Lecture Notes: Linear Algebra I.
Fall 2014

by Prof. Frederick P. Greenleaf
NYU/Courant Institute of Math Sciences

Prepared with the gratefully acknowledged assistance of

Dr. Sophie Marques

without whose efforts the Latex version of these Course Notes might never have seen the light of day.

Contents

Lecture Notes: Linear Algebra II. Fall 2014

by Prof. Frederick P. Greenleaf
and Dr. Sophie Marques
NYU/Courant Institute of Math Sciences

Prof. Greenleaf gratefully acknowledges the assistance of Dr. Sophie Marques without whose efforts the Latex version of these *Course Notes* might never have seen the light of day.

Contents

Chapter 0. A Few Preliminaries.

Course texts:

1. *Linear Algebra*, by S. Friedberg, A. Insel, L. Spence (latest edition). This will be the main backup text to accompany the present *Class Notes* . Various assignments will be taken from it.

2. Schaum's Outline Series: *Linear Algebra*, by Seymour Lipschutz, for a review of matrix algebra, row operations, and solution of linear systems (roughly the first 3-4 chapters). $\mathbb{K}$ is a field (see Appendix C of [F/I/S] text; read it). For us, $\mathbb{K} = \mathbb{C}, \mathbb{R}, \mathbb{Q}$, and occasionally the finite field $\mathbb{K} = \mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$, for a prime $p > 1$.

Recall that the finite field $\mathbb{Z}_p$ is modeled as $S = \{0, 1, 2, ..., p-1\}$, interpreting $a + b$ and ab (mod p). For example: if $p = 7$ then

$$5 \oplus 6 = 11 \equiv 4 \pmod 7 \qquad \text{and} \qquad 506 \odot 17 \equiv 6 \pmod 7.$$

Elements of $\mathbb{Z}_p$ are the (mod p) congruence classes $[k] = k + p\mathbb{Z} = \{\ell : \ell \equiv k \pmod p\}$. Using this notation the operations in $\mathbb{Z}_p$ take the form

$$[a] \oplus [b] = [a + b] \qquad [a] \odot [b] = [ab] \qquad \text{(add or multiply class representatives)}$$

The system $(\mathbb{Z}_p, \oplus, \odot)$ is a finite number field with additive zero element [0] and multiplicative identity element [1]. All nonzero elements $[k] \neq [0]$ have multiplicative inverses (reciprocals), but it may not be so easy to find the class $[k]^{-1} = [\ell],\ 0 < \ell < p$ such that $[k] \cdot [\ell] = [1]$. If $p = 7$ we have $[3]^{-1} = [5]$ because $3 \odot 5 = 15 = 14 + 1 \equiv 1 \pmod 7$. Notice that in $\mathbb{Z}_p$ the sum $[1] \oplus [1] \oplus .... \oplus [1]$ with p terms is equal to the zero element [0].

Chapter I

Section I.1. Vector Spaces over a Field $\mathbb{K}$

The objects of interest in this chapter will be vector spaces over arbitrary fields.

1.1 Definition. *A* **vector space over a field** $\mathbb{K}$ *is a set* V *equipped with two operations* $(+)$ *and* $(\cdot)$ *from* $V \times V \to V$ *and* $\mathbb{K} \times V \to V$ *having the following properties.*

1. **Axioms for** $(+)$**:**
   - COMMUTATIVE LAW: $x + y = y + x$
   - ASSOCIATIVE LAW: $(x + y) + z = x + (y + z)$
   - ZERO ELEMENT: There exists an element "0" in V such that $0+v = v$ for all v.
   - ADDITIVE INVERSE: For every $v \in V$ there is an element $-v \in V$, such that $v + (-v) = 0$.
2. **Axioms for** $(\cdot)$**:**
   - IDENTITY LAW: $1 \cdot v = v$ ($1 =$ the identity in $\mathbb{K}$)
   - ASSOCIATIVE LAW: $(ab) \cdot v = a \cdot (b \cdot v)$ for $a, b \in \mathbb{K},\ v \in V$
   - DISTRIBUTIVE LAW: $a \cdot (x + y) = (a \cdot x) + (b \cdot y)$
   - DISTRIBUTIVE LAW: $(a + b) \cdot x = (a \cdot x) + (b \cdot x)$

As a consequence,

1.2. Lemma. *The zero element is unique: if* $0, 0' \in V$ *are elements such that* $0 + v = v$ *and* $0' + v = v$*, for all* $v \in V$*, then* $0' = 0$.

Proof: $0 + 0' = 0'$ and $0 + 0' = 0$, so $0' = 0$. $\square$

1.3. Lemma. *The additive inverse is unique. That is, given* $v \in V$ *there is just one element* $u \in V$ *such that* $u + v = 0$.

Proof: Suppose $v \in V$ and we are given u and u' with $u + v = 0$ and $u' + v = 0$. Look at the combination $u + v + u'$; by associativity we get

$$u' = 0 + u' = (u + v) + u' = u + (v + u') = u + 0 = u \quad ,$$

so that $u' = u$. $\square$

1.4. Exercise. From the axioms and previous results prove:

(i) $0 \cdot v = 0_V$ (ii) $\lambda \cdot 0_V = 0_V$ (iii) $\lambda \cdot v = v$ and $v \neq 0_V \Rightarrow \lambda = 1$.

1.5. Exercise. Prove that $-v = (-1) \cdot v$ where -1 is the negative of $1 \in \mathbb{K}$.
Hint: $1 + (-1) = 0$ in $\mathbb{K}$ and $0 \cdot v = 0_V$. Remember: "$-v$" is the unique element that added to v is 0_V; prove that $(-1) \cdot v$ has this property and conclude by uniqueness of additive inverse.

1.6. Exercise. Prove that $-(-v) = v$, for all $v \in V$.
Hint: Same as the previous exercise.

1.7. Exercise (Cancellation Laws). If $a+v = a+w$ for $a, v, w \in V$ prove that $v = w$. Then use this to prove

(i) $\lambda \cdot v = 0_V$ and $v \neq 0_V$ implies that $\lambda = 0$ in $\mathbb{K}$

(ii) $\lambda \cdot v = v$ and $v \neq 0_V$ implies $\lambda = 1$.

1.8. Example. "*Coordinate space*" over the field $\mathbb{K}$ consists of all ordered n-tuples $\mathbb{K}^n = \{\mathbf{x} = (x_1, ..., x_n) : x_k \in \mathbb{K}\}$, equipped with the usual $(+)$ and $(\cdot)$ operations:

(i) $(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n)$

(ii) $\lambda \cdot (x_1, ..., x_n) = (\lambda \cdot x_n, ..., \lambda \cdot x_n)$ for $\lambda \in \mathbb{K}$. $\square$

1.9. Exercise. Explain why $(+)$ in $\mathbb{R}^2$ is described geometrically by the "parallelogram law" for vector addition shown in Figure 1.1.

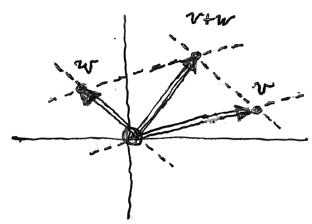

Figure 1.1. The Parallelogram Law for vector addition, illustrated in $\mathbb{R}^2$.

1.10. Example (Matrix Space). The space $\mathrm{M}(n \times m, \mathbb{K})$ of $n \times m$ matrices with entries in $\mathbb{K}$ becomes a vector space when equipped with the operations

$$\text{Addition Operation:} \quad (A + B)_{ij} = A_{ij} + B_{ij}$$

$$\text{Scaling Operation:} \quad (\lambda \cdot A)_{ij} = \lambda A_{ij}$$

The space of square matrices, with $m = n$, is denoted $\mathrm{M}(n, \mathbb{K})$.
Notation: Matrix entry A_{ij} is the one in the i^{th} row and j^{th} column. The pair (i, j) is referred to as its "address." $\square$

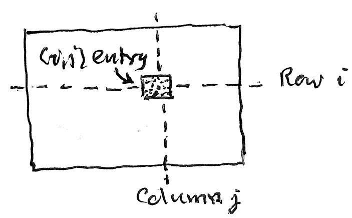

Figure 1.2. The entry in a matrix array with "address" (i, j) is the one in Row i and Column j.

There is also a matrix multiplication that makes $\mathrm{M}(n, \mathbb{K})$ an associative algebra with identity, but the matrix product AB can be defined more generally for non-square matrices as long as they are "compatible," with the number of columns in A equal to the number of rows in B. Thus if A is $m \times q$ and B is $q \times n$ we get an $m \times n$ matrix AB with entries

$$(AB)_{ij} = \sum_{k=1}^{q} A_{ik} B_{kj}$$

The algebra $\mathrm{M}(n, \mathbb{K})$ of square matrices is not commutative unless $n = 1$. $\square$

1.11. Example (Polynomial Ring $\mathbb{K}[x]$). The set $\mathbb{K}[x]$ consists of all finite "formal sums" $a_0 + a_1 x + ... + a_n x^n + ... = \sum_{k \geq 0} a_k x^k$ with $a_i \in \mathbb{K}$, and $a_i = 0$ for all but a finite number of indices. These sums can have arbitrary length. They include the "constant polynomials" which have form $c \cdot \mathbb{1}$ with $c \in \mathbb{K}$, where $\mathbb{1}$ is the particular constant polynomial $1 + 0 \cdot x + 0 \cdot x^2 + \ldots$; the zero polynomial $0 \cdot \mathbb{1}$ is written as "0", which might get confusing.

The algebraic operations in $\mathbb{K}[x]$ are

1. Addition: $(\sum_{k\geq 0} a_k x^k) + (\sum_{k\geq 0} b_k x^k) \;=\; \sum_{k\geq 0}(a_k + b_k)x^k$

2. Scaling: $\lambda \cdot (\sum_{k\geq 0} a_k x^k) \;=\; \sum_{k\geq 0}(\lambda a_k)x^k$.

There is also a multiplication operation, obtained by multiplying terms in the formal sums and gathering together those of the same degree

3. Product: $\Big(\sum_{k\geq 0} a_k x^k\Big) \times \Big(\sum_{l\geq 0} b_l x^l\Big) = \sum_{k,l\geq 0} a_k b_l\, x^{k+l} = \sum_{r\geq 0}\Big(\sum_{k,l\geq 0, k+l=r} a_k b_l\Big)\cdot x^r$

(the sum being finite for each r). This makes $\mathbb{K}[x]$ into a commutative associative algebra over $\mathbb{K}$ with $\mathbb{1}$ as its multiplicative identity.

All information about a polynomial resides in the symbol string $(a_0, a_1, a_2, ...)$ of coefficients, and the algebraic operations on $\mathbb{K}[x]$ can be performed as operations on symbol strings; the zero polynomial is represented by $(0, 0, ...)$, the identity by $\mathbb{1} = (1, 0, ..., 0)$, and x by $x = (0, 1, 0, ...)$, etc. $\square$

1.12. Exercise. If $f(x) = 3 + 3x + x^2$ and $g(x) = 4x^2 - 2x^3 + x^5$, compute the sum $f + g$ and product $f \cdot g$.

The **degree** $\deg(f)$ of $f = \sum_{k\geq 0} a_k x^k$ is n if $a_n \neq 0$ and $a_k = 0$ for all $k > n$. The degree of a constant polynomial $c\mathbb{1}$ is zero, except that no "degree" can be assigned to the zero polynomial 0. (For various reasons, the only possible assignment would be "$-\infty$").

1.13. Exercise. If $f, g \neq 0$ in $\mathbb{K}[x]$ prove that $fg \neq 0$ and $\deg(fg) = \deg(f) + \deg(g)$.

1.14. Exercise. If $f, g \neq 0$ in $\mathbb{K}[x]$, what (if anything) can you say about $\deg(f + g)$?

1.15. Example (Polynomials in Several Unknowns). The polynomial ring $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, ..., x_n]$ is handled using very efficient "multi-index notation." A **multi-index** is an element $\alpha = (\alpha_1, \ldots \alpha_n)$ of the Cartesian product set $\mathbb{Z}_+^n = \mathbb{Z}_+ \times ... \times \mathbb{Z}_+$ (n factors). Each multi-index determines a **monomial** $x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$, in which we interpret $x_k^0 = 1$. Elements of $\mathbb{K}[x_1, \ldots, x_n]$ are finite formal linear combinations of monomials

$$f(x_1, \ldots, x_n) = \sum_{\alpha\in\mathbb{Z}_+^n} c_\alpha x^\alpha \qquad (c_\alpha \in \mathbb{K})$$

The monomial $x^{(0,\ldots,0)}$ is the constant polynomial $\mathbb{1}$ in $\mathbb{K}[x_1, ..., x_n]$. With these ideas in mind,

1. The **total degree** of a multi-index is $|\alpha| = \alpha_1 + \ldots + \alpha_n$ and the degree of the corresponding monomial is $\deg(x^\alpha) = |\alpha|$. Note that many monomials can have same total degree, for example x^2y and xy^2.

2. The **degree of a polynomial** $f \in \mathbb{K}[\mathbf{x}]$ is

$$\deg(f) = \max\{\, |\alpha| : c_\alpha \neq 0 \,\}$$

   Nonzero constant polynomials $c\mathbb{1}$ have degree zero: if f is the zero polynomial (all coefficients $c_\alpha = 0$) $\deg(f)$ cannot be defined. The generators $f_k(\mathbf{x}) = x_k$ of the polynomial ring all have degree 1.

The following operations make $V = \mathbb{K}[\mathbf{x}]$ a vector space and a commutative associative algebra with identity $\mathbb{1} = x^{(0,\ldots,0)}$.

1. Sum: $f + g = \sum_\alpha (a_\alpha + b_\alpha)\, x^\alpha$

2. Scaling: $\lambda \cdot f = \sum_\alpha (\lambda a_\alpha) x^\alpha$

3. Product Operation:

$$\begin{aligned} f \cdot g &= (\sum_\alpha a_\alpha x^\alpha) \cdot (\sum_\beta b_\beta x^\beta) \\ &= \sum_{\alpha,\beta \in \mathbb{Z}_+^n} a_\alpha b_\beta x^{\alpha+\beta} \\ &= \sum_{\gamma \in \mathbb{Z}_+^n} \Big(\sum_{\alpha+\beta=\gamma} a_\alpha \cdot b_\beta \Big) \cdot x^\gamma \end{aligned}$$

where we define a "sum of exponents" to be $\alpha + \beta = (\alpha_1 + \beta_1, ..., \alpha_n + \beta_n)$.

As an example, the monomials of degree 2 in $\mathbb{K}[x_1, x_2, x_3]$ are

multi-index	monomial
$(0,0,2)$	x_3^2
$(0,1,1)$	x_2x_3
$(0,2,0)$	x_2^2
$(1,0,1)$	x_1x_3
$(1,1,0)$	x_1x_2
$(2,0,0)$	x_1^2

Here we have lined up the monomials in "lexicographic"or "dictionary" order (taking $A = 0, B = 1, C = 2, \ldots$), which is a useful way to manage them. This is a strict *linear* ordering of monomials; they are only partially ordered by their "total degree" $\deg(x^\alpha) = |\alpha|$. The system $\mathbb{K}[x_1, \ldots, x_n]$ is a commutative associative algebra with identity element $\mathbb{1}$. Its properties are quite a bit more complicated than those of polynomials $\mathbb{K}[x]$ in one unknown, but they do share two crucial algebraic properties. $\square$

1.16. Exercise. (Hard, but try it) If $f, g \neq 0$ in $\mathbb{K}[x_1, ..., x_n]$ prove that

1. Degree Formula: $\deg(f \cdot g) = \deg(f) + \deg(g)$ for all $f, g \neq 0$ in $\mathbb{K}[x_i, \ldots, x_n]$.

2. No Zero Divisors: $f, g \neq 0$ in $\mathbb{K}[x_1, \ldots, x_n] \Rightarrow f \cdot g \neq 0$. This implies we can perform "cancellation" – if $f \neq 0$ and $f \cdot h_1 = f \cdot h_2$ then $h_1 = h_2$.

Hint: Try it first for $n = 1$. For $n = 2$ try lexicographic ordering of monomials in $\mathbb{K}[x, y]$.
Note: The maximum possible degree for a nonzero monomial in the product fg is obviously $d = \deg(f) + \deg(g)$. The problem is that the coefficient c_γ of such a monomial will be a sum of products $(\sum_{\alpha+\beta=\gamma} a_\alpha b_\beta)$, and not a simple product as it is when there is just one variable. Such sums could equal zero even if all terms are nonzero, so why couldn't these coefficients (sums) be zero for *all* monomials with the maximum possible degree d, making $\deg(fg) < \deg(f) + \deg(g)$? $\square$

A more complete discussion of the Degree Formula for $n \geq 2$, and especially its proof using lexicographic ordering of monomials, is provided in Appendix A of this chapter.

1.17. Example (Function Spaces). If S is a set, $\mathcal{C}(S)$ = all scalar-valued functions $f : S \to \mathbb{K}$ become a vector space under the usual operations

$$(f + g)(x) = f(x) + g(x), \qquad (\lambda \cdot f)(x) = \lambda f(x), \ \forall x \in S$$

There is also a pointwise multiplication operation

$$(f \cdot g)(x) = f(x) \cdot g(x) \ ,$$

which makes $\mathcal{C}(S)$ a commutative associative algebra over $\mathbb{K}$ with identity element $1(x) = 1$ for all x, and zero element $0(x) = 0$, for $x \in S$. $\square$

1.18. Example (Polynomial Functions vs Formal Sums). The **polynomial functions** $\mathcal{P}_{\mathbb{K}}$ with values in $\mathbb{K}$ are the functions $\phi_f : \mathbb{K} \to \mathbb{K}$ of the form

$$\phi_f(t) = \left[\, f(x)|_{x=t} \,\right] = \sum_{k\geq 0} a_k\, t^k \qquad (t \in \mathbb{K})$$

for some $f \in \mathbb{K}[x]$. (Thus, $\phi(t) = sin(t)$ is not a polynomial function on $\mathbb{K}$). Note carefully that the elements of $\mathcal{P}_{\mathbb{K}}$ are functions while $\mathbb{K}[x]$ is made up of symbol strings or formal sums. They are not the same thing, though there is a close relation between them implemented by the surjective (="onto") mapping $\Phi : \mathbb{K}[x] \to \mathcal{P}_K$ such that

$$\Phi f(t) = \sum_{k\geq 0} a_k\, t^k \qquad (t \in \mathbb{K})$$

if $f(x) = \sum_{k\geq 0} a_k x^k$ in $\mathbb{K}[x]$. This surjective map is a *homomorphism*: it preserves, or "intertwines," the algebraic operations in $\mathbb{K}[x]$ and in the "target space" $\mathcal{P}_{\mathbb{K}}$, so that

$$\Phi(\lambda \cdot f) = \lambda \cdot \Phi(f) \qquad \Phi(f+g) = \Phi(f) + \Phi(g) \qquad \Phi(f \cdot g) = \Phi(f) \cdot \Phi(g) \quad \square$$

1.19. Exercise. If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ explain why Φ is a bijection, hence an "isomorphism" between commutative associative algebras. In fact, prove that this is so for polynomials over any *infinite* field $\mathbb{K}$.
Hint: Φ is linear, hence being one-to-one is equivalent to saying that $\Phi(f) = 0 \Rightarrow f = 0$ in $\mathbb{K}[x]$. If f is nonzero in $\mathbb{K}[x]$ the corresponding polynomial function $\Phi(f) : \mathbb{K} \to \mathbb{K}$ can take on the value zero at no more than $n = \deg(f)$ points – i.e. the number of roots in $\mathbb{K}$ cannot exceed $\deg(f)$. Since $\mathbb{R}$ and $\mathbb{C}$ (and even $\mathbb{Q}$) are infinite we cannot have $\Phi(f) \equiv 0$ on these fields unless f is the zero polynomial. $\square$

The finite fields $\mathbb{Z}_p$ (p a prime) are widely used in number theory, cryptography, image processing, etc. This one-to-one correspondence breaks down for these fields. For example if $\mathbb{K} = \mathbb{Z}_p$ the nonzero polynomial $f = x^p - x$ has value zero for every choice of $x \in \mathbb{Z}_p$ and there are precisely $p = \deg(f)$ roots.

A theorem of Fermat says: if p is a prime then $t^{p-1} = 1$ for all nonzero t in $\mathbb{Z}_p$, but then $t^p - t = t$ is zero at every $t \in \mathbb{Z}_p$ and $\Phi(f) \equiv 0$ (the zero function in $\mathcal{P}_{\mathbb{K}}$).

1.20. Exercise. For $p = 3$, verify that $t^3 - t = 0$ for the three elements $t = [0], [1], [2]$ in $\mathbb{Z}_3$. But the corresponding element of $\mathbb{Z}_3[x]$ is $f = x^3 - x$, whose symbol string $(0, -1, 0, 1, 0, 0, ...)$ differs from that of the zero polynomial in $\mathbb{Z}_3[x]$.

I.2. Vector Subspaces

2.1. Definition. *A nonempty subset W of a vector space V is a* **vector subspace** *if*

1. W is *closed under* $(+)$: $W + W \subseteq W$, so $w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W$.

2. W is *closed under* $(\cdot)$: $\mathbb{K} \cdot W \subseteq W$, so $\lambda \in \mathbb{K}, w \in W \Rightarrow \lambda \cdot w \in W$.

The vector 0 then lies in W, for if $w \in W$ then $-w = (-1) \cdot w$ is also in W and then $0 = w + (-w) \in W$. Thus W becomes a vector space over $\mathbb{K}$ in its own right under the $(+)$ and $(\cdot)$ operations applied to elements of W.

Subspaces of V include the trivial examples $W = (0)$ and $W = V$; all others are "proper" subspaces of V.

2.2. Definition. *Given a non empty set S of vectors in V, its* **linear span** $\langle S \rangle =$

$\mathbb{K}$-span(S) *is the smallest subspace* $W \subseteq V$ *such that* W *contains* S.

It is easy to verify that:

2.3. Exercise. If $\{W_\alpha : \alpha \in I\}$ is any family of subspaces in V, prove that their intersection $W = \bigcap_{\alpha \in I} W_\alpha$ is also a subspace.

Thus Definition 2.2 makes sense: Given S there is at least one subspace containing S, namely V. If E = intersection of all subspace Wthat contain S, then E is a subspace and is obviously the smallest subspace containing S. Thus $\mathbb{K}$-span(S) exists, even if V is "infinite dimensional," for instance $V = \mathbb{K}[x]$.

This "top down" definition has its uses, but an equivalent "bottom-up" version is often more informative.

2.4. Lemma. If $S \neq \emptyset$ in V, its linear span $\mathbb{K}$-span(S) is the set of finite sums

$$\left\{ \sum_{i=1}^{n} a_i v_i : a_i \in \mathbb{K}, v_i \in S, n < \infty \right\}$$

Proof: Let $E = \{\sum_{i=1}^{N} c_i v_i : N < \infty,\ c_i \in \mathbb{K},\ v_i \in S\}$. Since $S \subseteq \mathbb{K}$-span(S), every finite sum lies this span, proving $E \subseteq \mathbb{K}$-span(S). For $(\supseteq)$, it is clear that the family E of finite linear combination is closed under $(+)$ and $(\cdot)$ operations because a linear combination of linear combinations is just one big linear combination of elements of S. It is a subspace of V, and contains S because $1 \cdot s = s$ is a (trivial) linear combination. On the other hand every subspace $W \supseteq S$ must contain all these linear sums, so $S \subseteq E \subseteq W$. Hences E is the smallest subspace containing S and $E = K$-span(S). $\square$

2.5. Exercise. If $K = \mathbb{R}$, $V = \mathbb{R}^3$ show that $W = \{x \in \mathbb{R}^3 : 3x_1 + 2x_2 - x_3 = 0\}$ is a subspace and $W' = \{x \in \mathbb{R}^3 : 3x_1 + 2x_2 - x_3 = 1\}$ is not a subspace.
Hint: For one thing the zero vector $0 = (0, 0, 0)$ is not in W'. The situation is shown in Figure 1.3.

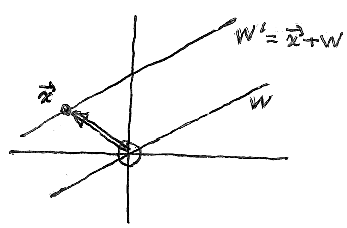

Figure 1.3. The subspace W in Exercise 2.5 and a translate $W' = \mathbf{x}_0 + W$ by some $\mathbf{x}_0 \in V$ such that $3x_1^0 + 2x_2^0 - x_3^0 = 1$, for instance $\mathbf{x}_0 = (0, 1, 1)$. The set W' is not a subspace.

System of Linear Equations. Systems of n linear equations in m unknowns are of two general types

Homogeneous

$$\begin{cases} a_{11}x_1 + \ldots + a_{1m}x_m = 0 \\ \vdots \\ a_{n1}x_1 + \ldots + a_{nm}x_m = 0 \end{cases}$$

Inhomogeneous

$$\begin{cases} a_{11}x_1 + ... + a_{1m}x_m = b_1 \\ \vdots \\ a_{n1}x_1 + ... + a_{nm}x_m = b_n \end{cases}$$

with a_{ij} and b_k in $\mathbb{K}$. $\square$

2.6. Exercise. Verify that the solutions $\mathbf{x} = (x_1, \ldots, x_m)$ of the homogeneous system form a vector subspace of $\mathbb{K}^m$. Explain why the solution set of an inhomogeneous system cannot be a vector subspace unless $\mathbf{b} = (b_1, \ldots, b_n) = \mathbf{0}$ in $\mathbb{K}^n$.

If we regard vectors $\mathbf{x} = (x_1, ..., x_m) \in \mathbb{K}^m$ as the entries in an $m \times 1$ column matrix,

$$\mathbf{x} = \mathrm{col}(x_1, \ldots, x_m) = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} ,$$

you will recognize that the solutions $\mathbf{x} \in \mathbb{K}^m$ of the homogeneous system of equations are precisely the solutions of the matrix equations

$$A\mathbf{x} = \mathbf{0} \qquad \text{where the zero vector is} \qquad \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}_{n\times 1}$$

and for inhomogeneous systems we must solve

$$A\mathbf{x} = B \qquad \text{where} \qquad B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}_{n\times 1}$$

for $B \in \mathbb{K}^n$.

The homogeneous system always has the zero vector $\mathbf{0} \in \mathbb{K}^m$ as a solution, and the solution set $\{\mathbf{x} \in \mathbb{K}^m : A\mathbf{x} = \mathbf{0}\}$ is a vector subspace in $\mathbb{K}^m$. If $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ then the number of solutions is either 1 or ∞ for this system. An inhomogeneous system might not have any solutions at all; otherwise, it has just one solution or infinitely many.

If A is an $n \times m$ matrix with entries in $\mathbb{K}$ we will find it useful to let A act by left multiplication as an operator $L_A : \mathbb{K}^m \rightarrow \mathbb{K}^n$ on column vectors

$$\mathbf{y} = L_A(\mathbf{x}) = A \cdot \mathbf{x} \qquad \big(\text{an } (n \times m)\cdot(m \times 1) \text{ matrix product}\big)$$

for $\mathbf{x} \in \mathbb{K}^m$. This is a *linear operator* in the sense that

$$L_A(\mathbf{x} + \mathbf{y}) = L_A(\mathbf{x}) + L_A(\mathbf{y}) \qquad \text{and} \qquad L_A(\lambda \cdot \mathbf{x}) = \lambda \cdot L_A(\mathbf{x})$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{K}^m$ and $\lambda \in \mathbb{K}$. Solving a system of linear equations is then equivalent to finding solutions of $L_A(\mathbf{x}) = 0$ or $L_A(\mathbf{x}) = B$ for $\mathbf{x} \in \mathbb{K}^m$. From this point of view, $A\mathbf{x} = B$ has solutions if and only if B lies in the range $R(L_A) = \{A\mathbf{x} : \mathbf{x} \in \mathbb{K}^m\}$ (a vector subspace in $\mathbb{K}^n$). If $B = \mathbf{0}$ the "homogeneous" equation $A\mathbf{x} = \mathbf{0}$ always has the trivial solution $\mathbf{x} = \mathbf{0}$. $\square$

2.7. Exercise. If A is an $n \times m$ matrix and $L_A ; \mathbb{K}^m \rightarrow \mathbb{K}^n$ is defined as above, verify that

1. The *range* $R(L_A) = L_A(\mathbb{K}^m) = \{A\cdot\mathbf{x} : x \in \mathbb{K}^m\}$ is a vector subspace in $\mathbb{K}^n$.

2. The *kernel* $K(L_A) = \ker(L_A) = \{x \in \mathbb{K}^m : L_A(\mathbf{x}) = A\cdot\mathbf{x} = \mathbf{0} \text{ in } \mathbb{K}^n\}$ is a vector subspace in $\mathbb{K}^m$.

2.8. Example. Given a particular solution $\mathbf{x}_0$ of $A\mathbf{x} = B$, the full solution set of this equation consists of the vectors $W_B = \mathbf{x}_0 + W$, where $W = \{x \in \mathbb{K}^m : A\mathbf{x} = 0\}$ is a vector subspace of $\mathbb{K}^m$ because $A\mathbf{x}_1$, $A\mathbf{x}_2 = 0$ implies $A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = 0 + 0 = 0$ and $A(\lambda \cdot \mathbf{x}) = \lambda \cdot A\mathbf{x} = \lambda \cdot 0 = 0$.
Note: The converse is also true: in $\mathbb{K}^m$ every vector subspace is the solution set of some homogeneous system of linear equation $A\mathbf{x} = 0$, but we are not ready to prove that yet. The situation is shown in Figure 1.4. $\square$

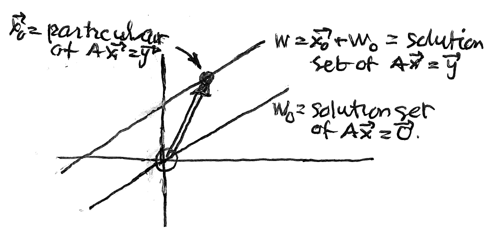

Figure 1.4. The subspace W_0 is the solution set for a homogeneous equation $A\mathbf{x} = \mathbf{0}$. If the inhomogeneous equation $A\mathbf{x} = \mathbf{y}$ has solutions and if $\mathbf{x}_0$ is a particular solution, so $A\mathbf{x}_0 = y$, the full solution set $W = \{\mathbf{x} : A\mathbf{x} = \mathbf{B}\}$ is the translate $W' = \mathbf{x}_0 + W$ of W_0.

This of course presumes that $A\mathbf{x} = B$ has any solutions at all; if it does not, we say that the system is *inconsistent*. Geometrically, that means B does not lie in the range $R(L_A)$. Here is an example of an inconsistent inhomogeneous system.

$$\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The corresponding system of linear equations

$$\begin{cases} x_1 + 0 \cdot x_2 &= 0 \\ 2x_1 + 0 \cdot x_2 &= 1 \end{cases}$$

implies that $x_1 = 0$ and $2x_1 = 1$, an obvious impossibility.

We will continue discussion of linear systems and their solutions via elementary row operations on A, or on the augmented matrix $[A : B]$, but first a few more examples of vector spaces we will encounter from time to time.

2.9. Example (Sequence Space ℓ^∞). Let ℓ^∞ = all sequences $a = (a_1, a_2, ...)$ with $a + b = (a_1 + b_1, a_2 + b_2, \dots)$ and $\lambda \cdot a = (\lambda a_1, \lambda a_2, ...)$. This infinite dimensional space has the following subspaces:

1. W_0 = {sequences such that $a_n \to 0$ as $n \to \infty$};
2. W_n = all sequences of the form $(a_1, ..., a_n, 0, 0, ...)$;
3. $\ell^1 = \{\, a : \sum_{n=1}^{\infty} |a_n| < \infty \,\}$

2.10. Example. In $M(n, \mathbb{K})$ we have various significant subspaces

1. Symmetric matrices: $A^{\mathrm{t}} = A$ where A^{t} = (transpose of A).

2. Diagonal matrices: $D = \begin{pmatrix} d_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & d_n \end{pmatrix}$

3. Block diagonal matrices: $D_{m_1,\ldots,m_r} = \begin{pmatrix} \boxed{B_{m_1\times m_1}} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \boxed{B_{m_r\times m_r}} \end{pmatrix}$

   for fixed indices $m_1, ..., m_r \geq 1$. (The "blocks" are allowed to have arbitrary entries and all other entries are zero; $m_1 + \ldots + m_r = n$.)

