \ldots, v_k$ is a basis for $[v]$. Extend it to a basis $v_1, \ldots, v_d$ for V. Let $v_1^*, \ldots, v_k^*$ be the dual basis of V^*, and put $v^* = v_k^*$. Then

$$\forall 1 \leq i \leq k-1,\ \langle v^*, v_i \rangle = 0,$$

$$\langle v^*, v_k \rangle = 1.$$

Let $U = \langle v^*, T^*v^*, \ldots, T^{*k-1}v^* \rangle$. Since $P^* = P$, U is T^*-invariant.
CLAIM: $U \cap [v]^\perp = 0$ and $\dim U = k$. If so, $V^* = U \oplus [v]^\perp$ and then

$$V = V^{**} = [v]^{\perp\perp} \oplus U^\perp = [v] \oplus U^\perp.$$

PROOF OF CLAIM: if either $\dim U < k$ or $U \cap [v]^\perp \neq 0$, there is $0 \leq s \leq k-1$ such that $a_s \neq 0$ and

$$a_0 v^* + a_1 T^* v^* + \ldots + a_s T^{*s} v^* \in [v]^\perp.$$

Then

$$T^{*k-1-s}\left(a_0 v^* + a_1 T^* v^* + \ldots + a_s T^{*s} v^*\right)$$
$$= a_0 T^{*k-1-s} v^* + a_1 T^{*k-s} v^* + \ldots + a_s T^{*k-1} v^* \in [v]^\perp.$$

Thus

$$0 = \langle (a_0 T^{*k-1-s} + a_1 T^{*k-s} + \ldots + a_s T^{*k-1}) v^*, v \rangle = \langle v^*, (a_0 T^{k-1-s} + \ldots + a_s T^{k-1}) v \rangle$$
$$= a_0 \langle v^*, v_{k-s} + \ldots + a_s \langle v^*, v_k \rangle = a_s,$$

a contradiction.

References

[A] S. Axler, *Linear algebra done right.* Second edition. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1997.

[Ax95] S. Axler, *Down with determinants!* Amer. Math. Monthly 102 (1995), no. 2, 139–154.

[BB03] K. Boulabiar and G. Buskes, *After the determinants are down: a criterion for invertibility.* Amer. Math. Monthly 110 (2003), no. 8, 737–741.

[Bu73] M.D. Burrow, *The minimal polynomial of a linear transformation.* Amer. Math. Monthly 80 (1973), 1129–1131.

[Cl12] P.L. Clark, *Covering numbers in linear algebra.* Amer. Math. Monthly 119 (2012), no. 1, 65–67.

[Co] K. Conrad, *The fundamental theorem of algebra via linear algebra.* `http://www.math.uconn.edu/~kconrad/blurbs/fundthmalg/fundthmalglinear.pdf`

[De03] H. Derksen, *The fundamental theorem of algebra and linear algebra.* Amer. Math. Monthly 110 (2003), no. 7, 620–623.

[FT] P.L. Clark, *Field Theory.*

[Ja73] H.G. Jacob, *Another proof of the rational decomposition theorem.* Amer. Math. Monthly 80 (1973), 1131–1134.

Department of Mathematics, Boyd Graduate Studies Research Center, University of Georgia, Athens, GA 30602-7403, USA

Email address: `pete@math.uga.edu`