4. Upper triangular and Strictly upper triangular matrices.

$$\begin{pmatrix} * & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & * \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & & & * \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & 0 \end{pmatrix}$$

2.11. Exercise. Which of these four subspaces, if any, are closed under matrix multiplication as well as $(+)$?

2.12. Exercise. Show that the vector subspace of upper triangular and strictly upper triangular matrices are closed under formation of matrix product AB.

2.13. Exercise. Show that the vector subspaces of upper triangular (or strictly upper triangular) matrices are **Lie algebras**: all **commutators** $[A, B] = AB - BA$ are (strictly) upper triangular if A, B are.

2.14. Exercise. If an $n \times n$ matrix A has the strictly upper triangular form shown in (a), prove that A^2 has the form in (b).

$$\text{(a)}\quad A = \begin{pmatrix} 0 & * & & & * \\ & 0 & * & & * \\ & & \ddots & & \\ & & & 0 & * \\ 0 & & & & 0 \end{pmatrix} \qquad \text{(b)}\quad A^2 = \begin{pmatrix} 0 & 0 & * & & & & * \\ & 0 & 0 & * & & & * \\ & & 0 & 0 & * & & * \\ & & & \ddots & \ddots & & \\ & & & & \ddots & \ddots & * \\ & & & & & 0 & 0 \\ 0 & & & & & & 0 \end{pmatrix}$$

Note: Further computations show that A^3 has three diagonal files of zeros, etc so that A is a *nilpotent operator*, with $A^n = 0_{n\times n}$.

I.3. Determining Linear Span: A Case Study

Given vectors $\{v_1, ..., v_r\} \subseteq V$ and $b \in V$, the basic problem is to decide whether there exist $x_1, .., x_r \in \mathbb{K}$ such that $b = \sum_{i=1}^{r} x_i v_i$ (and if so, for which choices of coefficients $x_1, .., x_r$). Row operations on matrices are the main tool for resolving such questions.

3.1. Example. Consider the vectors in $\mathbb{K}^3$

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},\ \mathbf{u}_2 = \begin{pmatrix} -2 \\ -4 \\ -2 \end{pmatrix},\ \mathbf{u}_3 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix},\ \mathbf{u}_4 = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix},\ \mathbf{u}_5 = \begin{pmatrix} -3 \\ 8 \\ 16 \end{pmatrix}$$

and let A be the matrix with these vectors as its columns

$$A = \begin{pmatrix} 1 & -2 & 0 & 2 & -3 \\ 2 & -4 & 2 & 0 & 8 \\ 1 & -2 & 3 & -3 & 16 \end{pmatrix}$$

If $B = \text{col}(2, 6, 8) = \begin{pmatrix} 2 \\ 6 \\ 8 \end{pmatrix}$, determine all column vectors

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

such that $\sum_i x_i \mathbf{u}_i = \mathbf{0}$ or $\sum_i x_i \mathbf{u}_i = B$ in $\mathbb{K}^3$. (In the second case we are determining whether B lies in the linear span of $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$.) Then do this for an arbitrary column vector $B = \text{col}(b_1, b_2, b_3)$ to to get all solutions of $A\mathbf{x} = B$.

Discussion: A solution $\mathbf{x} = \text{col}(x_1, \ldots, x_5)$ of $A\mathbf{x} = B$ statisfies the matrix equation

$$\begin{aligned} B &= \sum_{i=1}^{5} x_i \mathbf{u}_i = x_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} -2 \\ -4 \\ -2 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} + x_4 \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix} + x_5 \begin{pmatrix} -3 \\ 8 \\ 16 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -2 & 0 & 2 & -3 \\ 2 & -4 & 2 & 0 & 8 \\ 1 & -2 & 3 & -3 & 16 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = A\mathbf{x} \,. \end{aligned}$$

We shall determine the full solution sets of the systems $A\mathbf{x} = 0$ or $A\mathbf{x} = B$ for the 3×5 matrix $A = [\mathbf{u}_1; \mathbf{u}_2; \mathbf{u}_3; \mathbf{u}_4; \mathbf{u}_5]$.

Before analyzing this problem we recall a few basic facts about solving matrix equations using elementary row operations. These methods are based on the following observations with which you should already be familiar: see the early chapters of *Schaum's Outline*. The simple (but important) verification is left as an exercise.

3.2. Proposition. *The following elementary row operations on a matrix A do not change the set of solutions $\mathbf{x}$ of $A\mathbf{x} = 0$.*

1. $R_i \leftrightarrow R_j$: *switch two rows*;
2. $R_i \rightarrow \lambda R_i$: *scale* (*row* i) *by some* $\lambda \neq 0$ *in* $\mathbb{K}$;
3. $R_i \rightarrow R_i + \lambda R_j$: *for* $i \neq j$ *add any scalar multiple of* (*row* j) *to* (*row* i), *leaving* (*row* j) *unaltered.*

Applied to the "augmented matrix" $[A : B]$ associated with an inhomogeneous system $A\mathbf{x} = B$, the system $A'\mathbf{x} = B'$ associated with the modified matrix $[A' : B']$ has the same solution set as $A\mathbf{x} = B$.

The reason is that each of the moves 1.-3. is reversible, with $R_i \rightarrow R_i - \lambda R_j$ the inverse of $R_i \rightarrow R_i + \lambda R_j$. Although row operations do not change the solution set they can greatly simplify the system of equations to be solved, leading to easy systematic solution of matrix equations. For instance, when $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ it is always possible to find a sequence of row operations that put A into upper triangular **echelon form:**

$$
\text{(1)} \qquad \textsc{Echelon Form:} \qquad A = \left(\begin{array}{cccccccccc}
\boxed{1} & * & . & . & . & & & & & * \\
 & & & \boxed{1} & * & . & . & . & . & * \\
0 & & & & & \boxed{1} & * & . & . & * \\
\hline
 & & & & & & & & & \\
 & & & & & \mathbf{0} & & & &
\end{array}\right)
$$

The same moves put the augmented matrix $[A : B]$ into similar form

$$
\text{(2)} \qquad [A' : B'] = \left(\begin{array}{cccccccccc|c}
\boxed{1} & * & . & . & . & & & & & * & b'_1 \\
 & & & \boxed{1} & * & . & . & . & & * & \vdots \\
0 & & & & & \boxed{1} & * & . & . & * & b'_r \\
\hline
 & & & & & & & & & & b'_{r+1} \\
 & & & & \mathbf{0} & & & & & & \vdots \\
 & & & & & & & & & & b'_m
\end{array}\right)
$$

Solutions of the systems $A'\mathbf{x} = 0$, $A'\mathbf{x} = B'$ are quickly found by "backsolving" (illustrated bellow). One could go further, forcing A into even simpler form by knocking out all terms $*$ above the "step corners." These additional operations would of course affect B' in the augmented matrix yielding the **reduced echelon form**.

$$
[A'' : B''] = \left(\begin{array}{cccccccccc|c}
\boxed{1} & * & . & 0 & * & 0 & & & & * & b''_1 \\
 & & & \boxed{1} & * & 0 & . & . & . & * & \\
0 & & & & & \boxed{1} & * & . & . & * & b''_r \\
\hline
 & & & & & & & & & & b''_{r+1} \\
 & & & & \mathbf{0} & & & & & & \vdots \\
 & & & & & & & & & & b''_m
\end{array}\right)
$$

The "step corners" appearing in these echelon displays are often referred to as "*pivots*," and the columns in which they occur are the "*pivot columns*."

Notice that $A\mathbf{x} = B$ has the same solutions as $A'\mathbf{x} = B'$ where $[A' : B']$ is the echelon form of $[A : B]$. Solutions exist if and only if we have $b'_{r+1} = ... = b'_n = 0$ (the terms in B' below the row containing the last "step corner") because the last equations in the new linear system $A'\mathbf{x} = B'$ read $0 = b'_{r+1}$, ..., $0 = b'_n$ (the variables x_1,..., x_m don't appear!) These are inconsistent unless $b'_{r+1} = ... = b'_n = 0$.

Columns $C_i(A)$ that *do not* pass though a step corner correspond to "free variables" x_i in the solutions of the equation $A'\mathbf{x} = 0$; they are also free variables in solutions of $A'\mathbf{x} = B'$ if the consistency conditions $b'_{r+1} = ... = b'_n = 0$ have been met (without which there are no solutions at all.) If $I = \{1 \leq i_1 < ... < i_r \leq m\}$ are the indices labeling the pivot columns, the remaining indices correspond to free variables x_k $(k \notin I)$ in the solution. Once the values of the free variables have been specified, backsolving yields the values of the remaining "dependent" variables x_k $(k \in I)$. We get a unique solution $A'\mathbf{x} = 0$ for *every* choice of the free variables $(k \notin I)$; different choices yield different solutions and all solutions are accounted for. By Proposition 3.2 these are also the solutions of the original equation $A\mathbf{x} = 0$.

Example 3.1 (Resumed). Returning to our discussion, we put the original system into echelon form by applying row operations to

$$\left(\begin{array}{ccccc|c} 1 & -2 & 0 & 2 & -3 & 0 \\ 2 & -4 & 2 & 0 & 8 & 0 \\ 1 & -2 & 3 & -3 & 16 & 0 \end{array}\right)$$

Applying, $R_2 \leftarrow R_2 - 2R_1$ and $R_3 \leftarrow R_3 - R_1$ this becomes

$$\left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 2 & -3 & 0 \\ 0 & 0 & 2 & -4 & 14 & 0 \\ 0 & 0 & 3 & -5 & 19 & 0 \end{array}\right)$$

Now apply $R_3 \leftarrow R_3 - \frac{3}{2}R_2,\ R_2 \leftarrow \frac{1}{2}R_2$, and then $R_3 \leftarrow R_3 - 3R_2$ to get

$$\left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 2 & -3 & 0 \\ 0 & 0 & \boxed{1} & -2 & 7 & 0 \\ 0 & 0 & 0 & \boxed{1} & -2 & 0 \end{array}\right)$$

This is the desired echelon form. Some additional work, needless for most purposes, would yield the reduced echelon form,

$$\to \left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 0 & * & 0 \\ 0 & 0 & \boxed{1} & 0 & * & 0 \\ 0 & 0 & 0 & \boxed{1} & -2 & 0 \end{array}\right)$$

Recursively backsolving the corresponding system of linear equations, we see that

1. x_2, x_5 are free variables;
2. $x_4 - 2x_5 = 0 \ \Rightarrow\ x_4 = 2x_5$;
3. $x_3 - 2x_4 + 7x_5 = 0 \ \Rightarrow\ x_3 = -7x_5 + 2(2x_5) = -3x_5$;
4. $x_1 - 2x_2 + 2x_4 - 3x_5 = 0 \ \Rightarrow\ x_1 = 2x_2 - 2(2x_5) + 3x_5 = 2x_2 - x_5$.

The solutions of $A'\mathbf{x} = 0$ (which are also the solutions of $A\mathbf{x} = 0$) form a vector subspace in $\mathbb{K}^5$, each of whose points is uniquely labeled (parametrized) by the choice of the free variables x_2, x_5. Setting $x_2 = s$, $x_5 = t$ $(s, t \in \mathbb{K})$ we find that the solution set $W = \{\mathbf{x} \in \mathbb{K}^5 : A\mathbf{x} = 0\} = \{\mathbf{x} \in \mathbb{K}^5 : A'\mathbf{x} = 0\}$ is equal to

$$W = \left\{ \begin{pmatrix} 2s - t \\ s \\ -3t \\ 2t \\ t \end{pmatrix} : s, t \in \mathbb{K} \right\} = \left\{ \begin{pmatrix} 2x_2 - x_5 \\ x_2 \\ -3x_5 \\ 2x_5 \\ x_5 \end{pmatrix} : x_2, x_5 \in \mathbb{K} \right\}$$

These homogeneous solutions can be rewritten in a more instructive form

$$\mathbf{x} = \begin{pmatrix} 2s - t \\ s \\ -3t \\ 2t \\ t \end{pmatrix} = s \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ -3 \\ 2 \\ 1 \end{pmatrix} = s\,\mathbf{w}_1 + t\,\mathbf{w}_2 \ ,$$

which shows that every solution of $A\mathbf{x} = 0$ is a linear combination of two basic solutions

$$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \mathbf{w}_2 = \begin{pmatrix} -1 \\ 0 \\ -3 \\ 2 \\ 1 \end{pmatrix}$$

This approach describes the solution set W of $A\mathbf{x} = 0$ as linear span $\mathbb{K}$-span$\{\mathbf{w}_1, \mathbf{w}_2\}$ of a set of generators $\{\mathbf{w}_1, \mathbf{w}_2\}$. We will later observe that these vectors are a “basis” for the solution set W.

Solving the Inhomogeneous Equation $A\mathbf{x} = B$. The same elementary row operations that put A into echelon form may be applied to the augmented matrix $[A : B]$. We already know what happens to A; applying the same moves to the column vector $B = \mathrm{col}(b_1, b_2, b_3)$ with undetermined coefficients, the operations $R_2 \leftarrow R_2 - 2R_1$ and $R_3 \leftarrow R_3 - R_1$ transform

$$B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \rightarrow \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - b_1 \end{pmatrix}$$

Then $R_3 \leftarrow R_3 - \frac{3}{2}R_2$; $R_2 \leftarrow \frac{1}{2}R_2$, and $R_3 \leftarrow R_3 - 3R_2$ yield

$$\rightarrow \begin{pmatrix} b_1 \\ \frac{1}{2}b_2 - b_1 \\ b_3 - b_1 - \frac{3}{2}(b_2 - 2b_1) \end{pmatrix} = \begin{pmatrix} b_1 \\ \frac{1}{2}b_2 - b_1 \\ b_3 - \frac{3}{2}b_2 + 2b_1 \end{pmatrix}$$

The augmented matrix becomes

$$[A : B] \rightarrow \left(\begin{array}{ccccc|c} \boxed{1} & -2 & 0 & 2 & -3 & b_1 \\ 0 & 0 & \boxed{1} & -2 & 7 & \frac{1}{2}b_2 - b_1 \\ 0 & 0 & 0 & \boxed{1} & -2 & b_3 - \frac{3}{2}b_2 + 2b_1 \end{array}\right)$$

Again x_2 and x_5 are free variables and the general solution $\mathbf{x} = \mathrm{col}(x_1, x_2, x_3, x_4, x_5)$ of $A\mathbf{x} = B$ can be found by backsolving. Since we have already found the general solutions of $A\mathbf{x} = 0$, all we need is one particular solution $\mathbf{x}_B$. The simplest way to find one is to set $x_2 = x_5 = 0$ and backsolve to get

$$\begin{aligned}
&x_2, x_5 = 0 \\
&x_4 = b_3 - \tfrac{3}{2}b_2 + 2b_1 \\
&x_3 - 2x_4 = \tfrac{1}{2}b_2 - b_1 \Rightarrow x_3 = 2(b_3 - \tfrac{3}{2}b_2 + 2b_1) + \tfrac{1}{2}b_2 - b_1 = 2b_3 - \tfrac{5}{2}b_2 + 3b_1 \\
&x_1 - 2 \cdot 0 + 0 + 2x_4 + 0 = b_1 \Rightarrow x_1 = b_1 - 2x_4 = -2(b_3 - \tfrac{3}{2}b_2 + 2b_1) + b_1 = -2b_3 + 3b_2 - 3b_1.
\end{aligned}$$

So $\mathbf{x}_B = \mathrm{col}\big(-2b_3 + 3b_2 - 3b_1, 0, 2b_3 - \frac{5}{2}b_2 + 3b_1, b_3 - \frac{3}{2}b_2 + 2b_1, 0\big)$ is a particular solution and the full solution set is

$$W_B = \{x : A\mathbf{x} = B\} = \begin{pmatrix} -2b_3 + 3b_2 - 3b_1 \\ 0 \\ 2b_3 - \frac{5}{2}b_2 + 3b_1 \\ b_3 - \frac{3}{2}b_2 + 2b_1 \\ 0 \end{pmatrix} + \mathbb{K}\mathbf{w}_1 + \mathbb{K}\mathbf{w}_2$$

where $\mathbf{w}_1$ and $\mathbf{w}_2$ are the basis vectors for the space $W = \{x : A\mathbf{x} = 0\}$ of homogeneous solutions determined previously. Writing $s = x_2$, $t = x_5$ for the variable attached to $\mathbf{w}_1$, $\mathbf{w}_2$ we obtain a parametric description of the solution set, with each point in W_B tagged by a unique pair (s, t) in the parameter space $\mathbb{K}^2$.

In the problem originally posed we had $B = \mathrm{col}(2, 6, 8)$. Then the particular solution is $\mathbf{x}_0 = \text{col}(-4, 0, 7, 3, 0)$ and the solution set is

$$W_B = \begin{pmatrix} -4 + 2s - t \\ s \\ 7 - 3t \\ 3 + 2t \\ t \end{pmatrix} = \mathbf{x}_0 + \mathbb{K}\mathbf{w}_1 + \mathbb{K}\mathbf{w}_2$$

That concludes our discussion of the Case Study 3.1. $\square$

Further Remarks about Elementary Row Operations. Row operations can also be used to determine the subspace spanned by any finite set of vectors in $\mathbb{K}^m$. If these have the form $R_1 = (a_{11}, .., a_{1m})$,..., $R_n = (a_{n1}, ..., a_{nm})$ we may regard them as the rows of an $n \times m$ matrix

$$A = \begin{pmatrix} R_1 \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline R_n \end{pmatrix} = \begin{pmatrix} a_{11} & . & . & . & a_{1m} \\ a_{21} & . & . & . & a_{2m} \\ . & & & & . \\ . & & & & . \\ a_{n1} & . & . & . & a_{nm} \end{pmatrix}$$

The linear span $\text{Row}(A) = \mathbb{K}\text{-span}\{R_1, ..., R_n\} \subseteq \mathbb{K}^m$ is called the **row space** of A; the linear span of its columns $C_1, \dots, C_m$ is the **column space** $\text{Col}(A) = \mathbb{K}\text{-span}\{C_1, .., C_m\}$ in $\mathbb{K}^n$. One can show that:

3.3. Lemma. *Elementary row operations on a matrix A do not change the linear span of its rows.*

We leave the proof as a routine exercise. Note, however, that row operations will mess up column space!

As for columns, there is an obvious family of elementary column operations on A.

1. $C_i \leftrightarrow C_i$;
2. $C_i \rightarrow \lambda C_i$ for $\lambda \neq 0$ in $\mathbb{K}$;
3. $C_i \rightarrow C_i + \lambda C_j$, for $i \neq j$ where λ is any element in $\mathbb{K}$.

These do not change the linear span $\text{Col}(A)$. This can be verified by direct calculation, but it also follows by observing that row and column operations are related via a natural symmetry $A \mapsto A^{\text{t}}$ = the *transpose of* A, given by $(A^{\text{t}})_{ij} = A_{ji}$ (see Figure 1.5). Note that $(A^{\text{t}})^{\text{t}} = A$.

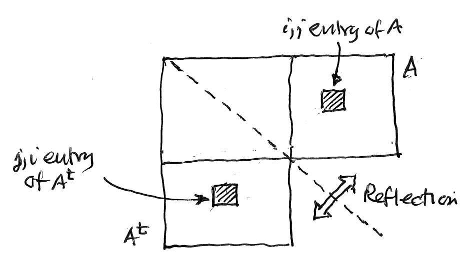

Figure 1.5. A matrix A and its transpose A^{t} are related by a reflection that sends rows in A to columns in A^{t}, and columns to rows.

The transpose operation takes rows of A to columns of A^{t} and vice-versa; elementary row operations on A become the corresponding elementary operations on the columns of A^{t}. It should also be evident that under the transpose operation the row space $\mathrm{Row}(A)$ = (span of the rows, regarded as vectors in $\mathbb{K}^m$) becomes the column space $\mathrm{Col}(A^{\mathrm{t}})$ = (columns in A^{t}, regarded as vectors in $\mathbb{K}^n$) of A^{t}. Invariance of $\mathrm{Col}(A)$ under column operations follows from invariance of $\mathrm{Row}(A^{\mathrm{t}})$ under row operations, discussed earlier.

3.4. Example. Let $v_1, .., v_n \in \mathbb{K}^m$. To find a basis for $W = \mathbb{K}\text{-span}\{v_1, .., v_n\}$, view the v_i as $1 \times m$ row vectors and assemble them as the $n \times m$ matrix

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_{n \times m}$$

If we perform row operations to put A in echelon form, this does not change row space $\mathrm{Row}(A) = \mathbb{K}\text{-span}\{v_1, .., v_n\}$, but it is now easy to pick out a minimal set of vectors with the same linear span, namely the rows R'_1, ..., R'_k that meet the step corners in the array.

$$A' = \begin{array}{c} R'_1 \\ R'_2 \\ \cdot \\ \cdot \\ \cdot \\ R'_k \\ \\ \end{array} \left(\begin{array}{cccccccccc} \boxed{1} & * & * & * & * & & & & & * \\ & & \boxed{1} & * & * & & & & & * \\ & & & & \boxed{1} & * & * & & & * \\ & & & & & \ddots & & & & \vdots \\ & & & & & & \boxed{1} & * & & * \\ 0 & & & & & & & & \boxed{1} & * \\ \hline 0 & & & & & & & & & 0 \end{array} \right)$$

We will say more about this in the next section. $\square$

3.5. Exercise. By invariance of row space $\mathrm{Row}(A)$ under row operations, the rows $R'_1, ..., R'_k$ also span $\mathrm{Row}(A)$. They are a *basis* for row space if they are also linearly independent in the following sense.

> LINEAR INDEPENDENCE: *If* $\sum_{i=1}^{k} c_i R'_i = \mathbf{0}$ *in* $\mathbb{K}^m$ *for coefficients* $c_1, .., c_k$ *in* $\mathbb{K}$*, we must have* $c_1 = ... = c_k = 0$ *in* $\mathbb{K}$.

Explain why the row vectors R'_1,..., R'_k in the previous example must have this independence property.
Hint: If $\sum_{i=1}^{n} c_i R'_i = (0, ..., 0)$ in $\mathbb{K}^m$, what conclusion can you draw about the first coefficient c_1? Etc.

A set of vectors $\{v_1, ..., v_n\} \subseteq V$ is a basis for V if they span V and are linearly independent. We will now show that this happens if and only if every $v \in V$ has a unique expansion $v = \sum_{i=1}^{n} c_i v_i$ with $c_i \in \mathbb{K}$. Independence simply says that the zero vector $v = 0$ in V has the unique expansion $0 = 0 \cdot v_1 + .. + 0 \cdot v_n$. But if some vector had two expansions $v = \sum_i c_i v_i = \sum_i d_i v_i$ then $0 = v - v = \sum (c_i - d_i) v_i$, so independence of the v_i implies $c_i = d_i$, and v has a unique expansion.

I.4. Linear Span, Independence and Bases

We now explain how to solve arbitrary systems of linear equations.

4.1. Definition. *A set of vectors* $S = \{v_1, .., v_r\}$ *in a vector space* V **spans** *a subspace* W *if*

$$W = \mathbb{K}\text{-span}\{S\} = \Big\{ \sum_{i=1}^{r} c_i v_i : c_i \in \mathbb{K} \Big\}$$

The vectors are **linearly independent** *if the only linear combination* $\sum_i c_i v_i = \mathbf{0}$ *adding up to zero in* V *is the trivial combination with* $c_1 = ... = c_r = 0$. *The vectors are a* **basis** *for* W *if they span* W *and are independent, so every* $w \in W$ *has a unique representation as* $\sum_{i=1}^n \lambda_i v_i$ $(\lambda_i \in \mathbb{K})$.

4.2. Exercise. If $\mathfrak{X} = \{v_1, \ldots, v_n\}$ span V and are independent, explain why every $v \in V$ has a unique representation as $\sum_{i=1}^n \lambda_i v_i$ $(\lambda_i \in K)$, so $\mathfrak{X}$ is a basis for V.

The next result exhibits two ways to construct a basis in a vector space. One starts with a spanning set and "prunes" it, deleting redundant vectors until we arrive at an independent subset with the same span as the original vectors. This yields a basis for V. The other constructs a basis recursively by adjoining "outside vectors" to an initial family of independent vectors in V. The initial family might consist of a single nonzero vector (obviously an independent set).

4.3. Proposition. *Every finite spanning set* $\{v_1, ..., v_n\}$ *in a vector space can be made into a basis by deleting suitably chosen entries from the list.*

Proof: We argue by induction on $n = \#(\text{vectors in list})$. There is nothing to prove if $n = 1$; then $V = \mathbb{K} \cdot v_1$ and $\{v_1\}$ is already a basis. The induction hypotheses (one for each index $n = 1, 2, ...$) are:

> HYPOTHESIS $P(n)$: *For any vector space* V *containing a spanning set of* n *vectors, we can delete vectors from the list to get a basis for* V.

We have proved this for $n = 1$. It is true for all n if we can prove $P(n+1)$ is true, using only the information that $P(n)$ is true – i.e. if we can verify that

$$P(n) \text{ true } \Rightarrow P(n+1) \text{ true}$$

(Remember: This is a *conditional* statement owing to the presence of the word "*If...*" It does not assert that $P(n)$ is actually true.)

So, assuming $P(n)$ true consider a spanning set $\mathfrak{X} = \{v_1, ..., v_n, v_{n+1}\}$ in V. If these vectors are already independent (which could be checked using row operations if $V = \mathbb{K}^m$), we already have a basis for V without deleting any vectors. If $\mathfrak{X}$ is not independent there must be coefficients $c_1, ..., c_{n+1} \in \mathbb{K}$ (not all equal to 0) such that $\sum_{i=1}^{n+1} c_i v_i = \mathbf{0}$. Relabeling, we may assume $c_{n+1} \neq 0$, and then ($\mathbb{K}$ being a field)

$$-c_{n+1} v_{n+1} = \sum_{i=1}^{n} c_i v_i \quad \text{ and } \quad v_{n+1} = \sum_{i=1}^{n} -\big(c_i / c_{n+1}\big) \cdot v_i$$

Thus $v_{n+1} \in \mathbb{K}\text{-span}\{v_1, ..., v_n\}$ and $\mathbb{K}\text{-span}\{v_1, ..., v_{n+1}\} = \mathbb{K}\text{-span}\{v_1, ..., v_n\}$ is all of V. By the induction hypotheses we may thin out $\{v_1, .., v_n\}$ to get a basis for V. $\square$

4.4. Proposition. *If* $\{v_1, ..., v_n\}$ *are independent in a vector space* V, *and* v_{n+1} *is a vector not in* $W_0 = \mathbb{K}\text{-span}\{v_1, ..., v_n\}$ *then*

1. $\{v_1, ..., v_n, v_{n+1}\}$ *are independent*;
2. $W_0 \subsetneq W_1 = \mathbb{K}\text{-span}\{v_1, ..., v_n, v_{n+1}\}$;
3. $\{v_1, ..., v_n, v_{n+1}\}$ *is a basis for* W_1.

Proof: If $v_1 \ldots, v_{n+1}$ are not independent there would be $c_i \in \mathbb{K}$ (not all zero) such that $\sum_{i=1}^{n+1} c_i v_i = \mathbf{0}$. We can't have $c_{n+1} = 0$, otherwise $\sum_{i=1}^{n} c_i v_i = \mathbf{0}$ contrary to assumed independence of $\{v_1, ..., v_n\}$. Thus $v_{n+1} = \sum_{i=1}^{n} -\big(c_i / c_{n+1}\big) \cdot v_i$ is in W_0, which contradicts the assumption $v_{n+1} \notin W_0$. Conclusion: $v_1, ..., v_{n+1}$ are independent. It follows immediately that $\{v_1, ..., v_{n+1}\}$ is a basis for $W_1 = \mathbb{K}\text{-span}\{v_1, ..., v_{n+1}\}$.

Note: This is an example of a "proof by contradiction," in which the assumption that $v_1, ..., v_n$ are *not* independent leads to an impossible conclusion. Therefore the statement "$v_1, ..., v_n$ are independent" must be true. $\square$

Important Remark: This process of "adjoining an outside vector" can be iterated to construct larger and larger independent sets and subspaces

$$\begin{aligned} W_0 = \mathbb{K}\text{-span}\{v_1, ..., v_n\} \ &\subsetneqq \ W_1 = \mathbb{K}\text{-span}\{v_1, ..., v_n, v_{n+1}\} \\ &\subsetneqq \ .... \subsetneqq W_r = \mathbb{K}\text{-span}\{v_1, ..., v_{n+r}\} \end{aligned}$$

Since $\{v_1, \ldots, v_n\}$ are independent they are a basis for the initial space W_0, and by Lemma 4.4 $v_1, \ldots, v_n, \ldots, v_{n+r}$ will be a basis for W_r. If this process stops in finitely many steps (because $W_r = V$ and we can no longer find a vector outside W_r), we have produced a basis for V. If the process never stops, no finite subset of vectors can span V and in this case we say V is **infinite dimensional**. To begin the process we need an initial set of independent vectors, but if $V \neq (0)$ we could start with any $v_1 \neq 0$ and $W_0 = \mathbb{K} \cdot v_1$. Then apply Lemma 4.4 recursively as above. $\square$

4.5. Definition. *A vector space V is* **finite dimensional** *if there is a finite set of vectors $S = \{v_1, ..., v_n\}$ that span V. Otherwise V is said to be* **infinite dimensional**, *which we indicate by writing* $\dim(V) = \infty$.

Coordinate space $\mathbb{K}^n$ and matrix spaces $\mathrm{M}(m \times n, \mathbb{K})$ are finite dimensional; the spaces of polynomials $\mathbb{K}[x]$ and $\mathbb{K}[x_1, \ldots, x_n]$ are infinite dimensional.

4.6. Example. Coordinate space $\mathbb{K}^n$ is finite dimensional and is spanned by the **standard basis vectors** $\mathfrak{X} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$

$$\mathbf{e}_1 = (1, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ \ldots \ , \mathbf{e}_n = (0, \ldots, 0, 1)$$

In fact $\mathfrak{X}$ is a basis for $\mathbb{K}^n$.

Discussion: Obviously $v = (a_1, ..., a_n) = a_1\mathbf{e}_1 + \ldots + a_n\mathbf{e}_n$ so the $\mathbf{e}_i$ span $\mathbb{K}^n$. But if $\sum_i c_i\mathbf{e}_i = \mathbf{0} = (0, ..., 0)$, that means $(c_1, ..., c_n) = (0, ..., 0)$ and $c_i = 0$ for all i. $\square$

4.7. Example. Polynomial space $\mathbb{K}[x]$ is infinite dimensional. Given any finite set of nonzero vectors $\mathfrak{X} = \{f_1, ..., f_r\}$, let $d_i = \deg(f_i)$. All coefficients of f_i are zero if $i > N = \max\{d_1, ..., d_r\}$, and the same is true for all linear combinations $\sum_{i=1}^{r} c_i f_i$. But then $\mathfrak{X}$ cannot span $\mathbb{K}[x]$ because x^{N+1} is not in $\mathbb{K}\text{-span}\{f_1, ..., f_r\}$.

Actually the vectors $f_0 = 1, f_1 = x, f_2 = x^2, ...$ *are* a basis for $\mathbb{K}[x]$. This (infinite) set of vectors clearly spans $\mathbb{K}[x]$, but it is also independent, for if $\sum_{i=0}^{r} c_i f_i = 0$ that means $c_0 + c_1 x + ... + c_r x^r = 0$ as a polynomial, so the symbol string $(c_0, ..., c_r, 0, 0, ...)$ is equal to $(0, 0, 0, ...)$. $\square$

4.8. Corollary. *Every finite dimensional vector space has a basis.*

Proof: If $\{v_1, ..., v_r\}$ span V, hen by Proposition 4.3 we may delete some of the vectors to get an independent set with the same linear span. $\square$

4.9. Lemma. *If $S \subseteq V$ is an independent set of vectors in V and T a finite set of vectors that span V, we can adjoin certain vectors from T to S to get a basis for V containing the original set of independent vectors S.*

Proof: Let $W = \mathbb{K}\text{-span}\{S\}$. If $W = V$, S is already a basis. If $W \neq V$, there exists some $v_1 \in T$ such that $v_1 \notin W$ and then $S \cup \{v_1\}$ is an independent set, a basis for the larger space $W_1 = \mathbb{K}\text{-span}\{S \cup \{v_1\}\} \supsetneqq W$. Continuing, we get vectors $v_1, ..., v_r$ in T such that $W \subsetneqq W_1 \subsetneqq W_2 \subsetneqq .... \subsetneqq W_r$ for $0 \leq i \leq r$, where $W_i = \mathbb{K}\text{-span}\{v_1, ..., v_i\}$. The process must terminate when no vector $v_{r+1} \in T$ can be found outside of W_r. Then

$T \subseteq W_r$, so $\mathbb{K}\text{-span}\{T\} = V \subseteq W_r$ and $W_r = V$. Therefore $S \cup \{v_1, ..., v_r\}$ is a basis for $V = W_r$ (and $S \cup \{v_1, ..., v_k\}$ is a basis for W_k for each $1 \leq k \leq r$). $\square$

4.10. Theorem (Dimension Defined). *All bases in a finite dimensional vector space have the same cardinality. More generally, if V is finite dimensional, and S is a finite spanning set (with $|S| = n$), every independent set of vectors $L \subseteq V$ has cardinality $|L| \leq |S|$. In other words, the size of any independent set is always less than or equal to that of any spanning set.*

Proof: We can eliminate vectors from S to get an independent spanning set $S' \subseteq S$, which is then a basis for V. We will show that $|L| \leq |S'| \leq |S|$. Let $S' = \{u_1, ..., u_n\}$ and $L = \{v_1, ..., v_m\}$. Every $v_i \in L$ can be written $v_i = \sum_{i=1}^{n} a_{ji} u_j$ since the $u_i \in S'$ are a basis for V. On the other hand, if $c_1, ..., c_m$ are scalars such that $0 = \sum_{j=1}^{m} c_j v_j$, we must have $c_1 = ... = c_m = 0$ because the v_j are independent. But the identity $\sum_{j=1}^{m} c_j v_j = 0$ can be written another way, as

$$0 = \sum_{i=1}^{m} c_i \Big(\sum_{j=1}^{n} a_{ji} u_j \Big) \; = \; \sum_{j=1}^{n} \Big(\sum_{i=1}^{m} a_{ji} c_i \Big) \, u_j$$

Since the $u_j \in S'$ span V and are independent each expression $(\ldots)$ is $= 0$ so the coefficients $c_1, ..., c_m$ satisfy the system of n equations in m unknowns

$$\sum_{i=1}^{m} a_{ji} c_i = 0, \qquad \text{for } 1 \leq j \leq n \tag{3}$$

(a solution $C = \text{col}\big(c_1, \ldots, c_m\big)$ of the matrix equation $AC = 0$).

A linear system such as (3) always has nontrivial solutions if the number of unknowns $m = |L|$ exceeds the number of equations $n = |S'|$; it follows that $|L| \leq |S'|$, as claimed. In fact, row operations on the coefficient matrix A yield an echelon form shown below. There are at most n step corners and if $M > n$ there must be at least one column that fails to meet one of these pivots.

$$\left(\begin{array}{cccccccccc}
\boxed{1} & * & . & . & . & & & & & * \\
 & & & . & . & . & . & . & . & \\
 & & & . & . & . & . & . & . & \\
 & & & \boxed{1} & * & . & . & . & . & * \\
0 & & & & & \boxed{1} & * & . & . & * \\
\hline
0 & . & & . & & . & & . & . & 0
\end{array} \right)_{n \times m}$$

Hence there is at least one free variable and the system $AC = 0$ has nontrivial solutions. But we showed above that $C = 0$ is the only solution, so we obtain a contradiction unless $|L| \leq |S'| \leq |S|$. The theorem is proved. $\square$

4.11. Corollary. *In a finite dimensional vector space all bases have the same cardinality, which we refer to hereafter as the* **dimension** $\dim_{\mathbb{K}}(V)$.

Notation: We will often simplify notation when the underlying ground field $\mathbb{K}$ is understood, by writing $\dim(V)$ or even $|V|$ for the dimension of V. $\square$

4.12. Example. We have already seen that $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$, with the standard basis vectors $\mathbf{e}_1 = (1, 0, ..., 0)$, ..., $\mathbf{e}_n = (0, .., 0, 1)$. We may view $\mathbb{C}^n$ (or any vector space over $\mathbb{C}$) as a vector space over $\mathbb{R}$ by restricting scalars in $\lambda \cdot v$ to be real. As a vector space over $\mathbb{C}$ we have $\dim_{\mathbb{C}}(V) = n$, but as a vector space over $\mathbb{R}$ we have $\dim_{\mathbb{R}}(V) = 2n$.

Discussion: In fact, any $v \in \mathbb{C}^n$ can be written as a complex sum $v = \sum_{j=1}^n z_j \mathbf{e}_j$, and if $z_j = x_j + iy_j$ we may write

$$v = x_1\mathbf{e}_1 + \ldots + x_n\mathbf{e}_n + y_1(i\mathbf{e}_1) + \ldots + y_n(i\mathbf{e}_n) \qquad \text{with } x_i, y_j \in \mathbb{R}.$$

Thus the vectors $\{\mathbf{e}_1, ..., \mathbf{e}_n, i\mathbf{e}_1, ..., i\mathbf{e}_n\} \subseteq \mathbb{C}^n$ span $\mathbb{C}^n$ as a vector space over $\mathbb{R}$. They are also independent over $\mathbb{R}$, for if

$$0 = \sum a_j\mathbf{e}_j + \sum b_j(i\mathbf{e}_j) = \sum (a_j + ib_j)\mathbf{e}_j \ ,$$

we must have $a_j + ib_j = 0$ and $a_j = b_j = 0$ because $\{\mathbf{e}_j\}$ is a basis over $\mathbb{C}$. $\square$

4.13. Exercise. If V is a finite dimensional vector space and $W \subseteq V$ a subspace, explain why W must also be finite dimensional.

4.14. Exercise. If V_1, V_2 are finite dimensional vector spaces prove that

1. If $V_1 \subseteq V_2$ then $\dim(V_1) \leq \dim(V_2)$;
2. If $\dim(V_1) = \dim(V_2)$ and $V_1 \subseteq V_2$, then $V_1 = V_2$ as sets.

4.15. Exercise. Explain why $W \subseteq V \Rightarrow \dim(W) \leq \dim(V)$, even if one or both of these spaces is infinite dimensional.

Describing Subspaces.

How can a subspace W in a vector space be specified? Every V of dimension n can be identified in a natural way with $\mathbb{K}^n$ once a basis $\{f_1, ..., f_n\}$ in V has been determined, so we may as well restrict attention to describing subspaces W of coordinate space $\mathbb{K}^n$. (Given a basis $\mathfrak{X} = \{f_i\}$ in V the map $j_{\mathfrak{X}} : \mathbb{K}^n \to V$ given by

$$\mathbf{x} = (x_1, ..., x_n) \mapsto j_{\mathfrak{X}}(\mathbf{x}) = \sum_{i=1}^n x_i f_i$$

is a bijection that respects all vector space operations in the sense that

$$j_{\mathfrak{X}}(\lambda \cdot \mathbf{x}) = \lambda \cdot j_{\mathfrak{X}}(\mathbf{x}) \qquad \text{and} \qquad j_{\mathfrak{X}}(\mathbf{x} + \mathbf{y}) = j_{\mathfrak{X}}(\mathbf{x}) + j_{\mathfrak{X}}(\mathbf{y})$$

It is an *isomorphism* between $\mathbb{K}^n$ and V, by which properties of one space can be matched with those of the other.

Subspaces $W \subseteq \mathbb{K}^n$ can be described in two ways.

1. By exhibiting a basis $\mathfrak{X} = \{f_1, ..., f_r\}$ in W, so $W = \mathbb{K}\text{-span}\{\mathfrak{X}\}$ and $\dim_{\mathbb{K}}(W) = r$. This is a "**parametric description**" of W since each $w \in W$ is labeled by a coordinate r-tuple $\mathbf{c} = (c_1, \ldots, c_r)$.

2. By finding a set of linear equations

$$\begin{aligned} a_{11}x_1 + ... + a_{1m}x_m &= 0 \\ \vdots \qquad & \quad \vdots \\ a_{n1}x_1 + ... + a_{nm}x_m &= 0 \end{aligned}$$

   described by a matrix equation $A\mathbf{x} = \mathbf{0}$ ($A = n \times m$, $\mathbf{0} = n \times 1$, $\mathbf{x} = m \times 1$) whose solution set $\{\mathbf{x} \in \mathbb{K}^m : A\mathbf{x} = \mathbf{0}\}$ is equal to W. Such an **"implicit description"** may include redundant equations. When there are no redundant equations we will see that $W = \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}$ has dimension $m - n = \dim(V) - \#(\textit{equations})$.

We illustrate this with some computational examples.

4.16. Example. Determine the dimension of the subspace $W = \mathbb{R}\text{-span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in $\mathbb{R}^3$ if

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad \mathbf{u}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} \qquad \mathbf{u}_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

Find a basis for W. Then describe W as the solution set of a system of linear equations:

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 & = & 0 \\ \vdots & & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 & = & 0 \end{array}$$

where $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$.

Solution: We write the vectors as the rows of the 3×3 matrix

$$A = \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$$

Row space $W = \text{Row}(A)$, the span of the rows, is unaffected by elementary row operations. These yield the echelon form

$$A \rightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & -2 & -4 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Therefore $\mathbf{w}_1 = (1, 2, 3)$ and $\mathbf{w}_2 = (0, 1, 2)$ span W; they are also independent because $0 = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 = (c_1\,,\,2c_1 + c_2\,,\,3c_1 + 2c_2)$ implies

$$\left\{ \begin{array}{c} c_1 = 0 \\ 2c_1 + c_2 = 0 \\ 3c_1 + 2c_2 = 0 \end{array} \right. \quad \Rightarrow \; c_1 = c_2 = c_3 = 0 \;\; .$$

Thus $\{\mathbf{w}_1, \mathbf{w}_2\}$ is a basis and $\dim(W) = 2$. A typical vector in W can be written (uniquely) as

$$s\,\mathbf{w}_1 + t\,\mathbf{w}_2 = (s\,,\,2s + t\,,\,3s + 2t) = (x_1, x_2, x_3) \qquad \text{with } s, t \in \mathbb{R}$$

To describe W as the solution set of a system of equations in x_1, x_2, x_3 we need to "eliminate" s, t from this parametric description of W. This can be done by writing

$$\left\{ \begin{array}{ll} x_1 = s & \Rightarrow s = x_1 \\ x_2 = 2s + t & \Rightarrow x_2 = 2s + t = 2x_1 + t \Rightarrow t = x_2 - 2x_1 \\ x_3 = 3s + 2t & \end{array} \right.$$

The last equation yields the "constraint" identity that determines W,

$$x_3 = 3s + 2t = 3x_1 + 2(x_2 - 2x_1) = -x_1 + x_2$$

or $x_1 - x_2 + x_3 = 0$ (1 equation in 3 unknows). Thus $W = \{\mathbf{x} \in \mathbb{R}^3 : x_1 - 2x_2 + x_3 = 0\}$, which has dimension $\dim(\mathbb{R}^3) - 1 = 2$. $\square$

4.17. Example. Let $W \subseteq \mathbb{R}^4$ be the solution set for the system of linear equations:

$$\left\{ \begin{array}{rcl} x_1 + x_2 - x_3 + 2x_4 & = & 0 \\ 3x_1 - x_2 + x_4 & = & 0 \end{array} \right.$$

so $A\mathbf{x} = \mathbf{0}$ ($\mathbf{x} \in \mathbb{R}^4$) where

$$A = \begin{pmatrix} 1 & 1 & -1 & 2 \\ 3 & -1 & 0 & 1 \end{pmatrix}_{2\times 4}$$

Find a basis for W and determine $\dim_{\mathbb{R}}(W)$. Do the answers change if we replace $\mathbb{R}$ by $\mathbb{Q}$ or $\mathbb{C}$?

Solution: Elementary row operations yield

$$A \rightarrow \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -4 & 3 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & 1 & -1 & 2 \\ 0 & \boxed{1} & -\frac{3}{4} & \frac{1}{2} \end{pmatrix}$$

and for any solution of $A\mathbf{x} = 0$, $\mathbf{x} = \mathrm{col}(x_1, x_2, x_3, x_4)$ has x_3, x_4 as free variables. Backsolving yields the dependent variables

$$\begin{aligned} x_2 &= \tfrac{3}{4}x_3 - \tfrac{1}{2}x_4 \\ x_1 &= -x_2 + x_3 - 2x_4 = (-\tfrac{3}{4}x_3 + \tfrac{1}{2}x_4) + x_3 - 2x_4 \;=\; \tfrac{1}{4}x_3 - \tfrac{3}{2}x_4 \end{aligned}$$

Thus solutions have the form

$$\mathbf{x} = \begin{pmatrix} \frac{1}{4}x_3 - \frac{3}{2}x_4 \\ \frac{3}{4}x_3 - \frac{1}{2}x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} \frac{1}{4} \\ \frac{3}{4} \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{3}{2} \\ -\frac{1}{2} \\ 0 \\ 1 \end{pmatrix} = x_3\mathbf{f}_1 + x_4\mathbf{f}_2$$

for every $x_3, x_4 \in \mathbb{K}$. The solution set is equal to the $\mathbb{R}$-span$\{(1,3,4,0)\,,\,(3,1,0,-2)\} =$ $\mathbb{R}$-span$\{\mathbf{f}_1, \mathbf{f}_2\}$. The vectors $\mathbf{f}_1$, $\mathbf{f}_2$ span the solution set W, but are also independent because

$$c_1(1,3,4,0) + c_2(3,1,0,-2) = (c_1 + 3c_2\,,\, 3c_1 + c_2\,,\, 4c_1\,,\, -2c_2) = (0,0,0,0)$$

implies that $c_1 = c_2 = 0$. Thus $\{\mathbf{f}_1, \mathbf{f}_2\}$ is a basis and $\dim_{\mathbb{R}}(W) = 2$. The result is the same if we replace the ground field $\mathbb{R}$ with $\mathbb{Q}$ or $\mathbb{C}$. $\square$

As a "rule of thumb," each constraint equation $a_{i1}x_1 + ... + a_{im}x_m = 0$ on $\mathbb{K}^m$ reduces the dimension of the solution set $W = \{\mathbf{x} \in \mathbb{K}^m : A\mathbf{x} = 0\}$ by 1, but this is not always the case.

4.18. Exercise. Consider the special case of one constraint equation

$$W = \{\mathbf{x} : \sum_{i=1}^{n} c_i x_i = 0\} \qquad \text{with } c_1, ..., c_n \in \mathbb{K}$$

1. Under what condition on $\{c_1, ..., c_n\}$ do we have $\dim_{\mathbb{K}}(W) = n - 1$?
2. Explain why $\dim(W) < n - 1$ is impossible.

4.19. Exercise. Same question but now with two constraint equations

$$\begin{cases} a_{11}x_1 + ... + a_{1m}x_m &= 0 \\ a_{21}x_1 + ... + a_{2m}x_m &= 0 \end{cases}$$

(or $A\mathbf{x} = \mathbf{0}$ with $A = 2 \times m$, $\mathbf{x} = m \times 1$, $\mathbf{0} = 2 \times 1$.) Now what condition on A make

1. $\dim_{\mathbb{K}}(W) = 0$
2. $\dim_{\mathbb{K}}(W) = 1$,

for the subspace $W = \{\mathbf{x} \in \mathbb{K} : A\mathbf{x} = \mathbf{0} \text{ in } \mathbb{K}^2\}$?

4.20. Example (Lagrange Interpolation Formula). For any infinite field such as $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, the problem of finding a polynomial $f \in \mathbb{K}[x]$ having specified values $f(p_j) = \lambda_j$ at a given set of distinct points $p_1, ..., p_n$ in $\mathbb{K}$ always has a solution. The solution is nonunique (the problem is underdetermined) unless we require that $\deg(f) = n - 1$; there may be no solution if $\deg(f) < n - 1$.

Discussion: The product $h(x) = \prod_{j=1}^{n}(x - p_j)$ has degree equal to n and is zero at each p_j (and zero nowhere else), so the solution to the interpolation problem cannot be unique without restrictions on $f(x)$: one can add h (or any scalar multiple thereof) to any proposed solution f. It is reasonable to ask for a solution $f(x)$ of minimal degree to reduce the ambiguity. The polynomial

$$f(x) = \sum_{i=1}^{n} \lambda_i \cdot \frac{\prod_{j\neq i}(x - p_j)}{\prod_{j\neq i}(p_i - p_j)} \tag{4}$$

has nonzero denominator, is equal to λ_i at p_i for each i, and has $\deg(f) = n - 1$.

This is the **Lagrange Interpolation Formula**, determined by direct methods. It is a bit complicated to rewrite this sum of products in the form $f = c_0 + c_1 x + ... + c_{n-1}x^{n-1}$. But the coefficients $c_0, \ldots, c_{n-1}$ can also be found directly as the solution of a system of linear equations

$$\lambda_j = f(p_j) = \sum_{k=0}^{n-1} p_j^k c_k \qquad \text{for } 1 \leq j \leq n - 1 \ ,$$

which is equivalent to the matrix equation $A\mathbf{c} = \lambda$ in which

$$A = \begin{pmatrix} p_1^0 & . & . & . & p_1^{n-1} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ p_n^0 & . & . & . & p_n^{n-1} \end{pmatrix}_{n\times n} \quad \text{and} \quad \mathbf{c} = \begin{pmatrix} c_0 \\ . \\ . \\ . \\ c_{n-1} \end{pmatrix}_{n\times 1} \quad \lambda = \begin{pmatrix} \lambda_0 \\ . \\ . \\ . \\ \lambda_{n-1} \end{pmatrix}_{n\times 1}$$

I.5 Quotient Spaces V/W.

If V is a vector space and W a subspace, the **additive cosets** of W are the translates of W by various vectors in V. They are the subsets $x + W = \{x + w : w \in W\}$ for some $x \in V$, which we shall often denote by $[x]$ when the subspace W is understood. In particular, W itself is the "**zero coset**": $[0] = 0 + W = W$. The key observation is that the whole space V gets partitioned into disjoint cosets that fill V. The collection of all cosets $[x]$ is the **quotient space** V/W. Observe that *points* in the space V/W are at the same time *subsets* in V.

5.1. Lemma. *If W is a subspace in V and $x, y \in V$,*

1. *Two cosets $x + W$ and $y + W$ either coincide or are disjoint, hence the distinct cosets of W partition the space V.*
2. *An additive coset can have various* **representatives** *$x \in V$. We have $y + W = x + W \Leftrightarrow$ there is some $w \in W$ such that $y = x + w$ (or $y - x \in W$).*
3. *If $y \in x + W$ then $y + W = x + W$.*

Proof: We start with an observation about sums $A + B = \{a + b : a \in A, b \in B\}$ of sets $A, B \subseteq V$ that will be invoked repeatedly in what follows.

5.2. Exercise. If W is a subspace of a vector space V and $w \in W$, prove that

1. $w + W = W$, for all $w \in W$;
2. $W + W = \{w_1 + w_2 : w_1, w_2 \in W\}$ is equal to W;
3. $W - W = W$.

Resuming the proof of Lemma 5.1, if cosets $x + W$ and $y + W$ have a point p in common there are $w_1, w_2 \in W$ such that $x + w_1 = p = y + w_2$, hence $y = x + (w_1 - w_2)$. By Exercise 5.2 the cosets are equal:

$$y + W = (x + (w_1 - w_2)) + W = x + \big((w_1 - w_2) + W\big) = x + W$$

For (2.), $x + W = y + W \Rightarrow y = y + 0 = x + w$ for some $w \in W$. Conversely, if $y = x + w$ for $w \in W$, then $y + W = x + (w + W) = x + W$ again by the Exercise. For (3.), it follows from (1.) that $y \in x + W \Rightarrow (y + W) \cap (x + W) \neq \emptyset \Rightarrow y + W = x + W$.

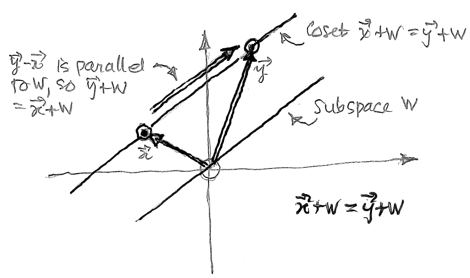

Figure 1.6. Additive cosets $\mathbf{x}+W$ of a subspace W are a family of parallel "hyperplanes" in a vector space V. When $V = \mathbb{R}^2$ and W a line through the origin, *all* lines parallel to W are cosets. Two vectors $\mathbf{x}, \mathbf{y}$ in the same coset yield the same translate of W: $\mathbf{x} + W = \mathbf{y} + W$ because $\mathbf{y} - \mathbf{x}$ is parallel to the subspace W.

As an example, if $V = \mathbb{R}^2$ and $W = \{(x, y) : x = y\}$ the cosets of W are precisely the distinct lines in the plane that make an angle of 45° with the positive x-axis. These lines are the "points" in the quotient space V/W, see Figure 1.6.

5.3. Definition. *There is a natural surjective* **quotient map** $\pi : V \to V/W$*, such that*

$$\pi(x) = [x] = x + W \tag{5}$$

If C is a coset, any point $v \in C$ such that $C = [v] = v + W$ is called a **representative of the coset**. *Part* (2.) *of Lemma 5.1 tells us when two vectors x, y represent the same coset.*

Algebraic Structure in V/W.

There are natural sum and scalar multiplication operations in V/W, inherited from the overlying vector space V.

5.4. Definition. *For any $x, y \in V$ and $\lambda \in \mathbb{K}$ we define operations in V/W*

1. Addition: $[x] \oplus [y] = [x + y]$;
2. Scalar Multiplication: $\lambda \odot [x] = [\lambda \cdot x]$

To spell out what is involved, this definition tells us how to form the sum $X \oplus Y$ of two cosets $X, Y \in V/W$ via the following algorithm:

1. Pick representatives $x, y \in V$ such that $X = [x], Y = [y]$.

2. Add the representatives to get $x+y \in V$.

3. Form the coset $[x+y] = (x+y)+W$ and report the output: $X \oplus Y = [x+y]$

But why should this make sense? The outcome depends on a choice of representatives for each coset X, Y and if different choices yield different outputs, everything written above is nonsense. Fortuately the outcome is independent of the choice of representatives and the operation $(\oplus)$ is *well-defined.* In fact, if $[x] = [x']$ and $[y] = [y']$ there must exist $w_1, w_2 \in W$ such that $x' = x + w_1,\ y' = y + w_2$, and

$$[x' + y'] = (x' + y') + W = (x+y) + \big((w_1 + w_2) + W\big) = (x+y) + W = [x+y]$$

Similarly, the scaling operation is well-defined: if $[x'] = [x]$ we have $x' = x + w$ for some $w \in W$, and then

$$[\lambda \cdot x'] = (\lambda \cdot x') + W = (\lambda \cdot x) + (\lambda w + W) = (\lambda \cdot x) + W = [\lambda \cdot x]$$

Once we know the operations $(\oplus)$ and $(\odot)$ make sense, direct calculations involving representatives show that all vector space axioms are satisfied by the system $(V/W, \oplus, \odot)$. For instance,

1. Associativity of $\oplus$ on V/W follows from associativity of $(+)$ on V: since $x+(y+z) = (x+y)+z$ in V we get

$$\begin{aligned} [x] \oplus \big([y] \oplus [z]\big) &= [x] \oplus [y+z] \ = \ [x + (y+z)] \\ &= [(x+y)+z] \ = \ [x+y] \oplus [z] \ = \ \big([x] \oplus [y]\big) \oplus [z] \end{aligned}$$

2. The zero element is $[0] = 0 + W = W$ because $[0] \oplus [x] = [0 + x] = [x]$

3. The additive inverse $-[x]$ of $[x] = x + W$ is $[-x] = (-x) + W$ since $[x] \oplus [-x] = [x + (-x)] = [0]$.

5.5. Exercise. Verify the remaining vector space axioms for $(V/W, \oplus, \odot)$. Then show that the quotient map $\pi : V \to V/W$ with $\pi(x) = [x] = x + W$ "*intertwines*" the algebraic operations in $(V, +, \cdot\,)$ with those in $(V/W, \oplus, \odot)$ in the sense that: for any $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$ we have

1. $\pi(v_1 + v_2) = \pi(v_1) \oplus \pi(v_2)$

2. $\pi(\lambda \cdot v_1) = \lambda \odot \pi(v_1)$

Thus $\pi : V \to V/W$ is a *linear operator* between these vector spaces. $\square$

When $W = (0)$ the quotient space consists of single points $[v] = v + W = \{v\}$, and V/W has a natural identification with V under the quotient map which is now a bijection. When $W = V$, there is just one coset, $v + W = v + V = V$; the quotient space reduces to a single point, the zero element $[0] = 0 + V = V$.

5.6. Exercise. Let $V = \mathbb{R}^3$ and $W = \{(x_1, x_2, x_3) : x_3 = 0\}$ = the x, y-plane in 3-dimensional space. The cosets in V/W are the distinct planes parallel to the x, y-plane: if $v = (v_1, v_2, v_3)$ then

$$\begin{aligned} v + W &= \{v + w : w \in W\} \\ &= \{(v_1, v_2, v_3) + (w_1, w_2, 0) : w_1, w_2 \in \mathbb{R}\} \\ &= \{(v_1 + s, v_2 + t, v_3) : s, t \in \mathbb{R}\} \\ &= \{(x_1, x_2, x_3) : x_1, x_2 \in \mathbb{R}, x_3 = v_3\} \end{aligned}$$

(the plane parallel to W passing through $(0, 0, v_3)$). Each value of $v_3 \in \mathbb{R}$ gives a different coset.

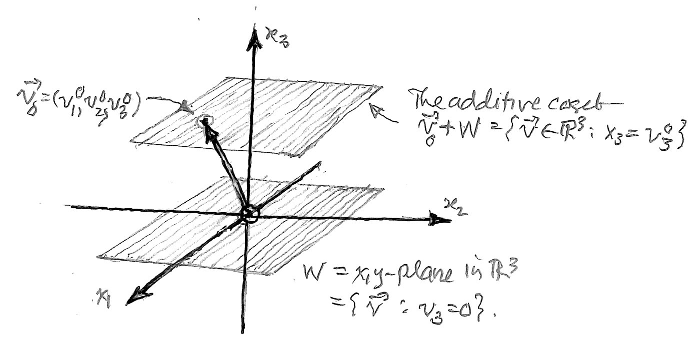

Figure 1.7. Additive cosets of $W = \{\mathbf{v} \in \mathbb{R}^3 : v_3 = 0\}$ are planes parallel to W in $\mathbb{R}^3$. A typical coset $\mathbf{v}_0 + W$ is shown.

One important viewpoint is to think of the quotient map $\pi : v \to V/W$ as "erasing" inessential aspects of the original vector space, retaining only those relevant to the problem at hand. Whole "bunches" of vectors in V, the cosets $v + W$, collapse to single points in the target space V/W (the planes in the last example become points in V/W). A lot of detail is lost in this collapse, but if W is suitably chosen the quotient map space will retain information that is buried in a lot of superfluous detail when we look at what is happening in the larger space V. We will soon give many examples of this, once we start looking at the structure of "linear operators" between vector spaces. For the moment we assemble a few more basic facts about quotients of vector spaces.

5.7. Theorem (Dimension Theorem for Quotients). *If V is finite dimensional and W is a subspace. Then:*

1. $\dim(V/W) \le \dim(V) < \infty$;
2. $\dim(W) \le \dim(V) < \infty$;

and

$$\dim(V) = \dim(W) + \dim(V/W) \tag{6}$$

By our notational conventions this identity can also be written in the abbreviated form $|V| = |W| + |V/W|$.

Proof: The quotient map $\pi : V \to V/W$ preserves linear combinations in the sense that

$$\pi\Big(\sum_{i=1}^{m} \lambda_i v_i\Big) = \sum_{i=1}^{m} \lambda_i \pi(v_i).$$

(recall Exercise 5.5), so if vectors $\{v_i\}$ span V their images $\overline{v}_i = \pi(v_i)$ span V/W. That proves

$$\dim(V/W) \le \#\{\overline{v}_i\} \le \#\{v_i\} = \dim(V) < \infty$$

as claimed in (1.).

As for item (2.), we know $\dim(V) < \infty$ but have no *a priori* information about W, but we showed earlier that no independent set in V can have more than $\dim(V)$ elements, and a basis for W would be such a set.

The identity (6) is proved by constructing a basis in V/W aligned with a specially chosen basis in V. Since $\dim(W) < \infty$ there is a basis $\{w_1, ..., w_m\}$ in W. If $W = V$ then

V/W is trivial and there is nothing more to do, but otherwise we can find an "outside vector" $v_{m+1} \notin W$ such that the larger set $\{w_1, ..., w_m, v_{m+1}\}$ is independent, and hence a basis for

$$W_1 = \mathbb{K}\text{-span}\{w_1, ..., w_m, v_{m+1}\} \supsetneq W_0 = W.$$

If $W_1 \neq V$, we can adjoin one more vector $v_{m+2} \notin W_1$ to get an independent set $\{w_1, ..., w_m, v_{m+1}, v_{m+2}\}$ with

$$W_0 \subsetneq W_1 \subsetneq W_2 = \mathbb{K}\text{-span}\{w_1, ..., w_m, v_{m+1}, v_{m+2}\}$$

This process must terminate, otherwise we would have arbitrary large independent sets in the finite dimensional space V. When the construction terminates we get an independent spanning set $\{w_1, ..., w_m, v_{m+1}, ..., v_{m+k}\}$ in $W_k = V$. This is a basis for V so $\dim(V) = m + k = \dim(W) + k$.

To conclude the proof we demonstrate that the $k = \dim(V/W)$ by showing that the π-images $\bar{v}_{m+1}, \ldots, \bar{v}_{m+k} \in V/W$ of the "outside vectors" are a basis for V/W. Since π is surjective the images $\pi(w_1), \ldots, \pi(v_{m+k})$ span V/W. But π "kills" all vectors in W, so

$$\pi(w_1) = \ldots = \pi(w_m) = [0] \quad \text{ in } V/W \ ,$$

and the remaining images $\bar{v}_{k+i} = \pi(v_{m+i})$ span V/W. They are also linearly independent. In fact, if some linear combination $\sum_{i=1}^{k} c_{m+i}\bar{v}_{m+i} = [0]$ in V/W, then by linearity of the quotient map π we get

$$[0] = \sum_{j=1}^{k} c_{m+j}\pi(v_{m+j}) = \pi\Big(\sum_{j=1}^{k} c_{m+j}v_{m+j}\Big)$$

But $\pi(v) = [0]$ for a vector $v \in V \Leftrightarrow [v] = v + W$ is equal to the zero coset $[0] = W$. Furthermore $v + W = W \Leftrightarrow v \in W$, so we can find coefficients $c_1, \ldots, c_m$ such that

$$\sum_{i=1}^{m} c_i w_i \ = \ v \ = \ \sum_{j=1}^{k} c_{m+j}v_{m+j}$$

or

$$0 = \sum_{i=1}^{m} c_i w_i + \sum_{j=1}^{k} (-1)c_{m+j}v_{m+j} \qquad \text{in } V.$$

Since $w_1, \ldots, w_m, v_{m+1}, \ldots v_{m+k}$ is a basis for V this can only happen if all coefficients in this sum are zero, and in particular $c_{m+1}, \ldots, c_{m+k} = 0$. Thus the $\{\bar{v}_i\}$ are independent and a basis for V/W, and $\dim(V/W) = k = \dim(V) - \dim(W)$. $\square$

Remark: The construction developed in proving Theorem 5.7 shows how to find bases in a quotient space V/W, and perform effective calculations with them. The key was to find representatives v_i back in V so we can transfer calculations involving cosets in V/W to calculations in V involving actual vectors v_i. The proof of Theorem 5.7 describes an explicit procedure for finding independent vectors $\{v_i\}$ outside of W, whose images $\pi(v_i) = \overline{v}_i$ are the desired basis in the quotient space.

5.8. Exercise. Find explicit bases for the following quotient spaces

1. $V = \mathbb{R}^3$, $W = \mathbb{R}\mathbf{e}_1 + \mathbb{R}\mathbf{e}_2$.
2. $V = \mathbb{R}^3$, $W = \mathbb{R}\text{-span}\{\mathbf{w}_1 = (1, 2, 3), \mathbf{w}_2 = (0, 1, -1)\}$;
3. $V = \mathbb{C}^4$, $W = \mathbb{C}\text{-span}\{\mathbf{z}_1 = (1\,,\,1 + i\,,\,3 - 2i\,,\,-i),\ \mathbf{z}_2 = (4 - i\,,\,0\,,\,-1\,,\,1 + i)\}$;

4. $V = \mathbb{R}^4$, $W = \{\mathbf{x} : x_1 + x_2 - x_3 + x_4 = 0 \text{ and } 4x_1 - 3x_2 + 2x_3 + x4 = 0\}$.

Here is a simple example involving bases in a quotient space V/W.

5.9. Example. Let $V = \mathbb{R}^4$ and $W = \{x \in \mathbb{R} : 2x_1 - x_2 + x_4 = 0\}$. The subspace W is the solution set of the matrix equation

$$A\mathbf{x} = \mathbf{0} \qquad \text{where} \qquad A = [\,2, -1, 0, 1\,]_{1\times 4}$$

that imposes a single linear constraint on $\mathbb{R}^4$. Find a basis for V/W

Solution: Row operations yield

$$A \rightarrow A' = \left[\ \boxed{1}\ , -\tfrac{1}{2}\ , 0\ , \tfrac{1}{2}\ \right]$$

The free variable are x_2, x_3, x_4 and $x_1 = \frac{1}{2}x_2 - \frac{1}{2}x_4$, so the solutions have the form

$$\mathbf{x} = \begin{pmatrix} \frac{1}{2}x_2 - \frac{1}{2}x_4 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

for $x_2, x_3, x_4 \in \mathbb{R}$. Thus the solution set for $A\mathbf{x} = \mathbf{0}$ is the linear span of the column vectors

$$\mathbf{u}_1 = \mathrm{col}(1, 2, 0, 0) \qquad \mathbf{u}_2 = \mathrm{col}(0, 0, 1, 0) \qquad \mathbf{u}_3 = \mathrm{col}(-1, 0, 0, 2)$$

These are a basis for W since they are easily seen to be linearly independent. Just row reduce the 3×4 matrix M that has these vectors as its rows

$$M = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 2 \end{pmatrix}$$

and see if you get a row of zeros; you do not. Therefore $\dim(\mathrm{Row}(M)) = 3$ and the vectors are independent.

Since $\dim(V) = \dim(W) + \dim(V/W)$ and $\dim(W) = 3$, we need only find one "outside" vector $\mathbf{u}_4 \notin W$ to complete a basis for $V = \mathbb{R}^4$; then $\pi(\mathbf{u}_4) = \mathbf{u}_4 + W$ will be nonzero, and a basis vector for the 1-dimensional quotient space. The vector $\mathbf{u}_4 = \mathbf{e}_4 = (0, 0, 0, 1)$ is not in W because it fails to satisfy the constraint equation $2x_1 - x_2 + x_4 = 0$. Thus the single vector $[\mathbf{e}_4] = \pi(\mathbf{e}_4) = \mathbf{e}_4 + W$ is a basis for V/W, and $\dim(V/W) = 1$. $\square$

5.10. Exercise (Another Dimension Formula). If E, F are subspaces in a finite-dimensional vector space V and $E + F = \{e + f : e \in E, f \in F\}$ is their linear span, prove that

$$\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F)$$

Hint: Choose appropriate bases related to E, F and $E \cap F$.

Appendix A: The Degree Formula for $\mathbb{K}[x_1, \ldots x_N]$.

Let $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \ldots, x_N]$ be the unital ring of polynomials with coefficients in an integral domain. Using the multi-index notation introduced in Section 10.1 we can write any such polynomial as a finite sum (finitely many nonzero coefficients)

$$f(\mathbf{x}) = \sum_{\alpha \in \mathbb{Z}_+^N} a_\alpha \, x^\alpha \qquad (x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_N^{\alpha_N}, \ c_\alpha \in R) \tag{7}$$

The degree of a monomial x^α is $|\alpha| = \alpha_1 + \ldots + \alpha_N$ and if $f \in \mathbb{K}[\mathbf{x}]$ is not the zero polynomial (all $a_\alpha = 0$) its degree is

$$m = \deg(f) = \max\{\, |\alpha| : c_\alpha \neq 0\}$$

When $N > 1$ there may be several different monomials x^α of the same total degree $|\alpha| = m$ with nonzero coefficients.

Let $f, g \neq 0$ in $\mathbb{K}[\mathbf{x}]$ with degrees $m = \deg(f), n = \deg(g)$. Their product is

$$\begin{aligned} (f\cdot g)(\mathbf{x}) &= \Big(\sum_\alpha a_\alpha \, x^\alpha\Big) \cdot \Big(\sum_\beta b_\beta \, x^\beta\Big) = \sum_{\alpha,\beta} a_\alpha b_\beta \, x^{\alpha+\beta} \\ &= \sum_\gamma \Big(\sum_{\alpha+\beta=\gamma} a_\alpha b_\beta\Big)\, x^\gamma = \sum_\gamma c_\gamma \, x^\gamma \end{aligned} \tag{8}$$

where $\alpha + \beta = (\alpha_1 + \beta_1, \ldots, \alpha_N + \beta_n)$. If $a_\alpha b_\beta x^{\alpha+\beta} \neq 0$ in (36) we must have $|\alpha| \leq m$ and $|\beta| \leq n$, so that $|\alpha + \beta| \leq m + n$; consequently $\deg(f \cdot g) \leq \deg(f) + \deg(g)$.

Let us split off the monomials of maximum degree, writing

$$\begin{aligned} f(\mathbf{x}) &= \sum_{|\alpha|=m} a_\alpha \, x^\alpha + (\cdots) \\ g(\mathbf{x}) &= \sum_{|\beta|=n} b_\beta \, x^\beta + (\cdots) \\ (f\cdot g)(\mathbf{x}) &= \sum_{|\gamma|=m+n} c_\gamma \, x^\gamma + (\cdots) \end{aligned}$$

where $(\cdots)$ are terms of lower degree. To prove the degree formula

(9) Degree Formula: $\quad \deg(f\cdot g) = \deg(f) + \deg(g) \qquad$ for $f, g \neq 0$ in $\mathbb{K}[\mathbf{x}]$

it suffice to show there is at least one monomial x^{γ_0} of maximal degree $m + n$ such that the coefficient

$$c_{\gamma_0} = \sum_{\alpha+\beta=\gamma_0} a_\alpha b_\beta \qquad \text{is nonzero.} \tag{10}$$

This is trivial for $N = 1$, but problematic when $N \geq 2$ because this sum of products can be zero if there is more than one term, even if the individual terms are nonzero. On the other hand the degree formula (37) follows immediately if we can prove

(11) *There exists some monomial x^γ of maximal degree $m + n$ for which the sum* (38) *consists of a single nonzero term.*

The key to proving (39) is to introduce a ranking of the monomials x^γ, $\gamma \in \mathbb{Z}_+^N$, more refined than ranking by total degree $\deg(x^\gamma) = |\gamma|$, which cannot distinguish between the various monomials of the same degree. The tool for doing this is "**lexicographic**," or "**lexical**," ordering of the indices in $\mathbb{Z}_+^N$, an idea that has proved useful in many parts of mathematics.

A.1. Definition (Lexicographic Order). *For $\alpha, \beta \in \mathbb{Z}_+^N$ we define the relation $\alpha \succ \beta$ to mean*

$$\alpha_i > \beta_i \text{ at the first index } i = 1, 2, \ldots, N \text{ at which } \alpha_i \text{ differs from } \beta_i$$

Thus

$$\alpha_1 = \beta_1, \ldots, \alpha_{i-1} = \beta_{i-1} \text{ and } \alpha_i > \beta_i \qquad (\text{other entries in } \alpha, \beta \text{ are irrelevant})$$

This is a **linear ordering** *of multi-indices: given α, β exactly one of the possibilities*

$$\alpha \succ \beta \qquad \alpha = \beta \qquad \beta \succ \alpha$$

holds. We write $\alpha \succeq \beta$ when the possibility $\alpha = \beta$ is allowed. □

Obviously $\alpha = (0, \ldots, 0)$ is the lowest multi-index in lexicographic order, and any finite set of multi-indices has a unique highest element. Note carefully that $\alpha \succ \beta$ does not imply that $|\alpha| \geq |\beta|$. For instance we have

$$\alpha = (1, 0, 0) \succ \beta = (0, 2, 2) \text{ in lexicographic order, but } |\beta| = 4 > |\alpha| = 1\,.$$

Other elementary properties of lexicographic order are easily verified once you understand the definitions.

A.2. Exercise. For lexicographic order in $\mathbb{Z}_+^N$ verify that

1. Linear Ordering. For any pair α, β we have exactly one of the possibilities $\alpha \succ \beta$, $\alpha = \beta$, $\beta \succ \alpha$.
2. Transitivity of Order. If $\alpha \succ \beta$ and $\beta \succ \gamma$ then $\alpha \succ \gamma$.
3. If $\alpha \succ \alpha'$ then $\alpha + \beta \succ \alpha' + \beta$ for all indices β.
4. If $\alpha \succ \alpha'$ and $\beta \succ \beta'$ then $\alpha + \beta \succ \alpha' + \beta'$.

Hint: It might help to make diagrams showing how the various N-tuples are related. You will have to do some "casework" in (3.) □

We now outline how the crucial fact (39) might be proved, leaving the final details as an exercise for the reader. If $f \neq 0$ with $m = \deg(f)$, so $f = \sum_{|\alpha| \leq m} a_\alpha\, x^\alpha$, there may be several monomials having maximal degree m with $a_\alpha \neq 0$, but just one of these is maximal with respect to lexicographic order, namely

$$\alpha_0 = \max_{\succ}\{\alpha : |\alpha| = m \text{ and } a_\alpha \neq 0\}$$

Likewise there is a unique index

$$\beta_0 = \max_{\succ}\{\beta : |\beta| = n \text{ and } b_\beta \neq 0\}$$

The multi-index $\gamma_0 = \alpha_0 + \beta_0$ has $|\gamma_0| = m + n$, and is a likely candidate for the solution to (39); note that $a_{\alpha_0} b_{\beta_0} \neq 0$ *by definition*. We leave the reader to verify a few simple properties of this particular multi-index.

A.3. Exercise. Explain why $\alpha_0 = \max_{\succ}\{\alpha : |\alpha| = m \text{ and } a_\alpha \neq 0\}$ might not be the same as $\alpha_0' = \max_{\succ}\{\alpha : a_\alpha \neq 0\}$. Is there any reason to expect α_0' to have maximal degree

$|\alpha_0'| = m$? $\square$

A.4. Exercise. In $\gamma_0 = \alpha_0 + \beta_0$ we have $|\alpha_0| = m$ and $|\beta_0| = n$, and $a_{\alpha_0} b_{\beta_0} \neq 0$, by definition. If α, β are any indices such that

$$|\alpha + \beta| = |\alpha_0 + \beta_0| = m + n \qquad \text{and} \qquad a_\alpha b_\beta \neq 0$$

prove that we must have $|\alpha| = |\alpha_0| = m$ and $|\beta| = |\beta_0| = n$. $\square$

Defining $\alpha_0, \beta_0, \gamma_0 = \alpha_0 + \beta_0$ as above, we make the following claim:

> CLAIM: If $\alpha + \beta = \alpha_0 + \beta_0$ and $a_\alpha b_\beta \neq 0$ then $\alpha = \alpha_0$ and $\beta = \beta_0$. Hence the sum
>
> $$c_{\gamma_0} = \sum_{\alpha+\beta=\gamma_0} a_\alpha b_\beta \tag{A.1}$$
>
> reduces to the single nonzero term $a_{\gamma_0} b_{\beta_0}$

A.5. Exercise. Prove the claim made in (A.1) using the facts assembled in the preceding discussion. $\square$

That will complete the proof of the Degree Formula.

Chapter II. Linear Operators $\mathbf{T : V \to W}$.

II.1 Generalities. A map $T : V \to W$ between vector spaces is a **linear operator** if for any $v, v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$

1. Scaling operations are preserved: $T(\lambda{\cdot}v) = \lambda{\cdot}T(v)$
2. Sums are preserved: $T(v_1 + v_2) = T(v_1) + T(v_2)$

This is equivalent to saying

$$T\Big(\sum_i \lambda_i v_i\Big) \;=\; \sum_i \lambda_i T(v_i) \qquad \text{in } W$$

for all finite linear combinations of vectors in V. A trivial example is the **zero operator** $T(v) = 0_W$, for every $v \in V$. If $W = V$ the **identity operator**, $\mathrm{id}_V : V \to V$ is given by $\mathrm{id}(v) = v$ for all vectors. Some basic properties of any linear operator $T : V \to W$ are:

1. $T(0_V) = 0_W$. [Proof: $T(0_V) = T(0 \cdot 0_V) = 0{\cdot}T(0_V) = 0_W$.]
2. $T(-v) = -T(v)$. [Proof: $T(-v) = T((-1) \cdot v) = (-1) \cdot T(0_V) = 0_W$.]
3. A linear operator is determined by its action on any set S of vectors that span V. If $T_1, T_2 : V \to W$ are linear operators and

   $$T_1(s) = T_2(s) \quad \text{ for all } \quad s \in S,$$

   then $T_1 = T_2$ everywhere on V. [Proof: Any $v \in V$ is a finite linear combination $v = \sum_i c_i s_i$; then $T_1(v) = \sum_i c_i T_i(s_i) = T_2(v)$.]

1.1. Exercise. If $S \subseteq V$ and $T : V \to W$ is a linear operator prove that

$$T\big(\mathbb{K}\text{-span}\{S\}\big) \text{ is equal to } \mathbb{K}\text{-span}\{T(S)\}.$$

1.2. Definition. *We write* $\mathrm{Hom}_{\mathbb{K}}(V, W)$ *for the space of linear operators* $T : V \to W$. *This becomes a vector space over* $\mathbb{K}$ *if we define*

1. $(T_1 + T_2)(v) = T_1(v) + T_2(v)$;
2. $(\lambda{\cdot}T)(v) = \lambda{\cdot}(T(v))$.

for any $v \in V$, $\lambda \in \mathbb{K}$ *. The vector space axioms are easily verified. The zero element in* $\mathrm{Hom}(V, W)$ *is the zero operator:* $0(v) = 0_W$ *for every* $v \in V$. *The additive inverse* $-T$ *is is the operator* $-T(v) = (-1){\cdot}T(v) = T(-v)$, *which is also a scalar multiple* $-T = (-1)T$ *of* T. □

If $V = W$ we can also define the **composition product** $S \circ T$ of two operators,

$$(S \circ T)(v) = S(T(v)) \qquad \text{for all } v \in V$$

This makes $\mathrm{Hom}_{\mathbb{K}}(V) = \mathrm{Hom}_{\mathbb{K}}(V, V)$ a noncommutative associative algebra with identity $I = \mathrm{id}_V$.

A linear operator $T : V \to W$ over $\mathbb{K}$ determines two important vector subspaces, the **kernel** $K(T) = \ker(T)$ in the initial space V and the **range** $R(T) = \mathrm{range}(T)$ in the target space W.

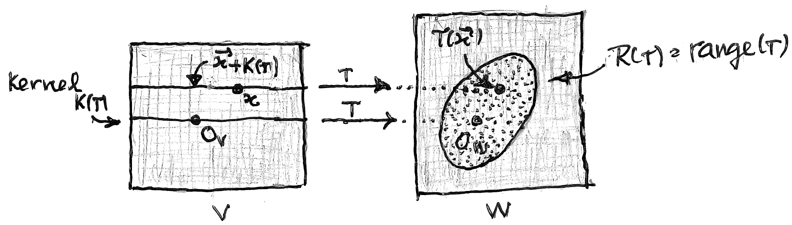

Figure 2.1. A linear map $T : V \to W$ sends all points in a coset $x + K(T)$ of the kernel to a single point $T(x)$ in W. Different cosets map to different points, and all images land within the range $R(T)$. The zero coset $0_V + K(T)$ collapses to the origin 0_W in W.

1. $K(T) = \ker(T) = \{v \in V : T(v) = 0_W\}$. The dimension of this space is often referred to as Nullity(T).
2. $R(T) = \text{range}(T)$ is the image set $T(V) = \{T(v) : v \in V\}$. Its dimension is the **rank**
$$\text{rank}(T) = \dim_{\mathbb{K}}\big(\text{range}(T)\big) \ ,$$
which is often abbreviated as $\text{rk}(T)$.

1.3. Exercise. Show that $\ker(T)$ and $\text{range}(T)$ are vector subspaces of V and W respectively.

We often have to decide whether a linear map is one-to-one, onto, or a bijection. For surjectivity, we must compute the range $R(T)$; determining whether T is one-to-one is easier, and amounts to computing the kernel $K(T)$. The diagram Figure 2.1 illustrates the general behavior of any linear operator. Each coset $v + K(T)$ of the kernel gets mapped to a single point in W because
$$T(v + K(T)) = T(v) + T(K(T)) = T(v) + 0_W = T(v) \quad ,$$
and distinct cosets go to different points in W. All points in V map into the range $R(T) \subseteq W$.

1.4. Lemma. *A linear operator $T : V \to W$ is one-to-one if and only if* $\ker(T) = 0$.

Proof: $(\Leftarrow)$. If $T(v_1) = T(v_2)$ for $v_1 \neq v_2$, then $0 = T(v_1) - T(v_2) = T(v_1 - v_2)$, so $v_2 - v_1 \neq 0$ is in $\ker(T)$ and the kernel is nontrivial. We have just proved the "contrapositive" $\neg\big(T \text{ is one-to-one}\big) \Rightarrow \neg\big(K(T) = \{0\}\big)$ of the statment $(\Leftarrow)$ we want, but the two are logically equivalent.
Moral: If you want to prove $(P \Rightarrow Q)$ it is sometimes easier to prove the equivalent *contrapositive statement* $(\neg Q \Rightarrow \neg P)$, as was the case here.

Proof: $(\Rightarrow)$. Suppose T is one-to-one. If $v \neq 0_V$ then $Tv \neq T(0_V) = 0_W$ so $v \notin K(T)$ and $K(T)$ reduces to $\{0\}$. $\square$

The following important result is closely related to Theorem 5.7 (Chapter I) for quotient spaces.

1.5. Theorem (The Dimension Theorem). *If $T : V \to W$ is a linear operator and V is finite dimensional, the range $R(T)$ is finite dimensional and is related to the kernel $K(T)$ via*

$$\dim(R(T)) + \dim(K(T)) = \dim(V) \tag{7}$$

In words, "rank + nullity = dimension of the initial space V." This can also be expressed in short form by writing $|R(T)| + |K(T)| = |V|$.

Proof: The kernel is finite dimensional because $K(T) \subseteq V \Rightarrow \dim(K(T)) \leq \dim(V) < \infty$. The range $R(T)$ is also finite dimensional. In fact, if $\{v_1, ..., v_n\}$ is a basis for V every $w \in R(T)$ has the form $w = T(v) = \sum_i c_i T(v_i)$, so the vectors $T(v_1), ..., T(v_n)$ span $R(T)$. Therefore $\dim(R(T)) \leq n = \dim(V)$.

Now let $\{w_1, \ldots, w_m\}$ be a basis for $K(T)$. By adjoining additional vectors from V we can obtain a basis $\{w_1, ..., w_m, v_{m+1}, ..., v_{m+k}\}$ for V. Obviously, $m = \dim(K(T))$ and $m + k = \dim(V)$. To prove $k = \dim(R(T))$ we show the vectors $T(v_{m+1}), ..., T(v_{m+k})$ are a basis for $R(T)$. They certainly span $R(T)$ because $w \in R(T) \Rightarrow w = T(v)$ for some $v \in V$, which can be written

$$v = c_1 w_1 + ... + c_m w_m + c_{m+1} v_{m+1} + \ldots + c_{m+k} v_{m+k} \quad (c_j \in \mathbb{K})$$

Since $w_j \in K(T)$ and $T(w_j) = 0_W$ we see that

$$w = T(v) = 0_W + ... + 0_W + \sum_{j=1}^{k} c_{m+j} T(v_{m+j})$$

so $v \in \mathbb{K}\text{-span}\{T(v_{r+1}), \ldots, T(v_{r+k})\}$. These vectors are also independent, for if

$$0_W = \sum_{i=1}^{k} c_{m+i} T(v_{m+i}) = T\Big(\sum_{i=1'}^{k} c_{m+i} v_{m+i}\Big)$$

that means $\sum_i c_{m+i} v_{m+i} \in K(T)$ and there are coefficients $c_1, ..., c_m$ such that $\sum_{i=1}^{m} c_i w_i = \sum_{j=1}^{k} c_{m+j} v_{m+j}$, or

$$0_V = -c_1 w_1 - ... - c_m w_m + c_{m+1} v_{m+1} + ... + c_{m+k} v_{m+k}$$

Because $\{w_1, ..., v_{m+k}\}$ is a basis for V we must have $c_1 = ... = c_{m+k} = 0$, proving independence of $T(v_{m+1}), ..., T(v_{m+k})$. Thus $\dim(R(T)) = k$ as claimed. $\square$

1.6. Corollary. *Let $T : V \to W$ be a linear operator between finite dimensional vector spaces such that $\dim(V) = \dim(W)$, which certainly holds if $V = W$. Then the following assertions are equivalent:*

(i) *T is one-to-one* (ii) *T is surjective* (iii) *T is bijective.*

Proof: By the Dimension Theorem we have $|K(T)| + |R(T)| = |V|$. If T is one-to-one then $K(T) = (0)$, so by (7) $|R(T)| = |V| = |W|$. Since $R(T) \subseteq W$ the only way that this can happen is to have $R(T) = W$ – i.e. T is surjective. Finally, if T is surjective then $|R(T)| = |W| = |V|$ by hypothesis. Invoking (7) we see that $|K(T)| = 0$, the kernel is trivial, and T is one-to-one.

We just proved that T is one-to-one if and only if T is surjective, so either condition implies T is bijective. $\square$

1.7. Exercise. Explain why a spanning set $\mathfrak{X} = \{v_1, ..., v_n\}$ is a basis for a finite dimensional space if and only the vectors in $\mathfrak{X}$ are independent ($\sum_{i=1}^{n} c_i v_i = 0_V \Rightarrow c_1 = \ldots = c_n = 0$).

1.8. Proposition. *Let V be a finite dimensional vector space and $\{v_1, ..., v_n\}$ a basis. Select any n vectors $w_1, ..., w_n$ in some other vector space W. Then, there is a unique linear operator $T : V \to W$ such that $T(v_i) = w_i$ for $1 \leq i \leq n$.*

Proof: Uniqueness of T (if it exists) was proved in our initial comments about linear operators. To construct such a T define $T(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i w_i$ for all choices of $\lambda_1, ..., \lambda_n \in \mathbb{K}$. This is obviously well-defined since $\{v_i\}$ is a basis, and is easily seen to be a linear operator from $V \to W$. $\square$

One prolific source of linear operators is the correspondence between $n \times m$ matrices A with entries in $\mathbb{K}$ and the linear operators $L_A : \mathbb{K}^m \to \mathbb{K}^n$ determined by matrix multiplication

$$L_A(\mathbf{v}) = A \cdot \mathbf{v} \qquad \Big(\text{matrix product } (n \times m)\cdot(m \times 1) = (n \times 1)\Big) ,$$

if we regard $\mathbf{v} \in \mathbb{K}^m$ as an $m \times 1$ column vector.

1.9. Example. Let $L_A : \mathbb{R}^3 \to \mathbb{R}^4$ (or $\mathbb{C}^3 \to \mathbb{C}^4$, same discussion) be the linear operator associated with the 4×3 matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Describe $\ker(L_A)$ and $\mathrm{range}(L_A)$ by finding explicit basis vectors in these spaces.

Solution: The range $R(L_A)$ of L_A is determined by finding all $\mathbf{y}$ for which there is an $\mathbf{x} \in \mathbb{K}^3$ such that $A\mathbf{x} = \mathbf{y}$. Row reduction of the augmented matrix $[A : \mathbf{y}]$ yields

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & y_1 \\ 1 & 0 & 2 & y_2 \\ 2 & 1 & 1 & y_3 \\ 1 & 1 & 1 & y_4 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 3 & y_1 \\ 0 & -2 & -1 & y_2 - y_1 \\ 0 & -3 & -5 & y_3 - 2y_1 \\ 0 & -1 & -2 & y_4 - y_1 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 2 & 3 & y_1 \\ 0 & 1 & 2 & y_1 - y_4 \\ 0 & 2 & 1 & y_1 - y_2 \\ 0 & 3 & 5 & 2y_1 - y_3 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|c} 1 & 2 & 3 & y_1 \\ 0 & 1 & 2 & y_1 - y_4 \\ 0 & 0 & -3 & y_1 - y_2 - 2(y_1 - y_4) \\ 0 & 0 & -1 & 2y_1 - y_3 - 3(y_1 - y_4) \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 2 & 3 & y_1 \\ 0 & \boxed{1} & 2 & y_1 - y_4 \\ 0 & 0 & \boxed{1} & y_1 + y_3 - 3y_4 \\ 0 & 0 & 0 & 2y_1 - y_2 + 3y_3 - 7y_4 \end{array}\right)$$

There are no solutions $\mathbf{x} \in \mathbb{K}^3$ unless $\mathbf{y} \in \mathbb{K}^4$ lies the 3-dimensional solution set of the equation

$$2y_1 - y_2 + 3y_3 - 7y_4 = 0.$$

When this constraint is satisfied, backsolving yields exactly one solution for each such $\mathbf{y}$; there are no free variables.

Thus $R(L_A)$ is the solution set of equation

$$2y_1 - y_2 + 3y_3 - 7y_4 = 0$$

When this is written as a matrix equation $B\mathbf{y} = \mathbf{0}$ ($B =$ the 1×4 matrix $[2, -1, 3, -7]$), y_2, y_3, y_4 are free variables and then $y_1 = \frac{1}{2}(y_2 - 3y_3 + 7y_4)$, so a typical vector in $R(L_A)$ has the form

$$\mathbf{y} = \begin{pmatrix} \frac{1}{2}(y_2 - 3y_3 + 7y_4) \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = y_2 \cdot \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix} + y_3 \cdot \begin{pmatrix} -\frac{3}{2} \\ 0 \\ 1 \\ 0 \end{pmatrix} + y_4 \cdot \begin{pmatrix} \frac{7}{2} \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

with $y_1, y_2, y_3 \in \mathbb{K}$. The re-scaled column vectors $\mathbf{u}_2 = \mathrm{col}(1, 2, 0, 0)$, $\mathbf{u}_3 = \mathrm{col}(-3, 0, 2, 0)$, $\mathbf{u}_4 = \mathrm{col}(7, 0, 0, 2)$ obviously span $R(L_A)$ and are easily seen to be linearly independent, so they are a basis for the range, which has dimension $|R(L_A)| = 3$.

The kernel: Now we want to find all solutions $\mathbf{x} \in \mathbb{K}^3$ of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. The same row operations used above transform $[A : \mathbf{0}]$ to:

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & 3 & 0 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Then $\mathbf{x} = \text{col}(x_1, x_2, x_3)$ has entries $x_1 = x_2 = x_3 = 0$. Therefore

$$K(L_A) = \{\mathbf{x} \in \mathbb{K}^3 : L_A(\mathbf{x}) = A\mathbf{x} = \mathbf{0}\} \text{ is the trivial subspace } \{\mathbf{0}\}$$

and there is no basis to be found.

Note that $|R(L_A)| + |K(L_A)| = 3 + 0 =$ dimension of the *initial space* $\mathbb{K}^3$, while the target space $W = \mathbb{K}^4$ has dimension $= 4$. $\square$

1.10. Example. Let $L_A : \mathbb{R}^4 \to \mathbb{R}^4$ be the linear operator associated with the 4×4 matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 1 & 0 & 2 & -1 \\ 2 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Describe the kernel $K(L_A)$ and range $R(L_A)$ by finding basis vectors.

Solution: The range of L_A is determined by finding all $\mathbf{y}$ such that $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^4$. Row reduction of $[A : \mathbf{y}]$ yields

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & y_1 \\ 1 & 0 & 2 & -1 & y_2 \\ 2 & 1 & 1 & 2 & y_3 \\ 1 & 1 & 1 & 1 & y_4 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & y_1 \\ 0 & -2 & -1 & -1 & y_2 - y_1 \\ 0 & -3 & -5 & 2 & y_3 - 2y_1 \\ 0 & -1 & -2 & 1 & y_4 - y_1 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & y_1 \\ 0 & 1 & 2 & -1 & y_1 - y_4 \\ 0 & 2 & 1 & 1 & y_1 - y_2 \\ 0 & 3 & 5 & -2 & 2y_1 - y_3 \end{array}\right)$$

$$\to \left(\begin{array}{cccc|c} 1 & 2 & 3 & 0 & y_1 \\ 0 & 1 & 2 & -1 & y_1 - y_4 \\ 0 & 0 & -3 & 3 & y_1 - y_2 - 2(y_1 - y_4) \\ 0 & 0 & -1 & 1 & 2y_1 - y_3 - 3(y_1 - y_4) \end{array}\right) \to \left(\begin{array}{cccc|c} \boxed{1} & 2 & 3 & 0 & y_1 \\ 0 & \boxed{1} & 2 & -1 & y_1 - y_4 \\ 0 & 0 & \boxed{1} & -1 & y_1 + y_3 - 3y_4 \\ \hline 0 & 0 & 0 & 0 & 2y_1 - y_2 + 3y_3 - 7y_4 \end{array}\right)$$

There are *no* solutions $\mathbf{x}$ in $\mathbb{R}^4$ unless $\mathbf{y}$ lies the 3-dimensional solution set of the linear equation

$$2y_1 - y_2 + 3y_3 - 7y_4 = 0$$

– i.e. $\mathbf{y}$ is a solution of the matrix equation

$$C\mathbf{y} = \mathbf{0} \qquad \text{where} \qquad C = \begin{bmatrix} 2\,, -1\,, 3\,, -7 \end{bmatrix}_{1\times 4}$$

Then there exist multiple solutions, and $R(L_A) = \{\mathbf{y} \in \mathbb{R}^4 : C\mathbf{y} = \mathbf{0}\}$ is nontrivial. Multiplying C by $\frac{1}{2}$ puts it in echelon form, so y_2, y_3, y_4 are free variables in solving $C\mathbf{y} = \mathbf{0}$ and $y_1 = \frac{1}{2}(y_2 - 3y_3 + 7y_4)$. Thus a vector $\mathbf{y} \in \mathbb{R}^4$ is in the range $R(L_A) \Leftrightarrow$

$$\mathbf{y} = \begin{pmatrix} \frac{1}{2}(y_2 - 3y_3 + 7y_4) \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = y_2 \cdot \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix} + y_3 \cdot \begin{pmatrix} -\frac{3}{2} \\ 0 \\ 1 \\ 0 \end{pmatrix} + y_4 \cdot \begin{pmatrix} \frac{7}{2} \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

with $y_1, y_2, y_3 \in \mathbb{R}^3$. The column vectors $\mathbf{u}_2 = \text{col}(1, 2, 0, 0)$, $\mathbf{u}_3 = \text{col}(-3, 0, 2, 0)$, $\mathbf{u}_4 = \text{col}(7, 0, 0, 2)$, obviously span $R(L_A)$ and are easily seen to be linearly independent, so they are a basis for the range and $\dim(R(L_A)) = 3$. (Hence also $|K(L_A)| = |V| - |R(L_A)| = 1$ by the Dimension Theorem.)

The kernel: $K(L_A)$ can be found by setting $\mathbf{y} = 0$ in the preceding echelon form of $[A : \mathbf{y}]$, which becomes

$$\left(\begin{array}{cccc|c} \boxed{1} & 2 & 3 & 0 & 0 \\ 0 & \boxed{1} & 2 & -1 & 0 \\ 0 & 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Now x_4 is the free variable and

$$\begin{aligned} x_3 &= x_4 \\ x_2 &= -2x_3 + x_4 = -x_4 \\ x_1 &= -2x_2 - 3x_3 = -x_4 \ , \end{aligned}$$

hence,

$$K(L_A) = \mathbb{K} \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

is one dimensional as expected.

Given a vector $\mathbf{y}$ in the range $R(L_A)$ we can find a particular solution $\mathbf{x}_0$ of $A\mathbf{x} = \mathbf{y}$ by setting the free variable $x_4 = 0$. Then

$$\begin{aligned} x_3 &= x_4 + y_1 + y_3 - 3y_4 = y_1 + y_3 - 3y_4, \\ x_2 &= -2x_3 + x_4 + y_1 - y_4 = -y_1 - 2y_3 + 5y_4, \\ x_1 &= -2x_2 - 3x_3 + y_1 = y_1 + (2y_1 + 4y_3 - 10y_4) + (-3y_1 - 3y_3 + 9y_4) \\ &= y_3 - y_4 \end{aligned}$$

and

$$\mathbf{x}_0 = \begin{pmatrix} y_3 - y_4 \\ -y_1 - 2y_3 + 5y_4 \\ y_1 + y_3 - 3y_4 \\ 0 \end{pmatrix}$$

is a particular solution for $A\mathbf{x} = \mathbf{y}$. The full set of solutions is the additive coset $\mathbf{x}_0 + K(L_A)$ of the kernel of L_A. $\square$

1.11. Exercise. If $A \in \text{M}(n \times m, \mathbb{K})$ prove that

1. $\text{Range}(L_A)$ is equal to column space $\text{Col}(A) = \mathbb{K}\text{-span}\{C_1, \ldots, C_m\}$, the subspace of $\mathbb{K}^n$ spanned by the columns of A.

so

$$\dim\Big(\text{Range}(L_A)\Big) = \dim\big(\text{Col}(A)\big) = 3$$

II.2. Invariant Subspaces.

If $T : V \to V$ $(V = W)$ is a linear operator, a subspace W is **T-invariant** if $T(W) \subseteq W$. Invariant subspaces are important in determining the structure of T, as we shall see. Note that the subspaces (0), $R(T) = \text{range}(T) = T(V)$, $K(T) = \ker(T)$, and V are all T-invariant. Structural analysis of T proceeds initially by searching for **eigenvectors**: nonzero vectors $v \in V$ such that $T(v)$ is a scalar multiple $T(v) = \lambda \cdot v$, for some $\lambda \in \mathbb{K}$.

These are precisely the vectors in $\ker(T - \lambda I)$ where $I : V \to V$ is the identity operator on V. Eigenvectors may or may not exist; when they do they have a story to tell.

2.1. Definition. *Fix a linear map $T : V \to V$ and scalar $\lambda \in \mathbb{K}$. The* **λ-eigenspace** *is*

$$E_\lambda(T) = \{v \in V : T(v) = \lambda \cdot v\} = \{v \in V : \big(T - \lambda I\big)(v) = 0\} = \ker(T - \lambda I)$$

We say that $\lambda \in \mathbb{K}$ is an **eigenvalue** *for T if the eigenspace is nontrivial, $E_\lambda(T) \neq (0)$.*

If V is finite dimensional we will eventually see that the number of eigenvalues is $\leq n$ (possibly zero) because it is the set of roots in $\mathbb{K}$ of the "**characteristic polynomial**"

$$p_T(x) = \det(T - xI) \in \mathbb{K}[x] \ ,$$

which has degree $n = \dim_{\mathbb{K}}(V)$. The set of DISTINCT eigenvalues in $\mathbb{K}$ is called the **spectrum of** T and is denoted

$$\begin{aligned} \mathrm{sp}_{\mathbb{K}}(T) &= \{\lambda \in \mathbb{K} : \text{ such that } T(v) = \lambda \cdot v \text{ for some } v \neq 0\} \\ &= \{\lambda \in \mathbb{K} : E_\lambda \neq (0)\} \end{aligned}$$

Depending on the nature of the ground field, $\mathrm{sp}_{\mathbb{K}}(T)$ may be the empty set; it is always nonempty if $\mathbb{K} = \mathbb{C}$, because every nonconstant polynomial has at least one root in $\mathbb{C}$ (Fundamental Theorem of Algebra). The point is that all eigenspaces E_λ are T-invariant subspaces because T and $(T - \lambda I)$ "commute," hence

$$(T - \lambda I)(Tv) = T\big((T - \lambda I)v\big) = T(0) = 0 \qquad \text{if} \qquad v \in E_\lambda$$

The E_λ are also "essentially disjoint" from each other in the sense that $E_\mu \cap E_\lambda = (0)$ if $\mu \neq \lambda$. (You can't have $\lambda \cdot v = \mu \cdot v$ (or $(\lambda - \mu) \cdot v = 0$) for nonzero v if $\mu \neq \lambda$.) Note that $\ker(T)$ is the eigenspace corresponding to $\lambda = 0$ since

$$E_{\lambda=0} = \{v : (T - 0 \cdot I)(v) = T(v) = 0\} = \ker(T) \ ,$$

and $\lambda = 0$ is an eigenvalue in $\mathrm{sp}_{\mathbb{K}}(T) \Leftrightarrow$ this kernel is nontrivial. When $\lambda = 1$, $E_{\lambda=1}$ is the set of "fixed points" under the action of T.

$$E_{\lambda=1} = \mathrm{Fix}(T) = \{v : T(v) = v\} \qquad \text{(the fixed points in } V)$$

Decomposition of Operators.

We now show that if $W \subseteq V$ is an **invariant subspace**, so $T(W) \subseteq W$, then T induces linear operators in W and in the quotient space V/W:

1. RESTRICTION: $T|_W : W \to W$ is the restriction of T to W, so $\big(T|_W\big)(w) = T(w)$, for all $w \in W$.

2. QUOTIENT OPERATOR: The operator $\tilde{T} : V/W \to V/W$, sometimes denoted $T_{V/W}$, is induced by the action of T on additive cosets:

   (8) $$T_{V/W}(x + W) = T(x) + W \qquad \text{for all cosets in } V/W \ .$$

   The outcome is determined using a representative x for the coset, but as shown below different representatives yield the same result, so $T_{V/W}$ is well defined.

2.2. Theorem. *Given a linear operator $T : V \to V$ and an invariant subspace W, there is a unique linear operator $\tilde{T} : V/W \to V/W$ that makes the following diagram "commute" in the sense that $\pi \circ T = \tilde{T} \circ \pi$, where $\pi : V \to V/W$ is the quotient map.*

$$\begin{array}{ccc} V & \xrightarrow{T} & V \\ \pi \downarrow & & \downarrow \pi \\ V/W & \xrightarrow{\tilde{T}} & V/W \end{array}$$

Note: We have already shown that the quotient map $\pi : V \to V/W$ is a linear operator between vector spaces.

Proof: Existence. The restriction $T|_W$ is clearly a linear operator on W. As for the induced map $\tilde{T}$ on V/W, the fact that $T(W) \subseteq W$ implies

$$T(x+W) = T(x) + T(W) \subseteq T(x) + W \qquad \text{for } x \in V$$

This suggests that (8) is the right definition. It automatically insures that $\tilde{T} \circ \pi = \pi \circ T$; the problem is to show the outcome $\tilde{T}(x+W) = T(x) + W$ is independent of the choice of coset representative $x \in V$. But if $x' + W = x + W$ we have $x' = x' + 0 = x + w_0$ for some $w_0 \in W$ and

$$T(x') + W = T(x + w_0) + W = T(x) + \big(T(w_0) + W\big)$$

Since W is invariant $T(w_0) \in W$ and $T(w_0) + W = W$ by Exercise 5.2 of Chapter 1. Hence $T(v') + W = T(v) + W$ and the induced operator in (8) is well-defined. Linearity of $\tilde{T}$ is easily checked. Once we know $\tilde{T}$ is well-defined we get:

$$\begin{aligned}
\tilde{T}\big((v_1+W) \oplus (v_2+W)\big) &= \tilde{T}\big(v_1+v_2+W\big) \qquad \text{(definition of}(\oplus) \text{ in } V/W)\\
&= T(v_1+v_2)+W \;=\; T(v_1)+T(v_2)+W\\
&= \big(T(v_1)+W\big) + \big(T(v_2)+W\big) \qquad (\text{since } W+W=W)\\
&= \tilde{T}(v_1+W) \oplus \tilde{T}(v_2+W)
\end{aligned}$$

and similarly

$$\tilde{T}\big(\lambda \odot (v+W)\big) = \tilde{T}(\lambda\cdot v + W) = \lambda \odot \tilde{T}(v+W)$$

Uniqueness. If $\tilde{T}_1$, $\tilde{T}_2$ both satisfy the commutation relation $\tilde{T}_i \circ \pi = \pi \circ T_i$, then

$$\tilde{T}_1(v+W) = \tilde{T}_1 \circ \pi(v) = \pi(T(v)) = \tilde{T}_2(\pi(v)) = \tilde{T}_2(v+W)$$

so $\tilde{T}_1 = \tilde{T}_2$ on V/W. $\square$

A Look Ahead.

If $T : V \to V$ is a linear operator on a finite dimensional space, we will explain in Section II.4 how a matrix $[T]_{\mathfrak{X}}$ is associated with T once a basis $\mathfrak{X}$ in V has been specified. If W is a T-invariant subspace we will see that much of the structural information about T resides in the induced operators $T|_W$ and $T_{V/W}$, and that in some sense (to be made precise) T is assembled by "joining together" these smaller pieces. This is a big help in trying to understand the action of T on V, but it does depend on being able to find invariant subspaces – the more the better! To illustrate: if a basis $\mathfrak{X} = \{w_1, \ldots, w_m\}$ in W is augmented to get a basis for V,

$$\mathfrak{Z} = \{w_1, \ldots, w_m, v_{m+1}, \ldots, v_{m+k}\} \qquad (m + k = n = \dim(V))$$

we have seen that the image vectors $\bar{v}_{m+i} = \pi(v_{m+i})$ are a basis $\mathfrak{Y}$ in the quotient space V/W. In Section II.4 we will show that the matrix $[T]_{\mathfrak{Z}}$ assumes a special "**block-upper triangular form**" with respect to such a basis.

$$[T]_{\mathfrak{Z}} \;=\; \begin{pmatrix} \boxed{\mathrm{A}}_{m\times m} & \boxed{\;*\;}_{m\times k} \\ \mathbf{0}_{k\times m} & \boxed{\mathrm{B}}_{k\times k} \end{pmatrix}$$

where $A = [T|_W]_{\mathfrak{X}}$ and $B = [T_{V/W}]_{\mathfrak{Y}}$. Clearly, much of the information about T is encoded in the two diagonal blocks A, B; but some information is lost in passing from T to $T|_W$ and $T_{V/W}$ – the "cross-terms" in the upper right block $\boxed{\;*\;}$ cannot be determined if we only know the two induced operators. Additional information is needed to piece them together to recover T, and there may be more than one operator T yielding a particular pair of induced operators $\big(T|_W, T_{V/W}\big)$. $\square$

Isomorphisms of Vector Spaces. A linear map $T : V \to W$ is an **isomorphism** between vector spaces if it is a bijection. Since T is a bijection there is a well-defined "set-theoretic" inverse map in the opposite direction $T^{-1} : W \to V$

$$T^{-1}(w) = (\textit{the unique } v \in V \textit{ such that } T(v) = w)$$

for any $w \in W$. In general it might not be easy to describe the inverse of a bijection $f : X \to Y$ between two point sets in closed form

$$f^{-1}(y) = (\textit{some explicit formula})$$

(Try finding $x = f^{-1}(y)$ where $f : \mathbb{R} \to \mathbb{R}$ is $y = f(x) = x^3 + x + 1$, which is a bijection because $df/dx > 0$ for all x.) But the inverse of a linear map is automatically linear, if it exists. We write $V \cong W$ if there is an isomorphism between them.

2.3. Exercise. Suppose $T : V \to W$ is linear and a bijection. Prove that the set-theoretic inverse map $T^{-1} : W \to V$ must be linear, so

$$T^{-1}(\lambda w) = \lambda \cdot T^{-1}(w) \qquad \text{and} \qquad T^{-1}(w_1 + w_2) = T^{-1}(w_1) + T^{-1}(w_2)$$

Thus T and T^{-1} are both *isomorphisms* between V and W.
Hint: T^{-1} reverses the action of T and vice-versa, so $T \circ T^{-1} = \mathrm{id}_W$ and $T^{-1} \circ T = \mathrm{id}_V$.

We now observe that an isomorphisms between vector spaces V and W identifies important features of V with those of W. It maps

$$\left\{ \begin{array}{c} \text{independent sets} \\ \text{spanning sets} \\ \text{bases} \end{array} \right. \quad \text{in } V \quad \longrightarrow \quad \left\{ \begin{array}{c} \text{independent sets} \\ \text{spanning sets} \\ \text{bases} \end{array} \right. \quad \text{in } W$$

To illustrate, if $\{v_1, ..., v_n\}$ are independent in V, then

$$0_W = \sum_{i=1}^n c_i T(v_i) = T\Big(\sum_i c_i v_i\Big) \;\;\Rightarrow\;\; 0_V = \sum_{i=1}^n c_i v_i \text{ in } V \;\;\Rightarrow\;\; c_1 = ... = c_n = 0 \text{ in } \mathbb{K}$$

because $T(0_V) = 0_W$ and T is one-to-one. Thus the vectors $\{T(v_1), ..., T(v_n)\}$ are independent in W. Similar arguments yield the other two assertions.

2.4. Exercise. If $T : V \to W$ is an isomorphism of vector spaces, verify that:

1. $\mathbb{K}\text{-span}\{v_i\} = V \;\Rightarrow\; \mathbb{K}\text{-span}\{T(v_i)\} = W$;
2. $\{v_i\}$ is a basis in $V \;\Rightarrow\; \{T(v_i)\}$ is a basis in W.

In particular, isomorphic vector spaces V, W are either both infinite dimensional, or both finite dimensional with $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$.

The following result which relates linear operators, quotient spaces, and isomorphisms will be cited often in analyzing the structure of linear operators. It is even valid for infinite dimensional spaces.

2.5. Theorem (First Isomorphism Theorem). *Let $T : V \to R(T) \subseteq W$ be a linear map with range $R(T)$. If $K(T) = \ker(T)$, T induces a unique bijective linear map $\tilde{T} : V/K(T) \to R(T)$ that makes the following diagram "commute" ($\tilde{T} \circ \pi = T$).*

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\pi"'] & R(T) \subseteq W \\ V/K(T) \arrow[ur, "\tilde{T}"'] & \end{tikzcd}$$

where $\pi : V \to V/K$ is the quotient map. Furthermore $R(\tilde{T}) = R(T)$,
Hints: Try defining $\tilde{T}(v + K(T)) = T(v)$. Your first task is to show that the outcome is independent of the particular coset representative – i.e. $v' + K(T) = v + K(T) \Rightarrow T(v') = T(v)$, so $\tilde{T}$ is well-defined. Next show $\tilde{T}$ is linear, referring to the operations $\oplus$ and $\odot$ in the quotient space $V/K(T)$. The commutation property $\tilde{T} \circ \pi = T$ is built into the definition of $\tilde{T}$. For uniqueness, you must show that if $S : V/K(T) \to W$ is any other linear map such that $S \circ \pi = T$, then $S = \tilde{T}$; this is trivial once you clearly understand the question. $\square$

Note that range$(\tilde{T})$ = range(T) because the quotient map $\pi : V \to V/K$ is surjective: $w \in R(\tilde{T}) \Leftrightarrow$ there is a coset $v + K(T)$ such that $\tilde{T}(v + K(T)) = w$; but then $T(v) = w$ and $w \in R(T)$. Since $\tilde{T}$ is an isomorphism between $V/K(T)$ and $R(T)$, $\dim(V/K(T)) = \dim(V) - \dim(K(T))$ is equal to $\dim(R(T))$.

II.3. (Internal) Direct Sum of Vector Spaces.

A vector space V is an **(internal) direct sum** of subspaces V_1, ... ,V_n, indicated by writing $V = V_1 \oplus \ldots \oplus V_n$, if

1. The linear span $\sum_{i=1}^n V_i = \{\sum_{i=1}^n v_i : v_i \in V\}$ is all of V;
2. Every $v \in V$ has a *unique* representation as a sum $v = \sum_{i=1}^n v_i$ with $v_i \in V_i$.

Once we know that the V_i span V, condition (2.) is equivalent to saying

$$2^* \quad \sum_{i=1}^n v_i = 0 \quad \text{with } v_i \in V_i \quad \to \quad v_i = 0 \text{ for all } i.$$

In fact, if a vector v has two different representations $v = \sum_i v_i = \sum v_i'$ then $0 = \sum_{i=1}^n v_i''$ with $v_i'' = (v_i' - v_i) \in V_i$. Then ($2^*$.) implies $v_i'' = 0$ and $v_i' = v_i$ for all i. Conversely, if we can write $0 = \sum w_i$ with $w_i \in V_i$ not all zero, then the representation of a vector as $v = \sum_i v_i$ ($v_i \in V_i$) cannot be unique, since we could also write

$$0 = \sum_i w_i \qquad \text{with } w_i \neq 0 \text{ for some } i \ ,$$

and then $v = v + 0 = \sum_i (v_i + w_i)$ in which $v_i + w_i \in V_i$ is $\neq v_i$. Thus the condition (2.) is equivalent to (2^*.)

3.1. Example. We note the following examples of direct sum decompositions.

1. $\mathbb{K}^n = V \oplus W$ where

   $$V = \{(x_1, x_2, 0, ..., 0) : x_1, x_2 \in \mathbb{K}\} \text{ and } W = \{(0, 0, x_3, ..., x_n) : x_k \in \mathbb{K}\};$$

   More generally, in an obvious sense we have $\mathbb{K}^{m+n} \cong \mathbb{K}^m \oplus \mathbb{K}^n$.

2. The space of polynomials $\mathbb{K}[x] = V \oplus W$ is a direct sum of the subspaces
   - Even Polynomials: $V = \{\, \sum_{i=0}^\infty a_i x^i : a_i = 0 \text{ for odd indices}\}$
   - Odd Polynomials: $W = \{\, \sum_{i=0}^\infty a_i x^i : a_i = 0 \text{ for even indices }\}$

3. For $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}$, matrix space $\mathrm{M}(n, \mathbb{K})$ is a direct sum $\mathcal{A} \oplus \mathcal{S}$ of
   - Antisymmetric Matrices: $\mathcal{A} = \{A : A^{\mathrm{t}} = -A\}$
   - Symmetric Matrices: $\mathcal{S} = \{A : A^{\mathrm{t}} = A\}$.

In fact, since $(A^{\mathrm{t}})^{\mathrm{t}} = A$ we can write any matrix as

$$A = \frac{1}{2}(A^{\mathrm{t}} + A) + \frac{1}{2}(A^{\mathrm{t}} - A)$$

The first term is symmetric and the second antisymmetric, so $\mathrm{M}(n, \mathbb{K}) = \mathcal{A} + \mathcal{S}$, (linear span).

If $B \in \mathcal{A} \cap \mathcal{S}$ then $B^{\mathrm{t}} = -B$ and also $B^{\mathrm{t}} = B$, hence $B = -B$ and $B = \mathbf{0}$ (the zero matrix). Thus $\mathcal{A} \cap \mathcal{S} = (0)$ and Exercise 3.2 (below) implies that $\mathrm{M}(n, \mathbb{K}) = \mathcal{A} \oplus \mathcal{S}$.

Note: This actually works for any field $\mathbb{K}$ in which $2 = 1 + 1 \neq 0$ because the "$\frac{1}{2}$" in the formulas involves division by 2. In particular it works for the finite fields $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ except for $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$, in which $[1] \oplus [1] = [1 + 1] = [2] = [0]$ $\square$

If subspaces $V_1, \ldots, V_n$ span V it can be tricky to verify that V is a direct sum when $n \geq 3$, but if there are just two summands V_1 and V_2 (the case most often encountered) there is a simple and extremely useful criterion for deciding whether $V = V_1 \oplus V_2$.

3.2. Exercise. If E, F are subspaces of V show that V is the direct sum $E \oplus F$ if and only if

1. They span V: $E + F = \{a + b : a \in E, b \in F\}$ is all of V;
2. Trivial intersection: $E \cap F = \{0\}$.

It is important to note that this is NOT true when $n \geq 3$. If $\sum_{i=1}^{n} V_i = V$ and the spaces are only "pairwise disjoint,"

$$V_i \cap V_j = (0) \text{ for } i \neq j,$$

this is not enough to insure that V is a direct sum of the given subspaces (see the following exercise).

3.3. Exercise. Find three distinct 1-dimensional subspaces V_i in the two dimensional space $\mathbb{R}^2$ such that

1. $V_i \cap V_j = (0)$ for $i \neq j$;
2. $\sum_{i=1}^{3} V_i = \mathbb{R}^2$

Explain why $\mathbb{R}^2$ is not a *direct* sum $V_1 \oplus V_2 \oplus V_3$ of these subspaces.

3.4. Exercise. If $V = V_1 \oplus \ldots \oplus V_n$ and V is finite dimensional, we have seen that each V_i must be finite-dimensional with $\dim(V_i) \leq \dim(V)$.

1. Given bases $\mathfrak{X}_1 \subseteq V_1$, ... , $\mathfrak{X}_n \subseteq V_n$, explain how to create a basis for all of V;
2. Prove

$$\text{(9)} \qquad \textsc{Dimension Formula for Sums:} \quad \dim_{\mathbb{K}}(V_1 \oplus \ldots \oplus V_n) = \sum_{i=1}^{n} \dim_{\mathbb{K}}(V_i)$$

Direct sum decompositions play a large role in understanding the structure of linear operators. Suppose $T : V \to V$ and $V = V_1 \oplus V_2$, and that *both* subspaces are T-invariant. We get restricted operators $T_1 = T|_{V_1} : V_1 \to V_1$ and $T_2 = T|_{V_2} : V_2 \to V_2$, but because both subspaces are T-invariant we can fully reconstruct the original operator T in V from its "components" T_1 and T_2. In fact, every $v \in V$ has a unique decomposition $v = v_1 + v_2$ ($v_i \in V_i$) and then

$$T(v) = T(v_1) + T(v_2) = T_1(v_1) + T_2(v_2)$$

We often indicate this decomposition by writing $T = T_1 \oplus T_2$.

This does not work if only one subspace is invariant. But when both are invariant and we take bases $\mathfrak{X}_1 = \{v_1, \ldots, v_m\}$, $\mathfrak{X}_2 = \{v_{m+1}, \ldots, v_{m+k}\}$ for V_1, V_2, we will soon see that the combined basis $\mathfrak{Y} = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_{m+k}\}$ for all of V yields a matriix of particularly simple "**block-diagonal**" form

$$[T]_{\mathfrak{Y}} = \begin{pmatrix} \boxed{\mathrm{A}}_{m\times m} & \mathbf{0}_{\,m\times k} \\ \mathbf{0}_{\,k\times m} & \boxed{\mathrm{B}}_{k\times k} \end{pmatrix}$$

where A, B are the matrices of T_1, T_2 with respect to the bases $\mathfrak{X}_1 \subseteq V_1$ and $\mathfrak{X}_2 \subseteq V_2$.

Projections and Direct Sums. If $V = V_1 \oplus \ldots \oplus V_n$ then for each i there is a natural **projection operator** $P_i : V \to V_i \subseteq V$, the "*projection of V onto V_i along the complementary subspace $\bigoplus_{j\neq i} V_j$.*" By definition we have

$$(10) \qquad P_i(v) = v_i \text{ if } v = \sum_{j=1}^{n} v_j \text{ is the unique decomposition with } v_j \in V_j.$$

Note that $\ker(P_i) \supseteq$ all V_j with $j \neq i$, so $K(P_i) \supseteq \bigoplus_{j\neq i} V_j$. A number of properties of these projection operators are easily verified.

3.5. Exercise. Show that the projections P_i associated with a direct sum decomposition $V = V_1 \oplus \ldots \oplus V_n$ have the following properties.

1. Linearity: Each $P_i : V \to V$ is a linear operator;
2. Idempotent Property: $P_i^2 = P_i \circ P_i = P_i$ for all i;
3. $P_i \circ P_j = 0$ if $i \neq j$;
4. $\mathrm{Range}(P_i) = V_i$ and $\ker(P_i)$ is the linear span $\sum_{j\neq i} V_j$;
5. $P_1 + \ldots + P_n = I$ (identity operator on V).

If we represent vectors $v \in V$ as ordered n-tuples $(v_1, ..., v_n)$ in the Cartesian product set $V_1 \times \ldots \times V_n$, the i^{th} projection takes the form

$$P_i(v_1, ..., v_n) = (0, ..., 0, v_i, 0, ..., 0) \in V_i \subseteq V.$$

Don't be misled by this notation into thinking that we are speaking of orthogonal projections (onto orthogonal subspaces in $\mathbb{R}^n$). The following example and exercises illustrate what's really happening.

Note: A vector space must be equipped with additional structure such as an *inner product* if we want to speak of "orthogonality of vectors," or their "lengths." Such notions are meaningless in an unadorned vector space. Nevertheless, inner product spaces are important and will be fully discussed in Chapter VI.

3.6. Example. The plane $\mathbb{R}^2$ is a direct sum of the subspaces $V_1 = \mathbb{R}\mathbf{e}_1$ and $V_2 = \mathbb{R}(\mathbf{e}_1 + \mathbf{e}_2)$, where $\mathfrak{X} = \{\mathbf{e}_2, \mathbf{e}_2\}$ are the standard basis vectors $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ in $\mathbb{R}^2$. The maps

- P_1 projecting V onto V_1 along V_2,
- P_2 projecting V onto V_2 along V_1

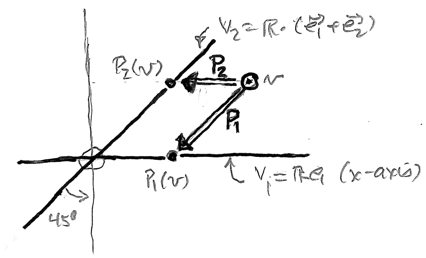

Figure 2.2. Projections P_1, P_2 determined by a direct sum decomposition $V = V_1 \oplus V_2$. Here $V = \mathbb{R}^2$, $V_1 = \mathbb{R}\mathbf{e}_1$, $V_2 = \mathbb{R}(\mathbf{e}_1 + \mathbf{e}_2)$; P_1 projects vectors $v \in V$ onto V_1 along V_2, and likewise for P_2.

are *oblique projections*, not the familiar *orthogonal projections* sending $\mathbf{x} = (x_1, x_2)$ to $(x_1, 0)$ and to $(0, x_2)$ respectively, see Figure 2.1. Find an explicit formula for $P_i(v_1, v_2)$ in $\mathbb{R}^2$, for arbitrary pairs (v_1, v_2) in $\mathbb{R}^2$.

Discussion: To calculate these projections we must write an arbitrary vector $\mathbf{v} = (v_1, v_2) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2$ in the form $a + b \in V_1 \oplus V_2$ where $V_1 = \mathbb{R}\mathbf{e}_1$ and $V_2 = \mathbb{R}(\mathbf{e}_1 + \mathbf{e}_2)$. The vectors $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$

$$\mathbf{f}_1 = \mathbf{e}_1 \qquad \text{and} \qquad \mathbf{f}_2 = \mathbf{e}_1 + \mathbf{e}_2 \tag{11}$$

that determine the 1-dimensional spaces V_1, V_2 are easily seen to be a new basis for $\mathbb{R}^2$. If $v = c_1\mathbf{f}_1 + c_2\mathbf{f}_2$ in the new basis, the action of the projections P_1, P_2 can be written immediately based on the definitions:

$$\begin{aligned} P_1(c_1\mathbf{f}_1 + c_2\mathbf{f}_2) &= c_1\mathbf{f}_1 \qquad \big(= c_1\mathbf{e}_1\big) \\ P_2(c_1\mathbf{f}_1 + c_2\mathbf{f}_2) &= c_2\mathbf{f}_2 \qquad \big(= c_2(\mathbf{e}_1 + \mathbf{e}_2)\,\big) \end{aligned} \tag{12}$$

Now $v = (v_1, v_2)$ is $v_1\mathbf{e}_1 + v_2\mathbf{e}_2$ in terms of the standard basis $\mathfrak{X}$ in $\mathbb{R}^2$ and we want to describe the outcomes $P_1(v), P_2(v)$ in terms of the same basis.

First observe that the action of P_2 is known as soon as we know the action of P_1: by the Parallelogram Law for vector addition (see Figure 2.2) we have $P_1 + P_2 = I$, so

$$P_2(v) = (I - P_1)(v) = v - P_1(v) \qquad \text{for all } v \in \mathbb{R}^2$$

Second, the projections P_i are linear so their action is known once we determine the images $P_i(\mathbf{e}_k)$ of the basis vectors $\mathbf{e}_k$ because

$$P_i(v) = P_1(v_1, v_2) = P_i\big(v_1\mathbf{e}_1 + v_2\mathbf{e}_2\big) \;=\; v_1 \cdot P_i(\mathbf{e}_1) + v_2 \cdot P_i(\mathbf{e}_2) \qquad (v_1, v_2 \in \mathbb{R})$$

The last step is to use the vector equations (11) to write the standard basis vectors $\{\mathbf{e}_1, \mathbf{e}_2\}$ in terms of the new basis $\{\mathbf{f}_1, \mathbf{f}_2\}$; then the action of P_1 in the standard basis is easily evaluated by applying (12). From (11) we get

$$\mathbf{e}_1 = \mathbf{f}_1 \qquad \text{and} \qquad \mathbf{e}_2 = \mathbf{f}_2 - \mathbf{f}_1$$

and then from (12),

$$\begin{aligned} P_1(\mathbf{e}_1) &= P_1(\mathbf{f}_1) \;=\; \mathbf{f}_1 \;=\; \mathbf{e}_1 \\ P_1(\mathbf{e}_2) &= P_1(\mathbf{f}_2 - \mathbf{f}_1) \;=\; -\mathbf{f}_1 \;=\; -\mathbf{e}_1 \end{aligned}$$

and similarly

$$\begin{aligned} P_2(\mathbf{e}_1) &= P_2(\mathbf{f}_1) \;=\; \mathbf{0} \\ P_2(\mathbf{e}_2) &= P_2(\mathbf{f}_2 - \mathbf{f}_1) \;=\; \mathbf{f}_2 \;=\; \mathbf{e}_1 + \mathbf{e}_2 \end{aligned}$$

The projections P_i can now be re-written in terms of the Cartesian coordinates in $\mathbb{R}^2$ as

$$\begin{aligned} P_1(v_1, v_2) &= v_1 P_1(\mathbf{e}_1) + v_2 P_1(\mathbf{e}_2) \\ &= v_1 \cdot \mathbf{e}_1 - v_2 \cdot \mathbf{e}_1 \;=\; (v_1 - v_2, 0) \\ P_2(v_1, v_2) &= v_1 P_2(\mathbf{e}_1) + v_2 P_2(\mathbf{e}_2) \\ &= v_1 \cdot \mathbf{0} + v_2 \cdot (\mathbf{e}_1 + \mathbf{e}_2) \;=\; (v_2, v_2) \end{aligned}$$

It is interesting to calculate P_1^2, P_2^2 and $P_1 \circ P_2$ using the preceding formulas to verify the properties listed in Exercise 3.5 $\square$

The "idempotent property" $P^2 = P$ for a linear operator is characteristic of projections associated with a direct sum decomposition $V = V_1 \oplus V_2$. We have already seen that if $P, Q = (I - P)$ are the projections associated with such a decomposition, then

(i) $P^2 = P$ and $Q^2 = Q$ $\quad$ (ii) $PQ = QP = 0$ $\quad$ (iii) $P + Q = I$ (identity operator)

But the converse is also true.

3.7. Proposition. *If $P : V \to V$ is any linear operator such that $P^2 = P$, then V is a direct sum $V = R(P) \oplus K(P)$ and P is the projection of V onto the range $R(P)$, along the kernel $K(P)$. The operator $Q = I - P$ is also idempotent, with*

$$R(Q) = R(I - P) = K(P) \qquad \textit{and} \qquad K(Q) = K(I - P) = R(P)\ , \tag{13}$$

and projects V onto $R(Q) = K(P)$ along $K(Q) = R(P)$.

Proof: First observe that $Q = (I - P)$ is also idempotent since

$$(I - P)^2 = I - 2P + P^2 = I - 2P + P = (I - P)\ .$$

Next, note that

$$v \in K(Q) \;\Leftrightarrow\; Q(v) = (I - P)v = 0 \;\Leftrightarrow\; P(v) = v \;\Leftrightarrow\; v \in R(P)\,.$$

[Implication ($\Rightarrow$) in the last step is obvious. Conversely, if $v \in R(P)$ then $v = P(w)$ for some w and then $P(v) = P^2(w) = P(w) = v$, proving ($\Leftarrow$).] Thus

1. $K(Q) = K(I - P)$ is equal to $R(P)$
2. $R(Q) = R(I - P)$ is equal to $K(P)$,

proving (13).

Obviously $P + Q = I$ because $v = P(v) + (I - P)(v)$ implies $P(v) \in R(P)$, while $(I - P)(v) \in R(Q) = K(P)$ by (13); thus the span $R(P) + K(P) = R(P) + R(Q)$ is all of V. Furthermore $K(P) \cap R(P) = (0)$, for if v is in the intersection we have $v \in K(P) \Rightarrow P(v) = 0$. But we also have $v \in R(P)$, so $v = P(w)$ for some w, and then

$$0 = P(v) = P^2(w) = P(w) = v\ .$$

By Exercise 3.2 we conclude that $V = R(P) \oplus K(P)$.

Finally, let $\tilde{P}$ be the projection onto $R(P)$ along $K(P)$ associated with this decomposition; we claim that $\tilde{P} = P$. By definition $\tilde{P}$ maps $v = r + k \in R(P) \oplus K(P)$ to r; that, however, is exactly what our original operator P does:

$$P(r + k) = P(r) + P(k) = r + 0 = r\ .$$

Therefore $\tilde{P} = P$ as operators on V. $\square$

Direct Sums and Eigenspaces. Let $T : V \to V$ be a linear operator on a finite dimensional space V. As above, the spectrum of T is the set of *distinct* eigenvalues $\mathrm{sp}(T) = \{\lambda \in \mathbb{K} : E_\lambda \neq 0\}$, where E_λ is the (nontrivial) λ-eigenspace

$$E_\lambda = \{v \in V : T(v) = \lambda \cdot v\} = \ker(T - \lambda I) \qquad\qquad (I = \mathrm{id}_V)$$

3.8. Definition. *A linear operator $T : V \to V$ is* **diagonalizable** *if V is the direct sum of the nontrivial eigenspaces,*

$$V = \bigoplus_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$$

We will see below that this happens if and only if V has a basis $\mathbf{f}_1, \dots, \mathbf{f}_n$ of eigenvectors,

$$T(\mathbf{f}_i) \ = \ \mu_i \cdot \mathbf{f}_i \qquad \textit{for some } \mu_i \in \mathbb{K}$$

for $1 \leq i \leq n$.

Our next result shows that T is actually diagonalizable if we only know that the eigenspaces *span* V, with

$$V \ = \sum_{\lambda \in \mathrm{sp}(T)} E_\lambda(T) \ = \ \mathbb{K}\text{-span}\{E_\lambda : \lambda \in \mathrm{sp}_{\mathbb{K}}(T)\}$$

(a property much easier to verify).

3.9. Proposition. *If $T : V \to V$ is a linear operator on a finite dimensional space, let W be the span $\sum_{\lambda \in \mathrm{sp}_{\mathbb{K}}(T)} E_\lambda(T)$ of the eigenspaces. This space is T-invariant and is in fact a direct sum $W = \bigoplus_\lambda E_\lambda$ of the eigenspaces.*

Proof: Since each E_λ is invariant their span W is also T-invariant. The E_λ span W by hypothesis, so each $w \in W$ has some decomposition $w = \sum_\lambda w_\lambda$ with $w_\lambda \in E_\lambda$. For uniqueness of this decomposition it suffices to show that

$$0 = \sum_{\lambda \in \mathrm{sp}(T)} w_\lambda \text{ with } w_\lambda \in E_\lambda \quad \Rightarrow \quad \text{each } w_\lambda = 0\ .$$

The operators $T, (T - \lambda I)$, and $(T - \mu I)$ commute for all $\mu, \lambda \in \mathbb{K}$ since the identity element I and its scalar multiples commute with everybody. Let us fix an eigenvalue λ_0; we will show $w_{\lambda_0} = 0$. With this λ_0 in mind we define the product

$$A \ = \prod_{\lambda \neq \lambda_0, \lambda \in \mathrm{sp}(T)} (T - \lambda I)\ .$$

Then

$$0 = A(0) = A\Big(\sum_\lambda w_\lambda\Big) = \sum_\lambda A(w_\lambda) = A(w_{\lambda_0}) + \sum_{\lambda \neq \lambda_0} A(w_\lambda)$$

If $\lambda \neq \lambda_0$ we have

$$A(w_\lambda) = \left(\prod_{\mu\neq\lambda_0}(T-\mu I)\right)\,w_\lambda = \left(\prod_{\mu\neq\lambda_0,\lambda}(T-\mu I)\right)\cdot(T-\lambda I)\,w_\lambda = 0$$

because $w_\lambda \in E_\lambda$. On the other hand, by writing $(T-\lambda_0 I)+(\lambda_0-\mu)I$ we find that

$$A(w_{\lambda_0}) = \left(\prod_{\mu\neq\lambda_0}(T-\mu I)\right)w_{\lambda_0} \;=\; \left(\prod_{\mu\neq\lambda_0}(T-\lambda_0 I)+(\lambda_0-\mu)I\right)w_{\lambda_0}$$

When we expand this product of sums, every term but one includes a factor $(T-\lambda_0 I)$ that kills w_{λ_0}:

$$(\mathit{Term}) = (\mathit{product\ of\ operators})\cdot(T-\lambda_0 I)\,w_{\lambda_0} \;=\; 0$$

The one exception is the product

$$\left(\prod_{\mu\neq\lambda_0}(\lambda_0-\mu)\right)\cdot w_{\lambda_0}.$$

The scalar out front cannot be zero because each $\mu \neq \lambda_0$, so

$$A(w_{\lambda_0}) = \prod_{\mu\neq\lambda_0}(\lambda_0-\mu)\cdot w_{\lambda_0} \neq 0$$

But we already observed that

$$0 = A\Big(\sum_\lambda w_\lambda\Big) = 0 + A(w_{\lambda_0})$$

so we get a contradiction unless $w_{\lambda_0}=0$. Thus each term in $\sum_\lambda w_\lambda$ is zero and W is the *direct* sum of the eigenspaces E_λ. $\square$

If $W \subsetneq V$ this result tells us nothing about the behavior of T off of the subspace W, but if we list the distinct eigenvalues as $\lambda_1,\ldots,\lambda_r$ we can construct a basis for W that runs first through E_{λ_1}, then through E_{λ_2}, etc to get a basis for W,

$$\mathfrak{X} = \{f_1^{(1)},\ldots,f_{d_1}^{(1)}; f_1^{(2)},\ldots,f_{d_2}^{(2)};\ldots; f_1^{(r)},\ldots,f_{d_r}^{(r)}\}$$

where $d_i = \dim(E_{\lambda_i})$ and $\sum_i d_i = m = \dim(W)$. The corresponding matrix describing $T|_W$ is *diagonal*, so $T|_W$ is a diagonalizable operator on W even if T is not diagonalizable on all of V.

$$[T|_W]_{\mathfrak{X},\mathfrak{X}} \;=\; \begin{pmatrix} \lambda_1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & \\ & & & & & \ddots & \\ 0 & & & & & & \lambda_r \end{pmatrix} \tag{14}$$

where $\mathrm{sp}(T) = \{\lambda_1,\ldots,\lambda_r\}$.

3.10. Exercise. If a linear operator $T: V\to V$ acts on a finite dimensional space, prove that the following statements are equivalent.

1. T is diagonalizable: $V = \bigoplus_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$
2. There is a basis $f_1, \ldots, f_n$ for V such that each f_i is an eigenvector, with $Tf_i = \mu_i f_i$ for some $\mu_i \in \mathbb{K}$.

II.4. Representing Linear Operators as Matrices.

Let $T : V \to W$ be a linear operator between finite dimensional vector spaces with $\dim(V) = m$, $\dim(W) = n$. An **ordered basis** in V is an ordered list $\mathfrak{X} = \{e_1, \ldots, e_n\}$ of vectors that are independent and span V; let $\mathfrak{Y} = \{f_1, ..., f_n\}$ be an ordered basis for the target space W.

The behavior of a linear map $T : V \to W$ is completely determined by what it does to the basis vectors in V because every $v = \sum_{i=1}^m c_i e_i$ (uniquely) and

$$T\big(\sum_{i=1} c_i e_i\big) = \sum_{i=1}^{n} c_i T(e_i)$$

Each image $T(e_i)$ can be expressed uniquely as a linear combination of vectors in the $\mathfrak{Y}$ basis,

$$T(e_i) = \sum_{j=1}^{n} t_{ji} f_j \qquad \text{for } \; 1 \le i \le m$$

yielding a system of $m = \dim(V)$ vector equations that tell us how to rewrite vectors in the $\mathfrak{X}$-basis in terms of vectors in the $\mathfrak{Y}$-basis

$$\begin{array}{rcl} T(e_1) & = & b_{11} f_1 + ... + b_{1n} f_n \\ & \vdots & \\ T(e_m) & = & b_{m1} f_1 + ... + b_{mn} f_n \end{array} \tag{15}$$

We *define* the **matrix of T with respect to the bases $\mathfrak{X}$, $\mathfrak{Y}$** to be the $n \times m$ matrix

$$[T]_{\mathfrak{Y}\mathfrak{X}} = [t_{ij}] \quad \text{(the } \textit{transpose} \text{ of the array of coefficients } B = [b_{ij}] \text{ in (15))}$$

Since $(B^{\mathrm{t}})_{k\ell} = B_{\ell,k}$ that means $t_{ij} = b_{ji}$; to put it differently, the entries t_{ij} in $[T]_{\mathfrak{Y}\mathfrak{X}}$ satisfy the following identities derived from (15)

$$T(e_i) = \sum_{j=1}^{n} t_{ji} f_j \qquad \text{or } 1 \le i \le m \tag{16}$$

Note carefully: the basis vector f_j in (16) is paired with t_{ji} and not t_{ij}.

The matrix description of $T : V \to W$ changes if we take different bases; nevertheless, the same operator T (which has a coordinate-independent existence) underlies all of these descriptions. One objective in analyzing T is to find bases that yield the simplest matrix descriptions. When $V = W$ the best possible outcome is of course a basis that diagonalizes T as in (14), but alas, not all operators are diagonalizable.

Another issue worth considering is the following: If T is the identity operator $I = \mathrm{id}$ on a vector space V, and we compute $[\mathrm{id}]_{\mathfrak{X}\mathfrak{X}}$, the outcome is the same for all bases $\mathfrak{X}$,

$$[\mathrm{id}]_{\mathfrak{X}\mathfrak{X}} = I_{n\times n} \qquad \text{(the } n \times n \text{ identity matrix)}$$

But there is no reason why we couldn't take different bases in the initial and final spaces (even if they are the same space), regarding $T = \mathrm{id}_V$ as a map from $(V, \mathfrak{X})$ to $(V, \mathfrak{Y})$. Then there are some surprises when you compute $[T]_{\mathfrak{X}\mathfrak{Y}}$.

4.1. Exercise. Let V be 2-dimensional coordinate space $\mathbb{R}^2$. Let $I : V \to V$ be identity map $I = \mathrm{id}_V$, but take different bases $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ and $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ in the initial and final spaces. Letting $\mathbf{e}_1, \mathbf{e}_2$ be the standard basis vectors in $\mathbb{R}^2$ and $\mathbf{f}_1 = (1, 0), \mathbf{f}_2 = (1, 2)$, compute the matrices

$$(i)\ \ [I]_{\mathfrak{X}\mathfrak{X}} \qquad (ii)\ \ [I]_{\mathfrak{Y}\mathfrak{X}} \qquad (iii)\ \ [I]_{\mathfrak{Y}\mathfrak{Y}} \qquad (iv)\ \ [I]_{\mathfrak{X}\mathfrak{Y}}$$

4.2. Exercise. If V, W are finite dimensional and $T : V \to W$ is linear, prove that there are always bases $\mathfrak{X}, \mathfrak{Y}$ and $\mathfrak{X}', \mathfrak{Y}'$ in V, W such that

$$(i)\ \ [I]_{\mathfrak{Y}\mathfrak{X}} = \begin{pmatrix} 0 & 0 \\ 0 & I_{r\times r} \end{pmatrix} \qquad (ii)\ \ [I]_{\mathfrak{Y}'\mathfrak{X}'} = \begin{pmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{pmatrix}$$

where $r = \mathrm{rk}(T) = \dim(\mathrm{range}(T))$ is the rank and $I_{r\times r}$ is the $r \times r$ identity matrix.
Hint: Finding a basis that produces (i) is fairly easy; part (ii) requires some thought about the order in which basis vectors are listed. Both matrices represent the same operator $T : V \to W$.

The matrix description of T could hardly be simpler than those in Exercise 4.2, but at the same time much information about T has been lost in allowing arbitrary unrelated bases in V and W. Most operators encode far more information than can be captured by the single number $\mathrm{rk}(T)$.

In addition to our description of a linear operator $T : V \to W$ as a matrix, we can also describe vectors $v \in V$, $w \in W$ as column matrices once bases $\mathfrak{X} = \{e_i\}$ and $\mathfrak{Y} = \{f_j\}$ are specified. The correspondence $\phi_{\mathfrak{X}} : V \to \mathbb{K}^m$ is a linear bijection (an isomorphism of vector spaces) defined by letting

$$\phi_{\mathfrak{X}}(v) = [v]_{\mathfrak{X}} = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} \qquad \text{if } v = \textstyle\sum_{i=1}^m v_i e_i \text{ (unique expansion)}$$

Similarly $\phi_{\mathfrak{Y}} : W \to \mathbb{K}^n$ is given by

$$\phi_{\mathfrak{Y}}(w) = [w]_{\mathfrak{Y}} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \qquad \text{if } w = \textstyle\sum_{j=1}^n w_j f_j \text{ (unique expansion)}$$

These coordinate descriptions of linear operators and vectors are closely related.

4.3. Proposition. *If $T : V \to W$ is a linear operator and $\mathfrak{X}$, $\mathfrak{Y}$ are bases in V, W then for all $v \in V$:*

$$\phi_{\mathfrak{Y}}(Tv) = [T]_{\mathfrak{Y}\mathfrak{X}} \cdot \phi_{\mathfrak{X}}(v)$$

or equivalently,

$$[T(v)]_{\mathfrak{Y}} = [T]_{\mathfrak{Y}\mathfrak{X}} \cdot [v]_{\mathfrak{X}} \qquad \big(\textit{an } (n \times m)\cdot(m \times 1) \textit{ matrix product}\big)$$

Thus the i^{th} *component $(Tv)_i$ of $\phi_{\mathfrak{Y}}(Tv)$ is given by the familiar formula*

$$(Tv)_i = \sum_{k=1}^m t_{ik} v_k \qquad \textit{for } 1 \leq i \leq n$$

if $v = \sum_{k=1}^m v_k e_k$ and $[T]_{\mathfrak{Y}\mathfrak{X}} = [t_{ij}]$.

Proof: If $T(e_i) = \sum_{j=1}^{n} t_{ji} f_j$ and $v = \sum_{k=1}^{m} v_k e_k$, then

$$T(v) \;=\; T\Big(\sum_{k=1}^{m} v_k e_k\Big) \;=\; \sum_{k=1}^{m} v_k T(e_k) \;=\; \sum_{k=1}^{m} v_k\Big(\sum_{j=1}^{n} t_{jk} f_j\Big) \;=\; \sum_{j}\Big(\sum_{k} t_{jk} v_k\Big) f_j$$

So, the i^{th} component $(T(v))_i$ of $\phi_{\mathfrak{Y}}\big(T(v)\big)$ is $\sum_{k=1}^{m} t_{ik} v_k$, as claimed. □

The natural linear maps $\phi_{\mathfrak{X}} : V \to \mathbb{K}^m$, $m = \dim(V)$, and $\phi_{\mathfrak{Y}} : W \to \mathbb{K}^n$, $n = \dim(W)$, are bijective isomorphisms. Therefore a unique linear map $\tilde{T} = \phi_{\mathfrak{Y}} \circ T \circ \phi_{\mathfrak{X}}^{-1}$ is induced from $\mathbb{K}^m \to \mathbb{K}^n$ that makes the following diagram commute.

(17)
$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \phi_{\mathfrak{X}} \downarrow & & \downarrow \phi_{\mathfrak{Y}} \\ \mathbb{K}^m & \xrightarrow{\tilde{T}} & \mathbb{K}^n \end{array}$$

Figure 2.3. The diagram commutes, with $\phi_{\mathfrak{Y}} \circ T = \tilde{T} \circ \phi_{\mathfrak{X}}$.

It follows from Proposition 4.3 that the map $\tilde{T}$ we get when $T : V \to W$ is transferred over to a map between coordinate spaces is precisely the multiplication operator $L_A : \mathbb{K}^m \to \mathbb{K}^n$, where

$$A = [T]_{\mathfrak{Y}\mathfrak{X}}$$

is the coordinate matrix that describes T as in (15) and (16); see also Exercise 4.13 below.

Once bases $\mathfrak{X}, \mathfrak{Y}$ are specified there is also a natural linear isomorphism between the space of linear operators $\mathrm{Hom}_{\mathbb{K}}(V, W)$ and the space of matrices $\mathrm{M}(n \times m, \mathbb{K})$

4.4. Lemma. *If $\mathfrak{X}$, $\mathfrak{Y}$ are bases for finite dimensional vector spaces V, W the map ϕ from $\mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{M}(n \times m, \mathbb{K})$ given by*

$$\phi(T) = [T]_{\mathfrak{Y}\mathfrak{X}}$$

is a linear bijection, so these vector spaces are isomorphic, and

$$\dim_{\mathbb{K}}\Big(\mathrm{Hom}_{\mathbb{K}}(V, W)\Big) = \dim_{\mathbb{K}}\Big(\mathrm{M}(n \times m, \mathbb{K})\Big) = m{\cdot}n$$

Proof: Linearity of ϕ follows because if $\mathfrak{X} = \{e_i\}$ and $\mathfrak{Y} = \{f_j\}$ we have

$$(\lambda{\cdot}T)(e_i) = \lambda{\cdot}(T(e_i)) = \lambda{\cdot}\Big(\sum_{j=1}^{n} t_{ji} f_j\Big) = \sum_{j=1}^{m} (\lambda t_{ji}) f_j$$

for $i \le i \le m$, which means that $[\lambda T]_{ij} = \lambda{\cdot}[T]_{ij}$. Similarly, if we write $[T_k]_{\mathfrak{Y}\mathfrak{X}} = \big[t_{ij}^{(k)}\big]$ for $k = 1, 2$ we get

$$(T_1 + T_2)(e_i) = T_1(e_i) + T_2(e_i) = \sum_j t_{ji}^{(1)} f_j + \sum_j t_{ji}^{(2)} f_j = \sum_j \big(t_{ji}^{(1)} + t_{ji}^{(2)}\big) f_j$$

So $[T_1 + T_2]_{ij} = [T_1]_{ij} + [T_2]_{ij}$, proving linearity of ϕ as a map from operators to matrices.

One-to-One: The map ϕ is one-to-one if $\ker(\phi) = (0)$ – i.e. if $\phi(T) = [T]_{\mathfrak{Y}\mathfrak{X}} = [0]$ then T is the zero operator on V. This is clear: If $t_{ji} = 0$ for all i, j then $T(e_i) = \sum_{j=1}^{m} t_{ji} e_j = 0$ for all i, and $T(v) = 0$ for all v because $\{e_i\}$ is a basis.

Surjective: To prove ϕ surjective: given an $n \times m$ matrix $A = [a_{ij}]$ we must produce a linear operator $T : V \to W$ and bases $\mathfrak{X}, \mathfrak{Y}$ such that $[T]_{\mathfrak{Y}\mathfrak{X}} = [a_{ij}]$. This can done by working the definition of $[T]_{\mathfrak{Y}\mathfrak{X}}$ backward: we saw earlier that there is a unique linear operator $T : V \to W$ such that $T(e_i) = \sum_{j=1}^{n} a_{ji} f_j$, because $\{e_i\} = \mathfrak{X}$ is a basis in V.

Then, by definition of $[T]_{\mathfrak{Y}\mathfrak{X}}$ as in (15) - (16) we have $t_{ij} = a_{ij}$. $\square$

When $V = W$ composition of operators $S \circ T$ makes sense and the space of linear operators $\mathrm{Hom}_{\mathbb{K}}(V, V)$ becomes a (noncommutative) *associative algebra*, with the identity operator $I = \mathrm{id}_V$ as the multiplicative identity element. The set of matrices $\mathrm{M}(n, \mathbb{K})$ is also an associative algebra, under matrix multiplication; its identity element is the $n \times n$ diagonal identity matrix $I_{n\times n} = \mathrm{diag}(1, 1, \dots, 1)$. These systems are "isomorphic" as associative algebras, as well as vector spaces, because the bijection $\phi : \mathrm{Hom}(V, V) \to \mathrm{M}(n, \mathbb{K})$ intertertwines the multiplication operations $(\circ)$ and $(\cdot)$.

4.5. Proposition. *The bijective linear map* $\phi : \mathrm{Hom}_{\mathbb{K}}(V, V) \to \mathrm{M}(n, \mathbb{K})$ *intertwines the product operations in these algebras:*

$$\phi(S \circ T) = \phi(S) \cdot \phi(T) \qquad \textit{for all } S, T \in \mathrm{Hom}(V, V) \tag{18}$$

Under the correspondence between operators and their matrix representations, this is equivalent to saying that

$$[S \circ T]_{\mathfrak{X}\mathfrak{X}} = [S]_{\mathfrak{X}\mathfrak{X}} \cdot [T]_{\mathfrak{X}\mathfrak{X}}$$

for every basis $\mathfrak{X}$ *in* V*, where we take matrix product on the right.*

This is a special case of a much more general result.

4.6. Proposition. *Let* $U \xrightarrow{T} V \xrightarrow{S} W$ *be linear maps and let* $\mathfrak{X} = \{u_i\}$, $\mathfrak{Y} = \{v_i\}$, $\mathfrak{Z} = \{w_i\}$ *be bases in* U, V, W. *Then the correspondence between operators and their matrix realizations is "covariant" in the sense that*

$$[S \circ T]_{\mathfrak{Z}\mathfrak{X}} \;=\; [S]_{\mathfrak{Z}\mathfrak{Y}} \cdot [T]_{\mathfrak{Y}\mathfrak{X}}$$

(*a matrix product of compatible non-square matrices*).

Proof: We have $S \circ T(u_i) = \sum_k (S \circ T)_{ki} w_k$ by definition, and also

$$\begin{aligned}
S \circ T(u_i) &= S\big(T(u_i)\big) \;=\; S\Big(\sum_j t_{ji} v_j\Big) \;=\; \sum_j t_{ji} S(v_j) \\
&= \sum_j \Big(\sum_k s_{kl} w_k\Big)\, t_{ji} \;=\; \sum_k \Big(\sum_j s_{kj} t_{ji}\Big)\, w_k \\
&= \sum_k \big([S][T]\big)_{ki} w_k \qquad \text{(definition of matrix product)}
\end{aligned}$$

Thus $[S \circ T]_{ki} = ([S][T])_{ki}$, for all i, k. $\square$

4.7. Exercise. The $n \times m$ matrices E_{ij} with a "1" in the (i, j) spot and zeros elsewhere, are a basis for matrix space $\mathrm{M}(n \times m, \mathbb{K})$ since $[a_{ij}] = \sum_{i,j} a_{ij} E_{ij}$. When $m = n$ the matrices E_{ij} have useful algebraic properties. Prove that:

(a) These matrices satsify the identities

$$E_{ij} E_{k\ell} = \delta_{jk} \cdot E_{i\ell}$$

where δ_{rs} is the **Kronecker delta symbol**, equal to 1 if $r = s$ and zero otherwise.

(b) The "diagonal" elements E_{ii} are projections, with $E_{ii}^2 = E_{ii}$.

(c) $E_{11} + \ldots + E_{nn} = I_{n\times n}$ (the identity matrix).

If $T : V \to W$ is an invertible (bijective) linear operator between finite dimensional spaces, then $\dim(V) = \dim(W) = n$ and the inverse map $T^{-1} : W \to V$ is also linear (recall Exercise II.2.3). From the definition of the inverse map T^{-1} we have

$$T^{-1} \circ T = \mathrm{id}_V \qquad \text{and} \qquad T \circ T^{-1} = \mathrm{id}_W \ ,$$

so each operator undoes the action of the other. For any bases $\mathfrak{X}, \mathfrak{Y}$ in V, W the corresponding matrix realizations of T and T^{-1} are inverses of each other too. To see why, first recall

$$[\mathrm{id}_V]_{\mathfrak{X}\mathfrak{X}} = I_{n\times n} \qquad \text{and} \qquad [\mathrm{id}_W]_{\mathfrak{Y}\mathfrak{Y}} = I_{n\times n} \ .$$

Then by Proposition 4.6,

$$[T^{-1}]_{\mathfrak{X}\mathfrak{Y}} \cdot [T]_{\mathfrak{Y}\mathfrak{X}} = I_{n\times n} \qquad \text{and} \qquad [T]_{\mathfrak{Y}\mathfrak{X}} \cdot [T^{-1}]_{\mathfrak{X}\mathfrak{Y}} = I_{n\times n} \ ,$$

which means that $[T^{-1}]_{\mathfrak{X}\mathfrak{Y}}$ is the inverse $[T]^{-1}_{\mathfrak{Y}\mathfrak{X}}$ of the matrix of T. When $V = W$ and there is just one basis $\mathfrak{X}$ and all this reduces to the simpler statement $[T^{-1}] = [T]^{-1}$.

4.8. Exercise. Explain why *isomorphic* vector space must have the same dimension, even if one of them is infinite dimensional.

4.9. Exercise. If $T : V \to W$ is an invertible linear operator, prove that $(T^{-1})^{-1} = T$.

4.10. Exercise. If $U \xrightarrow{T} V \xrightarrow{S} W$ are invertible linear operators, explain why $S \circ T : U \to W$ is invertible, with $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$. (Note the reversal of order.)

For $A \in \mathrm{M}(n, \mathbb{K})$, we have defined the linear operator $L_A : \mathbb{K}^n \to \mathbb{K}^n$ via $L_A(\mathbf{x}) = A{\cdot}\mathbf{x}$, regarding vectors $\mathbf{x}$ as $n \times 1$ column matrices.

4.11. Exercise. Prove that the correspondence $L : \mathrm{M}(n, \mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$ has the following algebraic properties.

1. $L_{A+B} = L_A + L_B$ and $L_{\lambda \cdot A} = \lambda{\cdot}L_A$ for all $\lambda \in \mathbb{K}$;
2. $L_{AB} = L_A \circ L_B$;
3. If $I = I_{n\times n}$ is the identity matrix, then $L_I = \mathrm{id}_{\mathbb{K}^n}$.
4. L_A is an invertible linear operator if and only if the matrix inverse A^{-1} exists in $\mathrm{M}(n, \mathbb{K})$, and then we have $(L_A)^{-1} = \big(L_{A^{-1}}\big)$.

4.12. Exercise. Explain why the correspondence $L : \mathrm{M}(n, \mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^n)$ is a linear bijection and an isomorphism between these associative algebras.

4.13. Exercise. If $A \in \mathrm{M}(n \times m, \mathbb{K})$ and $\mathfrak{X}$, $\mathfrak{Y}$ are the *standard bases* in coordinate spaces $\mathbb{K}^m, K^m$ prove that the matrix $B = \big[L_A\big]_{\mathfrak{Y},\mathfrak{X}}$ that describes $L_A : \mathbb{K}^m \to \mathbb{K}^n$ for this particular choice of bases is just the original matrix A
Note: Does this work for arbitrary bases in $\mathbb{K}^n$?

4.14. Exercise. Let $\mathcal{P} = \mathbb{K}[x]$ be the infinite dimensional space of polynomials over $\mathbb{K}$. Consider the linear operators

1. Derivative: $D(a_0 + a_1x + ... + a_nx^n) = a_1 + 2a_2x + ... + na_nx^{n-1}$;
2. Antiderivative: $A(a_0 + a_1x + ... + a_nx^n) = a_0x + \frac{1}{2}a_1x^2 + \ldots + \frac{1}{n+1}a_nx^{n+1}$

Show that $D \circ A = \mathrm{id}_{\mathcal{P}}$ but that $A \circ D \neq \mathrm{id}_{\mathcal{P}}$. (What is $A \circ D$?) Show that D is surjective and A is one-to-one, but $\ker(D) \neq (0)$ and the range $R(D) \neq \mathcal{P}$.

This behavior is possible only in an infinite dimensional space. We have already observed (recall Corollary II.1.6) that if finite dimensional spaces V, W *have the same dimension*, the following statements regarding a linear maop $T : V \to W$ are equivalent.

T is one-to-one $\qquad$ T is surjective $\qquad$ T is bijective

4.15. Exercise. If $A \in \mathrm{M}(n, \mathbb{K})$ define its **trace** to be

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii} \qquad \text{(sum of the diagonal entries)}$$

Show that, for any $A, B \in \mathrm{M}(n, \mathbb{K})$,

1. $\mathrm{Tr} : \mathrm{M}(n, \mathbb{K}) \to \mathbb{K}$ is a linear map.
2. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$;
3. If $B = SAS^{-1}$ for some invertible matrix $S \in \mathrm{M}(n, \mathbb{K})$ then $\mathrm{Tr}\big(SAS^{-1}\big) = \mathrm{Tr}(A)$.

Change of Basis and Similarity Transformations. If $T : V \to W$ is a linear map between finite dimensional spaces and $\mathfrak{X}, \mathfrak{X}' \subseteq V$ and $\mathfrak{Y}, \mathfrak{Y}' \subseteq W$ are different bases, it is important to understand how the matrix models $[T]_{\mathfrak{Y}\mathfrak{X}}$ and $[T]_{\mathfrak{Y}'\mathfrak{X}'}$ are related as we seek particular bases yielding simple descriptions of T. For instance if $T : V \to V$ we may ask if T is diagonalizable over $\mathbb{K}$: Is there a basis such that

$$[T]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} \lambda_1 & & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

(repeats allowed among the λ_i)? Not all operators are so nice, and if T is not diagonalizable we will eventually work out a satisfactory but more complicated "Plan B" for dealing with such operators. All this requires a clear understanding of how matrix descriptions behave under a "change of basis."

4.16. Theorem (Change of Basis). *Let $T : V \to V$ be a linear operator on a finite dimensional space and let id_V be the identity operator on V. If $\mathfrak{X} = \{e_1, ..., e_n\}$ and $\mathfrak{Y} = \{f_1, ..., f_n\}$ are bases in V, then*

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = [\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}} \cdot [T]_{\mathfrak{X}\mathfrak{X}} \cdot [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}} \tag{19}$$

Futhermore $[\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}$ and $[\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}}$ are inverses of each other.

Proof: Since $T = \mathrm{id}_V \circ T \circ \mathrm{id}_V : V \to V \to V$, repeated application of Proposition 4.3 yields (19), as in the following system of commuting diagrams.

$$\begin{array}{ccccccc} \mathbb{K}^n & \xrightarrow{[\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}} & \mathbb{K}^n & \xrightarrow{L_A} & \mathbb{K}^n & \xrightarrow{[\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}}} & \mathbb{K}^n \\ \uparrow & & \uparrow & & \uparrow & & \uparrow \\ (V, \mathfrak{Y}) & \xrightarrow{\mathrm{id}_V} & (V, \mathfrak{X}) & \xrightarrow{T} & (V, \mathfrak{X}) & \xrightarrow{\mathrm{id}_V} & (V, \mathfrak{Y}) \end{array}$$

where $A = [T]_{\mathfrak{X}\mathfrak{X}}$. Applying the same proposition to the maps $\mathrm{id}_V = \mathrm{id}_V \circ \mathrm{id}_V$ we get

$$I_{n\times n} = [\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{Y}} = [\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}} \cdot [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}$$

which proves $[\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}$ and $[\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}}$ are mutual inverses. $\square$

To summarize: there is a unique, invertible "transition matrix" $S \in \mathrm{M}(n, \mathbb{K})$ such that

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = S \cdot [T]_{\mathfrak{X}\mathfrak{X}} \cdot S^{-1}, \qquad \text{where} \qquad S = [\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}} \text{ and } S^{-1} = [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}} \tag{20}$$

If we have explicit vector equations expressing the $\mathfrak{Y}$-basis vectors in terms of the $\mathfrak{X}$-basis vectors, the matrix $S^{-1} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}$ can be written down immediately; then we can compute $S = \left(S^{-1}\right)^{-1}$ from it.

4.17. Definition (Similarity Transformations). *Two matrices A, B in $\mathrm{M}(n, \mathbb{K})$ are* **similar** *if there is an invertible matrix $S \in \mathrm{GL}(n, \mathbb{K})$ such that $B = SAS^{-1}$. The mapping $\sigma_S : \mathrm{M}(n, \mathbb{K}) \to \mathrm{M}(n, \mathbb{K})$ given by $\sigma_S(A) = SAS^{-1}$ is referred to as a* **similarity transformation** *of A. It is also referred to by algebraists as* **"conjugation"** *of arbitrary matrices A by an invertible matrix S.*

Each individual conjugation operator $\sigma_S(A) = SAS^{-1}$ is an *automorphism* of the associative matrix algebra – it is a bijection that respects all algebraic operations in $\mathrm{M}(n, \mathbb{K})$:

$$\begin{array}{rcll} \sigma_S(A \cdot B) & = & \sigma_S(A) \cdot \sigma_S(B) & \\ \sigma_S(A + B) & = & \sigma_S(A) + \sigma_S(B) & \\ \sigma_S(\lambda \cdot A) & = & \lambda \cdot \sigma_S(A) & \quad \text{for } \lambda \in \mathbb{K} \\ \sigma_S\big(I_{n\times n}\big) & = & I_{n\times n} & \end{array}$$

for all matrices A, B and all "conjugators" $S \in \mathrm{GL}(n, \mathbb{K})$. But there is even more to be said: the correspondence $\psi : S \to \sigma_S$ has important algebraic properties of its own,

$$\begin{array}{rcl} \sigma_{S_1 S_2} & = & \sigma_{S_1} \circ \sigma_{S_2} \qquad \text{for all invertible matrices } S_1, S_2 \\ \sigma_{I_{n\times n}} & = & \big(\text{the identity operator } \mathrm{id}_{\mathrm{M}} \text{ on matrix space } \mathrm{M} = \mathrm{M}(n, \mathbb{K})\,\big) \end{array}$$

from which we automatically conclude that

The operator $\sigma_{S^{-1}}$ is the inverse $\left(\sigma_S\right)^{-1}$ of conjugation by S.

Thus the conjugation operators $\{\sigma_S : S \in \mathrm{GL}\}$ form a *group* of automorphisms acting on the algebra of $n \times n$ matrices.

When a linear operator $T : V \to V$ is described with respect to different bases in V, the resulting matrices must be similar as in (20). The converse is also true: if $A = [T]_{\mathfrak{X}\mathfrak{X}}$ and $B = SAS^{-1}$ for some invertible matrix S, there is a basis $\mathfrak{Y}$ such that $B = [T]_{\mathfrak{Y}\mathfrak{Y}}$. Thus, the different matrix models of T corresponding to bases $\mathfrak{Y}$ other than $\mathfrak{X}$ are precisely the similarity transforms $\{S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1} : S \text{ is invertible}\}$.

4.18. Lemma. *If $T : V \to V$ is a linear operator on a finite dimensional vector space V and if $A = [T]_{\mathfrak{X}\mathfrak{X}}$ then a $n \times n$ matrix B is equal to $[T]_{\mathfrak{Y}\mathfrak{Y}}$ for some basis $\mathfrak{Y}$ if and only if $B = SAS^{-1}$ for some invertible matrix S.*

Proof: ($\Rightarrow$) follows from (20). For ($\Leftarrow$): since S is invertible it has a matrix inverse S^{-1}. (Later we will discuss effective methods to compute matrix inverses such as S^{-1}.) According to Theorem 4.16, what we need is a basis $\mathfrak{Y} = \{f_1, ..., f_n\}$ such that $[\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}} = S^{-1}$; then $[\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}} = (S^{-1})^{-1} = S$ and $B = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1}$. If we write $S^{-1} = [b_{ij}]$ and $S = [s_{ij}]$ the identity $S^{-1} = [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}$ means that

$$f_i = \mathrm{id}_V(f_i) = \sum_j b_{ji} e_j \qquad \text{for } 1 \leq i \leq n$$

where $\mathfrak{X} = \{e_1, ..., e_n\}$. This is the desired new basis $\mathfrak{Y} = \{f_j\}$. To see it is a basis, we have $\{f_j\} \subseteq \mathbb{K}\text{-span}\{e_1, ..., e_n\}$ by definition, but $\{e_i\}$ is in $\mathbb{K}\text{-span}\{f_j\}$ because the matrix $S^{-1} = [b_{ij}]$ is invertible; in fact, $S^{-1}S = I$ implies that $\sum_i b_{ji} s_{ik} = \delta_{jk}$ (Kronecker delta). Then

$$\sum_i s_{ik} f_i = \sum_i s_{ik}\Big(\sum_j b_{ji} e_j\Big) = \sum_j \Big(\sum_i b_{ji} s_{ik}\Big)\, e_j = \sum_j \delta_{jk} e_j = e_k$$

for $1 \leq k \leq n$, so $\{e_k\} \subseteq \mathbb{K}\text{-span}\{f_j\}$ as claimed. Therefore $\{e_1, ..., e_n\}$ and $\{f_1, ... f_n\}$ both span V, and because $\{e_i\}$ is already a basis $\{f_j\}$ must also be a basis. $\square$

The next example shows that it can be difficult to tell by inspection whether an operator T is diagonalizable.

4.19. Example. Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be the linear operator whose action on the standard basis $\mathfrak{X} = \{\mathbf{e}_1 = (1, 0), \mathbf{e}_2 = (0, 1)\}$ is

$$T(\mathbf{e}_1) = 4\mathbf{e}_1 \qquad T(\mathbf{e}_2) = -\mathbf{e}_2$$

Clearly T is diagonalized by the $\mathfrak{X}$-basis since

$$[T]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix}$$

Compute $[T]_{\mathfrak{Y}\mathfrak{Y}}$ for the basis

$$\mathbf{f}_1 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 + \mathbf{e}_2) \qquad \mathbf{f}_2 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 - \mathbf{e}_2)$$

Solution: We have $[T]_{\mathfrak{Y}\mathfrak{Y}} = S\,[T]_{\mathfrak{X}\mathfrak{X}} S^{-1}$ where $S = [\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}}$ and $S^{-1} = [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}$. This inverse can be computed easily from our definition of the vectors $\mathbf{f}_1$, $\mathbf{f}_2$:

$$\begin{aligned} \mathbf{f}_1 = \mathrm{id}(\mathbf{f}_1) &= \frac{1}{\sqrt{2}}\mathbf{e}_1 + \frac{1}{\sqrt{2}}\mathbf{e}_2 \\ \mathbf{f}_2 = \mathrm{id}(\mathbf{f}_2) &= \frac{1}{\sqrt{2}}\mathbf{e}_1 - \frac{1}{\sqrt{2}}\mathbf{e}_2 \end{aligned} \tag{21}$$

which implies

$$S^{-1} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

The inverse of this matrix (found by standard matrix algebra methods or simply by solving (21) for $\mathbf{e}_1$, $\mathbf{e}_2$ in terms of $\mathbf{f}_1$, $\mathbf{f}_2$) is

$$[\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} = (S^{-1})^{-1} = S = -\frac{1}{\sqrt{2}} \cdot \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

(Notice that $S = S^{-1}$; this is not usually the case.) Then we get

$$\begin{aligned} [T]_{\mathfrak{Y}\mathfrak{Y}} &= [S]\,[T]_{\mathfrak{X}\mathfrak{X}} S^{-1} = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}} \\ &= \tfrac{1}{2} \cdot \begin{pmatrix} 4 & -1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \tfrac{1}{2} \cdot \begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & \frac{5}{2} \\ \frac{5}{2} & \frac{3}{2} \end{pmatrix} \end{aligned}$$

Diagonalizability of T would not be at all apparent if we used the basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ to represent T. $\square$

4.20. Exercise. Compute the matrix $[T]_{\mathfrak{Y}\mathfrak{Y}}$ for the linear operator $T : \mathbb{C}^2 \to \mathbb{C}^2$ of the previous example for each of the folowing bases $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$:

1. $\begin{cases} \mathbf{f}_1 = \frac{1}{\sqrt{2}}(\mathbf{e}_1 + \mathbf{e}_2) \\ \mathbf{f}_2 = \frac{1}{\sqrt{2}}(-\mathbf{e}_1 + \mathbf{e}_2) \quad \big(\text{obtained by rotating the standard basis vectors by } \theta = +45°\big) \end{cases}$

2. $\begin{cases} \mathbf{f}_1 = \frac{\sqrt{3}}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2 \\ \mathbf{f}_2 = -\frac{1}{2}\mathbf{e}_1 + \frac{\sqrt{3}}{2}\mathbf{e}_2 \quad \big(\text{the standard basis vectors rotated by } \theta = +30°\big) \end{cases}$

3. $\begin{cases} \mathbf{f}_1 = \mathbf{e}_1 + i\mathbf{e}_2 \\ \mathbf{f}_2 = \mathbf{e}_1 - i\mathbf{e}_2 \quad \big(\text{ where } i = \sqrt{-1} \text{ in } \mathbb{C} \text{ and } V = \mathbb{C}^2\,\big) \end{cases}$

4.21. Exercise. Let $\mathcal{P}_n$= polynomials of degree $\leq n$. Let $D = d/dx : \mathcal{P}_n \to \mathcal{P}_n$, the formal derivative of a polynomial. Compute $[D]_{\mathfrak{X}\mathfrak{X}}$ with respect to the basis $\mathfrak{X} = \{1, x, ..., x^n\}$. Compute $[D^2]_{\mathfrak{X}\mathfrak{X}}$ and $[D^{n+1}]_{\mathfrak{X}\mathfrak{X}}$ too.

An **RST equivalence relation** on a set X is rule declaring certain points $x, y \in X$ to be "related" (and others not). Writing $x \underset{R}{\sim} y$ when the points are related, the phrase "RST" means the relation is

1. Reflexive: $x \underset{R}{\sim} x$ for all $x \in X$.
2. Symmetric: $x \underset{R}{\sim} y \Rightarrow y \underset{R}{\sim} x$.
3. Transitive: $x \underset{R}{\sim} y$ and $y \underset{R}{\sim} z \Rightarrow x \underset{R}{\sim} z$

For each $x \in X$ we can then define its *equivalence class*, the subset

$$[x]_R = \{y \in X : y \underset{R}{\sim} x\}$$

The RST property forces distinct equivalence classes to be disjoint, so the whole space X decomposes into a the disjoint union of these classes.

One example of an RST equivalence is "*congruence mod a fixed prime p*" in the set $X = \mathbb{Z}$,

$$k \sim \ell \Leftrightarrow k \equiv \ell \ (\text{mod } p) \Leftrightarrow \ k \text{ and } \ell \text{ differ by a multiple of } p$$

It is easily verified that this is an RST relation and that the equivalence class of an integer m is its (mod p) congruence class

$$[m] = m + p\mathbb{Z} = \{k \in \mathbb{Z} : k \equiv m \ (\text{mod p})\,\}$$

There are only finitely many distinct classes, namely $[0], [1], \ldots, [p-1]$, which are disjoint and fill $\mathbb{Z}$. The finite field $\mathbb{Z}_p$ is precisely this set of equivalence classes equipped with suitable $\oplus$ and $\odot$ operations inherited from the system of integers $(\mathbb{Z}, +, \cdot\,)$.

Similarity of matrices

$$A \underset{R}{\sim} B \Leftrightarrow B = SAS^{-1} \quad \text{for some invertible matrix} \quad S \in \mathrm{GL}(n, \mathbb{K}) \tag{22}$$

is an important example of an RST relation on matrix space $X = \mathrm{M}(n, \mathbb{K})$. The RST properties are easily verified.

4.22. Exercise. Prove that similarity of matrices (22) has each of the RST properties.

The equivalence classes partition $\mathrm{M}(n, \mathbb{K})$ into disjoint **"similarity classes"** (aka **"conjugacy classes"**). All the matrices $[T]_{\mathfrak{X}\mathfrak{X}}$ associated with a linear operator $T : V \to V$ constitute a single similarity class in matrix space – they are all the possible representations of T corresponding to different choice of bases in V – and different operators correspond to disjoint similarity classes in $\mathrm{M}(n, \mathbb{K})$.

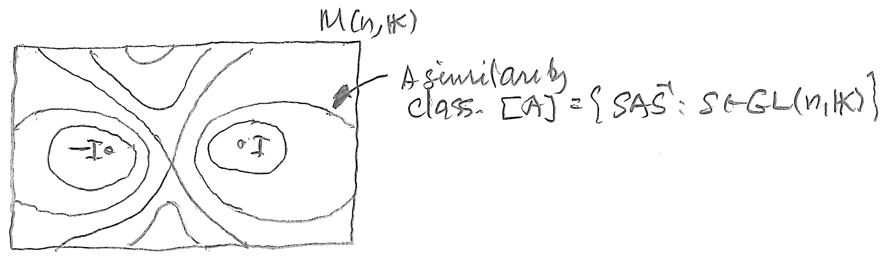

Figure 2.4. Similarity classes (= conjugacy classes) in $\mathrm{M}(n, \mathbb{K})$ partition matrix space into disjoint subsets $[A]$. Some classes are single points, for instance $[-I], [I]$, and $[0]$. Othere sre complicated hypersurfaces in $\mathbb{R}^{n^2} \cong \mathrm{M}(n, \mathbb{R})$. A similarity class could have seevral disconnected components.

The similarity classes don't all look the same. Some are trivial, consisting of a single point: for instance if

$$A = 0 \quad \text{or} \quad A = \lambda I_{n\times n} \ \ \text{(a scalar multiple of the identity matrix)}$$

we have

$$SAS^{-1} = \lambda \cdot SIS^{-1} = \lambda \cdot SS^{-1} = \lambda I = A \qquad \text{for all } S \in \mathrm{GL}(n, \mathbb{K})$$

The similarity class $[A]$ consists of the single point A. In particular, $[0] = \{0\}$, $[I] = \{I\}$ and $[-I] = \{-I\}$. When $\mathbb{K} = \mathbb{R}$ and we identify $\mathrm{M}(n, \mathbb{R})$ with $\mathbb{R}^{n^2}$, other similarity classes can be large curvilinear surfaces in Euclidean space. They can be quite a mess to compute.

4.23. Exercise. If $A \in \mathrm{M}(n, \mathbb{K})$ prove that

1. A commutes with *all* $n \times n$ matrices $\Leftrightarrow A = \lambda \cdot I_{n\times n}$, a scalar multiple if the identity matrix for some $\lambda \in \mathbb{K}$.
2. A commutes with all matrices in all *invertible* matrices $\mathrm{GL}(n, \mathbb{K}) = \{A : \det(A) \neq 0\} \Leftrightarrow A$ commutes with *all* $n \times n$ matrices, as in (1.)

Hint: Recall the matrices E_{ij} defined in Exercise 4.7, which are a basis for matrix space. In (2.), if $i \neq j$ then $I + E_{ij}$ is invertible (verify that $(I - E_{ij})$ is the inverse), and commutes with A. Hence E_{ij} commutes with A; we leave you to figure out what to do when $i = j$. If A commutes with all basis vectors E_{ij} it obviously commutes with all $n \times n$ matrices, and (1.) can be applied.

This shows that a similarity class $[A]$ in $\mathrm{M}(n, \mathbb{K})$ consists of a single point $\Leftrightarrow A = \lambda I$ (a *scalar matrix*).

4.24. Exercise. When we identify $\mathrm{M}(2, \mathbb{R}) \cong \mathbb{R}^4$ via the linear isomorphism

$$\mathbf{x} = \phi(A) = \big(a_{11}, a_{12}, a_{21}, a_{22}\big)\ ,$$

show that the similarity class $[A]$ of the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

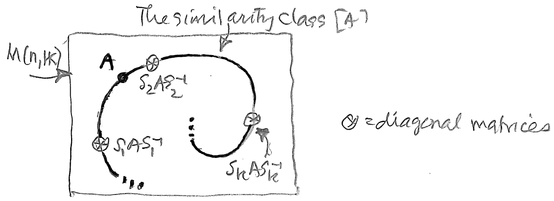

Figure 2.5. The diagonalization problem for $A \in \mathrm{M}(n, \mathbb{K})$ amounts to searching for one (or more) diagonal matrices lying in the similarity class $[A] = \{SAS^{-1} : S \in \mathrm{GL}(n, \mathbb{K})\}$.

is the 2-dimensional surface in $\mathbb{R}^4$ whose description in parametric form, described as the range of a polynomial map $\phi : \mathbb{R}^2 \to \mathrm{M}(2, \mathbb{R})$, is

$$[A(s,t)] = \left\{ \begin{pmatrix} 1 - st & s^2 \\ -t^2 & 1 + st \end{pmatrix} : s, t \in \mathbb{R} \ \text{ and } \ (s,t) \neq (0,0) \right\}$$

Note: A matrix $S = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible if the determinant $\det(S) = ad - bc$ is not 0, and then the inverse matrix is

$$S^{-1} = \frac{1}{\det(S)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \qquad \square$$

We will have a lot more to say about change of basis, similarity classes, and the diagonalization problem later on. Incidentally, not all matrices can be put into diagonal form by a similarity transformation. Our fondest hope is that in the equivalence class $[A]$ there will be at least one point SAS^{-1} that is diagonal (there may be several, as in Figure 2.5). If $A = [T]_{\mathfrak{X}\mathfrak{X}}$ for some linear operator $T : V \to V$ this is telling us which bases $\mathfrak{Y}$ make $[T]_{\mathfrak{Y}\mathfrak{Y}}$ diagonal, or whether there are any such bases at all.

4.25. Exercise. Suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear operator such that $T(\mathbf{e}_1) = 0$ and $T(\mathbf{e}_2) = \mathbf{e}_1$, so its matrix with respect to the standard basis $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ is $[T]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Prove that *no* basis $\mathfrak{Y} = \{f_1, f_2\}$ can make $[T]_{\mathfrak{Y}\mathfrak{Y}}$ diagonal.

Hint: $S = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is invertible $\Leftrightarrow$ the determinant $\det(S) = a_{11}a_{22} - a_{12}a_{21}$ is nonzero. We will eventually develop systematic methods to answer questions of this sort. For the moment, you will have to do it "bare-hands."

Chapter III. Dual Spaces and Duality.

III.1 Definitions and Examples.

The **linear functionals** (also known as **dual vectors**) on a vector space V over $\mathbb{K}$ are the linear maps $\ell : V \rightarrow \mathbb{K}$. We denote the space of functionals by V^*, or equivalently $\mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. It becomes a vector space when we impose the operations

1. Addition: $\big(\ell_1 + \ell_2\big)(v) = \ell_1(v) + \ell_2(v) \qquad$ for all $v \in V$
2. Scaling: $\big(\lambda\cdot\ell\big)(v) = \lambda\cdot\ell(v) \qquad$ for $\lambda \in \mathbb{K}, v \in V$

The zero element in V^* is the **zero functional** $\ell(v) = 0_{\mathbb{K}}$ for all $v \in V$, for which $\ker(\ell) = V$ and $\mathrm{range}(\ell) = \{0_{\mathbb{K}}\}$.

Notation: We will often employ "bracket" notation in discussing functionals, writing

$$\langle \ell, v \rangle \quad \text{instead of} \quad \ell(v)$$

This notation combines inputs ℓ, v to create a map $V^* \times V \rightarrow \mathbb{K}$ that is linear in each entry when the other entry is held fixed. In bracket notation both inputs play equal roles, and either one can be held fixed while the other varies. As we shall see this has many advantages. $\square$

We begin with an example that is central in understanding what dual vectors are and what they do.

1.1. Example. Let V be a finite dimensional space and $\mathfrak{X} = \{e_1, \ldots, e_n\}$ an ordered basis. Every $v \in V$ has a unique expansion

$$v = \sum_{i=1}^{n} c_i e_i \qquad (c_i \in \mathbb{K})$$

For each $1 \leq i \leq n$ the map $e_i^* : V \rightarrow \mathbb{K}$ that reads off the i^{th} coefficient

$$\langle e_i^*, v \rangle = c_i$$

is a linear functional in V^*. We will soon see that the set of functionals $\mathfrak{X}^* = \{e_1^*, \ldots, e_n^*\}$ is a basis for the dual space V^*, called the **dual basis** determined by $\mathfrak{X}$, from which it follows that the dual space is finite dimensional with $\dim(V^*) = \dim(V) = n$. $\square$

The following examples give some idea of the ubiquity of dual spaces in linear algebra.

1.2. Example. For $V = \mathbb{K}[x]$ an element $a \in \mathbb{K}$ determines an "evaluation functional" $\epsilon_a \in V^*$:

$$\langle \epsilon_a, f \rangle = \sum_{k=0}^{n} c_k a^k \qquad \text{if} \qquad f = \sum_{k=0}^{n} c_k x^k$$

These do not by themselves form a vector subspace of V^* because $\langle \epsilon_a - \epsilon_b, f \rangle = f(a) - f(b)$ cannot always be written as $f(c)$ for some $c \in \mathbb{K}$.

More generally, if $V = \mathcal{C}[a, b]$ is the space of *continuous* complex valued functions on the interval $X = [a, b] \subseteq \mathbb{R}$ we can define evaluation functionals $\langle \epsilon_s, f \rangle = f(s)$ for

$a \leq s \leq b$, but many element in V^* are of a quite different nature. Two examples:

$$\text{(i)} \qquad I(f) = \int_a^b f(t)\,dt \qquad \text{(Riemann integral of } f\text{)}$$

$$\text{(ii)} \qquad I^x(f) = \int_a^x f(t)\,dt \quad \text{(for any endpoint } a \leq x \leq b\text{)}$$

For another example, consider the space $V = \mathcal{C}^{(1)}(a,b)$ of real-valued functions on an interval $(a,b) \subseteq \mathbb{R}$ that have continuous first derivative $df/dx(s)$. We can define the usual evaluation functionals $\epsilon_s \in V^*$, but since differentiation is a linear operator on $\mathcal{C}^{(1)}(a,b)$ there are also functionals ℓ_s involving derivatives, such as

$$\ell_s : f \to \frac{df}{dx}(s) \qquad \text{for } a < s < b\ ,$$

or even linear combinations such as $\tilde{\ell}_s(f) = f(s) + \frac{df}{dx}(s)$. $\square$

1.3. Example. Suppose V is finite dimensional and that $l \in V^*$ is not the zero functional. The kernel $E = \ker(\ell) = \{v \in V : \langle \ell, v\rangle = 0\}$ is a "*hyperplane*" in V – a vector subspace of dimension $n-1$ where $n = \dim(V)$.

Proof: By the dimension formula,

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}\big(\ker(\ell)\big) + \dim_{\mathbb{K}}\big(\text{range}(\ell)\big)$$

But if $\ell \neq 0$, say $\langle \ell, v_0\rangle \neq 0$, then $\langle \ell, \mathbb{K}v_0\rangle = \mathbb{K}$, so range$(\ell) = \mathbb{K}$ has dimension 1. $\square$

1.4. Example. On $\mathbb{R}^n$ we have the standard Euclidean inner product

$$(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{n} x_k y_k \qquad \text{for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n\ ,$$

familiar from Calculus, but this is just a special case of the standard inner product on complex n-dimensional coordinate space $\mathbb{C}^n$,

$$(\mathbf{z}, \mathbf{w}) = \sum_{k=1}^{n} z_k \overline{w_k} \qquad \text{for complex } n\text{-tuples } \mathbf{z}, \mathbf{w} \text{ in } \mathbb{C}^n\ , \tag{23}$$

where $\overline{z} = x - iy$ is the complex conjugate of $z = x + iy$. We will focus on the complex case, because everything said here applies verbatim to the real case if you interpret "complex conjugation" to mean $\overline{x} = x$ for real numbers.

In either case, imposing an inner product on coordinate space $V = \mathbb{K}^n$ allows us to construct $\mathbb{K}$-linear functionals $\ell_{\mathbf{y}} \in V^*$ associated with individual vectors $\mathbf{y} \in V = \mathbb{K}^n$, by defining

$$\langle \ell_{\mathbf{y}}, \mathbf{x}\rangle = (\mathbf{x}, \mathbf{y}) \qquad \text{for any } \mathbf{x} \in V$$

In this setting the right hand vector $\mathbf{y}$ is fixed, and acts on the left-hand entry to produce a scalar in $\mathbb{K}$. (Think of $\mathbf{y}$ as the "actor" and $\mathbf{x}$ as the "actee" – the vector that gets acted upon.)

The functional $\ell_{\mathbf{y}}$ is $\mathbb{K}$-linear because the inner product is linear in its first entry when the second entry $\mathbf{y}$ is held fixed, hence $\ell_{\mathbf{y}}$ is a dual vector in V^*. Note carefully the placement of the "actee" on the left side of the inner product; the inner product on a vector space over $\mathbb{K} = \mathbb{C}$ is a *conjugate-linear* function of the right hand entry.

$$(\mathbf{z}, \lambda \cdot \mathbf{w}) = \overline{\lambda} \cdot (\mathbf{z}, \mathbf{w}) \quad \text{while} \quad (\lambda \cdot \mathbf{z}, \mathbf{w}) = \lambda \cdot (\mathbf{z}, \mathbf{w})$$

for $\lambda \in \mathbb{K}$. Placing the "actee" on the right would not produce a $\mathbb{C}$-*linear* operation on input vectors. (When $\mathbb{K} = \mathbb{R}$, complex conjugation doesn't do anything, and "conjugate-linear" is the same as "linear.")

The special case $\mathbb{K}^n = \mathbb{R}^n$ is of course important in geometry. The inner product on $\mathbb{R}^n$ and the functionals $\ell_{\mathbf{y}}$ then have explicit geometric interpretations:

$$\begin{aligned}
(\mathbf{x}, \mathbf{y}) &= \langle \ell_{\mathbf{y}}, \mathbf{x} \rangle = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cdot \cos(\theta) \\
&= \|\mathbf{x}\| \cdot \big(\|\mathbf{y}\| \cdot \cos\,\theta \big) \\
&= \|\mathbf{x}\| \cdot \left(\begin{array}{c} \text{orthogonally projected length of } \mathbf{y} \\ \text{on the 1-dimensional subspace } \mathbb{R}\cdot\mathbf{x} \end{array} \right) ,
\end{aligned}$$

where

$$\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2} = \Big(\sum_{k=1}^{n} |x_k|^2 \Big)^{1/2}$$

is the Euclidean length of vector $\mathbf{x} \in \mathbb{R}^n$. The angle $\theta = \theta(\mathbf{x}, \mathbf{y})$ is the angle in radians between $\mathbf{x}$ and $\mathbf{y}$, measured in the plane (two-dimensional subspace) spanned by $\mathbf{x}$ and $\mathbf{y}$ as shown in Figure 3.1. Notice that $\mathbf{x}$ and $\mathbf{y}$ are perpendicular if $(\mathbf{x}, \mathbf{y}) = 0$, so $\cos(\theta) = 0$.

Note: While the real inner product is natural in geometry, in physics the complex inner product is the notion of choice (in electrical engineering, quantum mechanics, etc, etc). *But beware*: physicists employ a convention opposite to ours. For them an inner product is linear in the *right-hand* entry and conjugate linear on the *left*. That can be confusing if you are not forwarned. $\square$

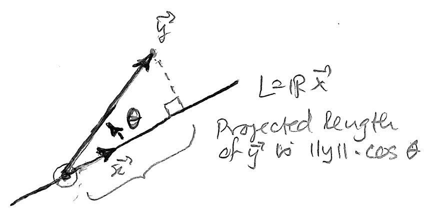

Figure 3.1. Geometric interpretation of the standard inner product $(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|\,\|\mathbf{y}\| \cdot \cos(\theta(\mathbf{x}, \mathbf{y}))$ in $\mathbb{R}^n$. The projected length of a vector $\mathbf{y}$ onto the line $L = \mathbb{R}\mathbf{x}$ is $\|\mathbf{y}\| \cdot \cos(\theta)$. The angle $\theta(\mathbf{x}, \mathbf{y})$ is measured within the two-dimensional subspace $M = \mathbb{R}\text{-span}\{\mathbf{x}, \mathbf{y}\}$. Vectors are *orthogonal* when $(\mathbf{x}, \mathbf{y}) = 0$, so $\cos\,\theta = 0$. The zero vector is orthogonal to everybody.

1.5. Example. In $V = \mathbb{R}^3$ with the standard inner product $(\mathbf{x}, \mathbf{y}) = \sum_i x_i y_i$, fix a vector $\mathbf{u} \neq \mathbf{0}$. The set of vectors $M = \{\mathbf{x} \in \mathbb{R}^3 : (\mathbf{x}, \mathbf{u}) = 0\}$ is the hyperplane of vectors orthogonal to $\mathbf{u}$ – see Figure 3.2. As an example, if $\mathbf{u} = (1, 0, 0) \in \mathbb{R}^3$ and $\ell_{\mathbf{u}}(x) = (\mathbf{x}, \mathbf{u})$ as in Example 1.4, this orthogonal hyperplane coincides with the kernel of $\ell_{\mathbf{u}}$:

$$M = \ker\big(\ell_{\mathbf{u}}\big) = \{(x_1, x_2, 0) : x_1, x_2 \in \mathbb{R}\} = \mathbb{R}\text{-span}\{\mathbf{e}_1, \mathbf{e}_2\} \quad \square \,.$$

1.6. Exercise. If $\mathbf{u} \neq \mathbf{0}$ in an inner product space of dimension n, explain why the orthogonal complement

$$M = \big(\mathbb{R}\cdot\mathbf{u}\big)^{\perp} = \{\mathbf{x} : (\mathbf{x}, \mathbf{u}) = 0\}$$

is a subspace of dimension $n - 1$.
Hint: Reread Example 1.3.

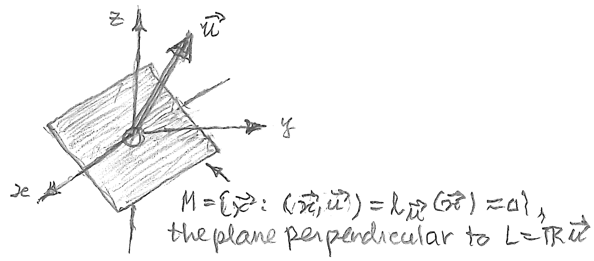

Figure 3.2. A nonzero vector u $\in \mathbb{R}^n$ determines a *hyperplane* $M = (\mathbb{R}\mathbf{u})^{\perp} = \{\mathbf{x} : (\mathbf{x}, \mathbf{y}) = 0\} = \ker(\ell_{\mathbf{u}})$, an $(n-1)$-dimensional subspace consisting of the vectors perpendicular to $\mathbf{u}$.

1.7. Example. Let $V = \mathcal{C}[0,1]$ be the ∞-dimensional space of all continuous complex-valued functions $f : [0,1] \rightarrow \mathbb{C}$. The **Fourier transform** of f is the function $f^\wedge : \mathbb{Z} \rightarrow \mathbb{C}$ defined by integrating $f(t)$ against the *complex trigonometric functions*

$$E_n(t) = e^{2\pi i n t} = \cos(2\pi t) + i\sin(2\pi t) \qquad (n \in \mathbb{Z})$$

on the real line. The n^{th} **Fourier coefficient** of $f(t)$ is the integral:

$$f^\wedge(n) = \int_0^1 f(t) e^{-2\pi i n t}\, dt \qquad (n \in \mathbb{Z})$$

(Note that the E_n are all periodic with period $\Delta t = 1$, so this integral is taken over the basic period $0 \leq t \leq 1$ common to them all.) If $f(t)$ is smooth and periodic with $f(t+1) = f(t)$ for all $t \in \mathbb{R}$, it can be synthesized as a superposition of the basic complex trigonometric functions E_n, with weights given by the Fourier coefficients:

$$f(t) = \sum_{n=-\infty}^{+\infty} f^\wedge(n) \cdot e^{2\pi i n t} = \sum_{n=-\infty}^{+\infty} f^\wedge(n) \cdot E_n(t)$$

The series converges pointwise on $\mathbb{R}$ if f is periodic and once continuously differentiable.

For each index $n \in \mathbb{Z}$ the map

$$f \in \mathcal{C}[0,1] \quad \xrightarrow{\phi_n} \quad f^\wedge(n) \in \mathbb{C}$$

is a linear functional in V^*. It is actually another example of a functional determined via an inner product as in Example 1.4. The standard inner product on $\mathcal{C}[0,1]$ is $(f,h) = \int_0^1 f(t)\overline{h(t)}\, dt$, and we have

$$\phi_n(f) = f^\wedge(n) = \int_0^1 f(t)\overline{E_n(t)}\, dt = \big(f, E_n\big)$$

for all $n \in \mathbb{Z}, f \in V$. So, ϕ_n is precisely the functional ℓ_{E_n} in Example 1.4. $\square$

III.2. Dual Bases in $\mathbf{V}^*$.

The dual space V^* of linear functionals can be viewed as the space of linear operators $\mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. For arbitrary vector spaces V, W of dimension m, n we saw earlier in Lemma 4.4 of Chapter II that $\mathrm{Hom}_{\mathbb{K}}(V, W)$ is isomorphic to the space $\mathrm{M}(n \times m, \mathbb{K})$ of $n \times m$

matrices, which obviously has dimension $m{\cdot}n$. In the special case when $W = \mathbb{K}$ we get $\dim(V^*) = m = \dim(V)$.

This can also be seen by re-examining Example 1.1, which provides a natural way to construct a basis $\mathfrak{X}^* = \{e_1^*, \dots, e_n^*\}$ in V^*, given an ordered basis $\mathfrak{X} = \{e_1, \dots, e_n\}$ in V. The functional e_i^* reads the i^{th} coefficent in the unique expansion $v = \sum_i c_i e_i$ of a vector $v \in V$, so that

$$\Big\langle e_i^*, \sum_{k=1}^{n} c_k e_k \Big\rangle = c_i \qquad \text{for } 1 \le i \le n \tag{24}$$

As an immediate consequence, the linear functional $e_i^* : V \to \mathbb{K}$ is completely determined by the property

$$\langle\, e_i^*, e_j \,\rangle = \delta_{ij} \qquad \text{(the Kronecker delta symbol} = 1 \text{ if } i = j \text{ and } 0 \text{ otherwise)} \tag{25}$$

Identity (25) follows because $e_j = 0 \cdot e_1 + \ldots + 1 \cdot e_j + \ldots + 0 \cdot e_n$; we recover (24) by observing that

$$\Big\langle e_i^*, \sum_{k=1}^{n} c_k e_k \Big\rangle = \sum_{k=1}^{n} c_k \langle e_i^*, e_k\rangle = \sum_{k=1}^{n} c_k \delta_{ik} = c_i$$

as expected.

We now show that the vectors $e_1^*, \dots, e_n^*$ form a basis in V^*, the **dual basis** to the original basis $\mathfrak{X}$ in V. This implies that $\dim(V^*) = \dim(V) = n$. Note, however, that to define the dual vectors e_i^* you must start with a basis in V; given a single vector "v" in V there is no way to define a dual vector "v^*" in V^*.

2.1. Theorem. *If V is finite dimensional and $\mathfrak{X}$ is a basis for V, the vectors $\mathfrak{X}^* = \{e_1^*, ..., e_n^*\}$ are a basis for V^*.*

Proof: Independence. If $\ell = \sum_{j=1}^n c_j e_j^*$ is the zero vector in V^* then $\langle \sum_j c_i e_j^*, v\rangle = 0$ for every $v \in V$, and in particular if $v = e_i$ we get

$$0 = \langle \ell, e_i\rangle = \sum_j c_j \langle e_j^*, e_i\rangle = \sum_j c_j \delta_{ji} = c_i$$

for $1 \le i \le n$, proving independence of the vectors e_i^*.

Spanning. If $\ell \in V^*$ and $c_i = \langle \ell, e_i\rangle$, we claim that ℓ is equal to $\ell' = \sum_{j=1}^n \langle \ell, e_j\rangle \cdot e_j^*$. It suffices to show that ℓ and ℓ' have the same values on the basis vectors $\{e_i\}$ in V, but that is obvious because

$$\begin{aligned} \langle \ell', e_i\rangle &= \Big\langle \sum_j \langle \ell, e_j\rangle e_j^* \,,\, e_i \Big\rangle \\ &= \sum_j \langle \ell, e_j\rangle \cdot \langle e_j^*, e_i\rangle \;=\; \sum_j \langle \ell, e_j\rangle \cdot \delta_{ij} \;=\; \langle \ell, e_i\rangle \end{aligned}$$

for $1 \le i \le n$ as claimed. $\square$

The formula developed in this proof is often useful in computing dual bases.

2.2. Corollary. *If V is finite dimensional, $\mathfrak{X} = \{e_i\}$ a basis in V, and $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis in V^*, then any $\ell \in V^*$ has*

$$\ell = \sum_{i=1}^{n} \langle \ell, e_i\rangle \cdot e_i^*$$

as its expansion in the $\mathfrak{X}^*$ basis.

2.3. Exercise. If $v_1 \neq v_2$ in a finite dimensional vector space V, prove that there is an $\ell \in V^*$ such that $\langle \ell, v_1 \rangle \neq \langle \ell, v_2 \rangle$. (Thus there are enough functionals in the dual V^* to distinguish vectors in V.)
Hint: It suffices to show $v_0 \neq 0 \Rightarrow \langle \ell, v_0 \rangle \neq 0$ for some $\ell \in V^*$. (Why?) Think about bases in V that involve v_0, and their duals.

Note: This result is actually true for all infinite dimensional spaces, but the proof is harder and requires "transcendental methods" involving the *Axiom of Choice*. These methods also show that every infinite dimensional space has a basis $\mathfrak{X}$ – an (infinite) set of independent vectors such that every $v \in V$ can be written as a *finite* $\mathbb{K}$-linear combination of vectors from $\mathfrak{X}$. As an example, the basic powers $\mathfrak{X} = \{1, x, x^2, \ldots\}$ are a basis for $\mathbb{K}[x]$ in this sense. A more challenging problem is to produce a basis for $V = \mathbb{R}$ when $\mathbb{R}$ is regarded as a vector space over the field of rationals $\mathbb{Q}$. Any such *Hamel basis* for $\mathbb{R}$ is necessarily uncountable. $\square$

2.4. Example. Consider the basis $\mathbf{u}_1 = (1,0,1)$, $\mathbf{u}_2 = (1,-1,0)$, $\mathbf{u}_3 = (2,0,-1)$ in $\mathbb{R}^3$. We shall determine the dual basis vectors $\mathbf{u}_i^*$ by computing their action as functionals on an arbitrary vector $v = (x_1, x_2, x_3)$ in $\mathbb{R}^3$.

Solution: Note that $(x_1, x_2, x_3) = \sum_{k=1}^3 x_k \mathbf{e}_k$ where $\{\mathbf{e}_k\}$ is the standard basis in $\mathbb{R}^3$. The basis $\{\mathbf{e}_k^*\}$ dual to the standard basis $\{\mathbf{e}_k\}$ has the following action:

$$\langle \mathbf{e}_k^*, (x_1, x_2, x_3) \rangle = \Big\langle \mathbf{e}_k^*, \sum_{i=1}^3 x_i \mathbf{e}_i \Big\rangle = x_k$$

because $\mathbf{e}_k^*$ reads the k^{th} coefficient in $v = \sum_i x_i \mathbf{e}_i$. For a different basis such as $\mathfrak{Y} = \{\mathbf{u}_i\}$, the dual vector $\mathbf{u}_k^*$ reads the k^{th} coefficient c_k when we expand a typical vector $v \in \mathbb{R}^3$ as $v = \sum_{j=1}^3 c_j \mathbf{u}_j$, so our task reduces to writing $v = (x_1, x_2, x_3) = \sum_{j=1}^3 x_i \mathbf{e}_j$ in terms of the new basis $\{\mathbf{u}_k\}$.

In matrix form, we have:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \sum_i c_i \mathbf{u}_i = c_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c_3 \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$$

To determine the coefficients c_k we must solve for C in the matrix equation

$$AC = X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \qquad \text{where} \qquad A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

Row operations on the augmented matrix for this system yield:

$$\begin{aligned} [A : X] \quad &= \quad \left(\begin{array}{ccc|c} 1 & 1 & 2 & x_1 \\ 0 & -1 & 0 & x_2 \\ 1 & 0 & -1 & x_3 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 1 & 2 & x_1 \\ 0 & 1 & 0 & -x_2 \\ 0 & -1 & -3 & x_3 - x_1 \end{array}\right) \\ &\to \left(\begin{array}{ccc|c} \boxed{1} & 1 & 2 & x_1 \\ 0 & \boxed{1} & 0 & -x_2 \\ 0 & 0 & \boxed{1} & \frac{1}{3}(x_1 + x_2 - x_3) \end{array}\right) \end{aligned}$$

There are no free variables; backsolving yields the unique solution

$$\begin{aligned} c_1 &= x_1 - c_2 - 2c_3 = \tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 + \tfrac{2}{3}x_3 \\ c_2 &= -x_2 \\ c_3 &= \tfrac{1}{3}(x_1 + x_2 - x_3) \end{aligned}$$

Thus,

$$v = \left(\tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 + \tfrac{2}{3}x_3\right)\mathbf{u}_1 - x_2\mathbf{u}_2 + \left(\tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 - \tfrac{1}{3}x_3\right)\mathbf{u}_3$$

Now read off the coefficients when $v = (x_1, x_2, x_3)$. Since $\langle \mathbf{u}_i^*, \mathbf{u}_j \rangle = \delta_{ij}$ we get

$$\begin{aligned}
\langle\, \mathbf{u}_i^*, v \,\rangle &= \langle\, \mathbf{u}_i^*, (x_1, x_2, x_3) \,\rangle \;=\; \Big\langle \mathbf{u}_i^*, \sum_i x_i \mathbf{e}_i \Big\rangle \\
&= \Big\langle \mathbf{u}_i^*, \sum_j c_j \mathbf{u}_j \Big\rangle \;=\; c_i \;=\; \begin{cases} \tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 + \tfrac{2}{3}x_3 & i = 1 \\ -x_2 & i = 2 \\ \tfrac{1}{3}x_1 + \tfrac{1}{3}x_2 - \tfrac{1}{3}x_3 & i = 3 \end{cases}
\end{aligned}$$

Since $\langle\, \mathbf{e}_k^*, (x_1, x_2, x_3) \,\rangle = x_k$ we can also rewrite this in the form

$$\begin{aligned}
\mathbf{u}_1^* &= \tfrac{1}{3}\mathbf{e}_1^* + \tfrac{1}{3}\mathbf{e}_2^* + \tfrac{2}{3}\mathbf{e}_3 \\
\mathbf{u}_2^* &= -\mathbf{e}_2^* \\
\mathbf{u}_3^* &= \tfrac{1}{3}\mathbf{e}_1^* + \tfrac{1}{3}\mathbf{e}_2^* - \tfrac{1}{3}\mathbf{e}_3^*
\end{aligned}$$

by Corollary 2.2. $\square$

III.3. The Transpose Operation.

There is a natural connection between linear operators $T : V \to W$ and operators in the opposite direction, from $W^* \to V^*$.

3.1. Theorem. *The* **transpose** $T^{\mathrm{t}} : W^* \to V^*$ *of a linear operator* $T : V \to W$ *between finite dimensional vector spaces is a linear operator that is uniquely determined in a coordinate-free manner by requiring that*

$$\langle\, T^t(\ell), v \rangle = \langle \ell, T(v) \rangle \qquad \text{for all } \ell \in W^*, v \in V \tag{26}$$

Proof: The right side of (26) defines a map $\phi_\ell : V \to \mathbb{K}$ such that $\phi_\ell(v) = \langle \ell, T(v) \rangle$. Observe that ϕ_ℓ is a linear functional on V (easily verified), so each $\ell \in W^*$ determines a well defined element of V^*. Now let $T^{\mathrm{t}} : W^* \to V^*$ be the map $T^{\mathrm{t}}(\ell) = \phi_\ell$. The property (26) holds by definition, but we must prove T^{t} is linear (and uniquely determined by the property (26)).

Uniqueness is easy: if $S : W^* \to V^*$ is another operator such that

$$\langle S(\ell), v \rangle = \langle \ell, T(v) \rangle = \langle T^{\mathrm{t}}\ell, v \rangle \qquad \text{for all } \ell \in W^* \text{ and } v \in V,$$

these identities imply $S(\ell) = T^{\mathrm{t}}(\ell)$ for all ℓ, which means $S = T^{\mathrm{t}}$ as maps on W^*.

The easiest proof that T^{t} is linear uses the scalar identities (26) and the following general observation.

3.2. Exercise. If V, W are finite dimensional vector spaces, explain why the following statements regarding two linear operators $A, B : V \to W$ are equivalent.

(a) $A = B$ as operators.

(b) $Av = Bv$ for all $v \in V$.

(c) $\langle \ell, Av \rangle = \langle \ell, Bv \rangle$ for all $v \in V, \ell \in W^*$.

Hint: Use Exercise 2.3 to prove (3.) $\Rightarrow$ (2.); implications (2.) $\Rightarrow$ (1.) $\Rightarrow$ (3.) are trivial.

To prove $T^{\mathrm{t}}(\ell_1 + \ell_2) = T^{\mathrm{t}}(\ell_1) + T^{\mathrm{t}}(\ell_2)$ just bracket these with an arbitrary $v \in V$ and compute:

$$\begin{aligned}
\langle T^{\mathrm{t}}(l_1 + l_2), v \rangle &= \langle l_1 + l_2, T(v) \rangle \\
&= \langle l_1, T(v) \rangle + \langle l_2, T(v) \rangle \qquad \text{(definition of } (+) \text{ in } W^*) \\
&= \langle T^{\mathrm{t}}(l_1) + T^{\mathrm{t}}(l_2), v \rangle \qquad \text{(definition of } (+) \text{ in } V^*)
\end{aligned}$$

for all $v \in V$. The other identity we need, $T^{\mathrm{t}}(\lambda\cdot\ell) = \lambda\cdot T^{\mathrm{t}}(\ell)$, is proved similarly. $\square$

Thus $T^{\mathrm{t}} : W^* \to V^*$ is a well-defined linear operator that acts in the opposite direction from $T : V \to W$.

Basic properties of the correspondence $T \to T^{\mathrm{t}}$ are left as exercises. The proofs are easy using the scalar identities (26).

3.3. Exercise. Verify that

(a) The transpose 0^{t} of the zero operator $0(v) \equiv 0_W$ from $V \to W$ is the zero operator from $W^* \to V^*$, so $0^{\mathrm{t}}(\ell) = 0_{V^*}$ for all $\ell \in W^*$.

1. When $V = W$ the transpose of the identity map $\mathrm{id}_V : V \to V$, with $\mathrm{id}_V(v) \equiv v$, is the identity map $\mathrm{id}_{V^*} : V^* \to V^*$ – in short, $\big(\mathrm{id}_V\big)^{\mathrm{t}} = \mathrm{id}_{V^*}$.
2. $(\lambda_1 T_1 + \lambda_2 T_2)^{\mathrm{t}} = \lambda_1 T_1^{\mathrm{t}} + \lambda_2 T_2^{\mathrm{t}}$, for any $\lambda_1, \lambda_2 \in \mathbb{K}$ and $T_1, T_2 : V \to W$.

3.4. Exercise. If $U \xrightarrow{T} V \xrightarrow{S} W$ are linear maps between finite dimensional vector spaces, prove that

$$\big(S \circ T\big)^{\mathrm{t}} = T^{\mathrm{t}} \circ S^{\mathrm{t}}$$

Note the reversal of order when we compute the transpose of a product.

3.5. Exercise. If V, W are finite dimensional and $T : V \to W$ is an invertible linear operator (a bijection), prove that $T^{\mathrm{t}} : W^* \to V^*$ is invertible too, and $\big(T^{-1}\big)^{\mathrm{t}} = \big(T^{\mathrm{t}}\big)^{-1}$ as maps from $V^* \to W^*$.

Now for some computational issues

3.6. Theorem. *Let $T : V \to W$ be a linear operator between finite dimensional spaces, let $\mathfrak{X} = \{v_1, ..., v_m\}$, $\mathfrak{Y} = \{w_1, ..., w_n\}$ be bases in V, W and let $\mathfrak{X}^* = \{v_i^*\}$, $\mathfrak{Y}^* = \{w_j^*\}$ be the dual bases in V^*, W^*. We have defined the transpose A^{t} of an $n \times m$ matrix to be the $m \times n$ matrix such that $(A^{\mathrm{t}})_{ij} = A_{ji}$. Then "$[T^{\mathrm{t}}] = [T]^{\mathrm{t}}$" in the sense that*

$$[T^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{Y}^*} = \big([T]_{\mathfrak{Y}\mathfrak{X}}\big)^{\mathrm{t}}$$

Important Note: This only works for the dual bases $\mathfrak{X}^*$, $\mathfrak{Y}^*$ in V^*, W^*. If $\mathfrak{A}, \mathfrak{B}$ are arbitrary bases in V^*, W^* unrelated to the dual bases there is no reason to expect that

$$[T^{\mathrm{t}}]_{\mathfrak{B}\mathfrak{A}} = \text{ the transpose of the matrix } [T]_{\mathfrak{Y}\mathfrak{X}} \qquad \square$$

Proof: To determine $[T^{\mathrm{t}}]$ we must calculate the coefficients in the system of vector equations

$$T^{\mathrm{t}}(w_i^*) = \sum_{j=1}^{m} [T^{\mathrm{t}}]_{ji} v_j^* \qquad 1 \le i \le n$$

These are easily found by applying each of these identities to a basis vector v_k in V:

$$\langle T^{\mathrm{t}}(w_i^*), v_k\rangle = \sum_{j=1}^{m} [T^{\mathrm{t}}]_{ji}\cdot\langle v_j^*, v_k\rangle = \sum_{j=1}^{m} [T^{\mathrm{t}}]_{ji}\delta_{jk} = [T^{\mathrm{t}}]_{ki} \tag{27}$$

for any $1 \le i \le n$ and $1 \le k \le m$. Thus

$$[T^{\mathrm{t}}]_{ki} = \langle T^{\mathrm{t}}(w_i^*), v_k\rangle$$

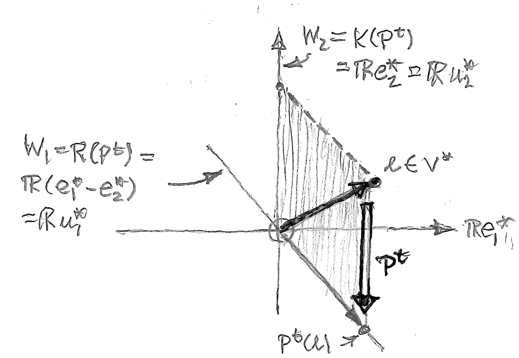

Figure 3.3. The decomposition $V = W_1 \oplus W_2$ determines the projection $P : \mathbb{R}^2 \to \mathbb{R}^2$ in Example 3.7 that maps V onto $W_1 = \mathbb{R}\mathbf{u}_1$ along $W_2 = \mathbb{R}\mathbf{u}_2$. We show that its transpose P^{t} projects V^* onto its range $R(P^{\mathrm{t}}) = \mathbb{R}\mathbf{u}_1^* = \mathbb{R}(\mathbf{e}_1^* - \mathbf{e}_2^*)$, along $K(P^{\mathrm{t}}) = \mathbb{R}\mathbf{u}_2^* = \mathbb{R}\mathbf{e}_2^*$. Horizontal axis in this picture is $\mathbb{R}\mathbf{e}_1^*$ and vertical axis is $\mathbb{R}\mathbf{e}_2^*$; a functional is then represented as $\ell = (\dot{x}_1\mathbf{e}_1^* + \dot{x}_2\mathbf{e}_2^*)$ with respect to the basis $\mathfrak{X}^* = \{\mathbf{e}_1^*, \mathbf{e}_2^*\}$ dual to the standard basis $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$.

By definition of T^{t} and the matrix $[T]_{\mathfrak{Y}\mathfrak{X}}$, we can also write (27) as

$$\begin{aligned}
\langle T^{\mathrm{t}}(w_i^*), v_k\rangle &= \langle w_i^*, T(v_k)\rangle = \Big\langle w_i^*, \sum_{j=1}^{n} [T]_{jk} w_j\Big\rangle \\
&= \sum_{j=1}^{j} [T]_{jk}\langle w_i^*, w_j\rangle = \sum_{j=1}^{n} [T]_{jk}\delta_{ij} = [T]_{ik}
\end{aligned}$$

for any $1 \le i \le n$, $1 \le k \le m$.

Upon comparison with (27) we conclude that $[T^{\mathrm{t}}]_{ki} = [T]_{ik} = \big([T]^{\mathrm{t}}\big)_{ki}$. Thus $[T^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{Y}^*}$ is the transpose of $[T]_{\mathfrak{Y}\mathfrak{X}}$. $\square$

3.7. Exercise (Computing Matrix Entries). If $T : V \to W$ and bases $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_i\}$ are given in V, W let $\mathfrak{X}^*$, $\mathfrak{Y}^*$ be the dual bases. Prove that

$$[T]_{\mathfrak{Y}\mathfrak{X}} = [t_{ij}] \text{ has entries } t_{ij} = \langle f_j^*, T(e_i)\rangle$$

The transpose of a projection $P : V \to V$ is a projection $P^{\mathrm{t}} : V^* \to V^*$ because $P^{\mathrm{t}} \circ P^{\mathrm{t}} = (P \circ P)^{\mathrm{t}} = P^{\mathrm{t}}$, so P^{t} maps V^* onto the range $R(P^{\mathrm{t}})$ along the nullspace $K(P^{\mathrm{t}}) = \ker(P^{\mathrm{t}})$ in the direct sum $V^* = R(P^{\mathrm{t}}) \oplus K(P^{\mathrm{t}})$. The following example shows how to calculate these geometric objects in terms of dual bases.

3.8. Example. Let $V = \mathbb{R}^2$ with basis $\mathfrak{Y} = \{\mathbf{u}_1, \mathbf{u}_2\}$ where $\mathbf{u}_1 = (1, 0)$, $\mathbf{u}_2 = (1, 1)$, and let P = projection onto $W_1 = \mathbb{R}\mathbf{u}_1$ along $W_2 = \mathbb{R}\mathbf{u}_2$. The standard basis $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ or the basis $\mathfrak{Y} = \{\mathbf{u}_1, \mathbf{u}_2\}$ can be used to describe P. The description with respect to $\mathfrak{X}$ has already been worked out in Example 3.6 of Chapter II.

1. Compute the dual bases $\mathfrak{X}^*$, $\mathfrak{Y}^*$ as functions $\ell : \mathbb{R}^2 \to \mathbb{R}$ and find the matrix descriptions of P^{t}:
$$[P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*} \qquad \text{and} \qquad [P^{\mathrm{t}}]_{\mathfrak{Y}^*\mathfrak{Y}^*}$$
2. Compute the kernel $K(P^{\mathrm{t}})$ and the range $R(P^{\mathrm{t}})$ in terms of the basis $\mathfrak{X}^*$ dual to the standard basis $\mathfrak{X}$.

3. Repeat (2.) for the basis $\mathfrak{Y}^*$.

Solution: First observe that

$$\begin{cases} \mathbf{u}_1 &= \mathbf{e}_1 \\ \mathbf{u}_2 &= \mathbf{e}_1 + \mathbf{e}_2 \end{cases} \quad \Rightarrow \quad \begin{cases} \mathbf{e}_1 &= \mathbf{u}_1 \\ \mathbf{e}_2 &= \mathbf{u}_2 - \mathbf{u}_1 \end{cases}$$

By definition, $\mathfrak{Y} = \{\mathbf{u}_1, \mathbf{u}_2\}$ is a diagonalizing basis for P, with

$$\begin{cases} P(\mathbf{u}_1) &= \mathbf{u}_1 \\ P(\mathbf{u}_2) &= \mathbf{0} \end{cases} \quad \text{which } \Rightarrow \; [P]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

We also have

$$\begin{cases} P(\mathbf{e}_1) &= P(\mathbf{u}_1) \,=\, \mathbf{u}_1 \,=\, \mathbf{e}_1 \\ P(\mathbf{e}_2) &= P(\mathbf{u}_2 - \mathbf{u}_1) \,=\, -\mathbf{u}_1 = -\mathbf{e}_1 \end{cases} \quad \text{which } \Rightarrow \; [P]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

The dual basis vectors are computed as functions $\mathbb{R}^2 \to \mathbb{R}$ by observing that

$$\begin{aligned} \mathbf{u}_1^*(v_1, v_2) &= \mathbf{u}_1^*(v_1\mathbf{e}_1 + v_2\mathbf{e}_2) \\ &= \mathbf{u}_1^*\big((v_1 - v_2)\mathbf{u}_1 + v_2\mathbf{u}_2\big) \qquad \text{which } \Rightarrow \mathbf{u}_1^* = \mathbf{e}_1^* - \mathbf{e}_2^* \\ &= v_1 - v_2 \,=\, (\mathbf{e}_1^* - \mathbf{e}_2^*)(v_1, v_2) \end{aligned}$$

and

$$\begin{aligned} \mathbf{u}_2^*(v_1, v_2) &= \mathbf{u}_2^*(v_1\mathbf{e}_1 + v_2\mathbf{e}_2) \\ &= \mathbf{u}_2^*\big((v_1 - v_2)\mathbf{u}_1 + v_2\mathbf{u}_2\big) \qquad \text{which } \Rightarrow \mathbf{u}_2^* = \mathbf{e}_2^* \\ &= v_2 \,=\, \mathbf{e}_2^*(v_1, v_2) \end{aligned}$$

No further calculations are needed to finish (1.), just apply Theorem 3.6 to get

$$[P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*} = \big([P]_{\mathfrak{X}\mathfrak{X}}\big)^{\mathrm{t}} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

Applying the same ideas we see that

$$[P^{\mathrm{t}}]_{\mathfrak{Y}^*\mathfrak{Y}^*} = \big([P]_{\mathfrak{Y}\mathfrak{Y}}\big)^{\mathrm{t}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^{\mathrm{t}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad .$$

That resolves Question 1.

For (2.), a functional $\ell = \dot{x}_1\mathbf{e}_1^* + \dot{x}_2\mathbf{e}_2^*$ ($\dot{x}_i \in \mathbb{R}$) is in $K(P^{\mathrm{t}}) \Leftrightarrow$

$$[P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \dot{x}_1 \\ -\dot{x}_1 \end{pmatrix} \text{ is equal to } \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

That happens $\Leftrightarrow \dot{x}_1 = 0$, so $K(P^{\mathrm{t}}) = \mathbb{R}\mathbf{e}_2^*$ with respect to the $\mathfrak{X}^*$ basis. Since we know $\mathbf{e}_2^* = \mathbf{u}_2^*$ we get $K(P^{\mathrm{t}}) = \mathbb{R}\mathbf{u}_2^*$ with respect to the $\mathfrak{Y}^*$ basis.

As for $R(P^{\mathrm{t}})$, if $\ell = b_1\mathbf{e}_1^* + b_2\mathbf{e}_2^*$ in the $\mathfrak{X}^*$-basis and we let $B = \mathrm{col}(b_1, b_2)$, we must solve the matrix equation $A\dot{X} = B$, where $A = [P^{\mathrm{t}}]_{\mathfrak{X}^*\mathfrak{X}^*}$. Row operations on $[A : B]$ yield

$$\left(\begin{array}{cc|c} 1 & 0 & b_1 \\ -1 & 0 & b_2 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 0 & b_2 + b_1 \end{array}\right)$$

so $B \in R(P^{\mathrm{t}}) \Leftrightarrow b_1 + b_2 = 0 \Leftrightarrow \ell \in \mathbb{R}\cdot(\mathbf{e}_1^* - \mathbf{e}_2^*)$. Thus $R(P^{\mathrm{t}}) = \mathbb{R}\cdot(\mathbf{e}_1^* - \mathbf{e}_2^*)$ in the $\mathfrak{X}^*$ basis, while in the $\mathfrak{Y}^*$ basis this becomes

$$R(P^{\mathrm{t}}) = \mathbb{R}\big(\mathbf{e}_1^* - \mathbf{e}_2^*\big) = \mathbb{R}\big((\mathbf{u}_2^* + \mathbf{u}_1^*) - \mathbf{u}_2^*\big) = \mathbb{R}\mathbf{u}_1^* \quad \square$$

The projection P^{t}, and the corresponding decomposition $V^* = R(P^{\mathrm{t}}) \oplus K(P^{\mathrm{t}})$, both have coordinate-independent geometric meaning. But the components of the direct sum have different descriptions according to which dual basis we use to describe vectors in V^*:

$$V^* = R(P^{\mathrm{t}}) \oplus K(P^{\mathrm{t}}) = \begin{cases} \mathbb{R}(\mathbf{u}_1^*) \oplus \mathbb{R}(\mathbf{u}_2^*) & \text{for the } \mathfrak{Y}^* \text{ basis} \\ \mathbb{R}(\mathbf{e}_1^* - \mathbf{e}_2^*) \oplus \mathbb{R}(\mathbf{e}_2^*) & \text{for the } \mathfrak{X}^* \text{ basis} \end{cases} \qquad \square$$

3.9. Exercise. Round out the previous discussion by verifying that

(a) For the standard basis $\mathfrak{X}$ we have $R(P) = \mathbb{R}\mathbf{e}_1$ and $K(P) = \mathbb{R}\cdot(\mathbf{e}_1 + \mathbf{e}_2)$.

(b) For the $\mathfrak{Y}$ basis we have $R(P) = \mathbb{R}\mathbf{u}_1$ and $K(P) = \mathbb{R}(\mathbf{u}_2)$.

3.9A. Exercise. Let $\mathfrak{X} = \{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ and $\mathfrak{Y} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be the standard bases in $V = \mathbb{K}^m, W = \mathbb{K}^n$. If S, T are linear maps such that

(a) $m \leq n$ and $S(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0)$

(b) $m \geq n$ and $T(x_1, \ldots, x_n, \ldots, x_m) = (x_1, \ldots, x_n)$

Compute the actions of S^{t} and T^{t} with respect to the dual bases $\mathfrak{X}^* = \{\mathbf{e}_1^*, \ldots, \mathbf{e}_m^*\}$ and $\mathfrak{Y}^* = \{\mathbf{f}_1^*, \ldots, \mathbf{f}_m^*\}$.

Reflexivity of Finite Dimensional Spaces. If V is finite dimensional there is a natural "bracketing map"

$$\phi : V^* \times V \to \mathbb{K} \qquad \text{given by} \qquad \phi : (\ell, v) \mapsto \langle \ell, v \rangle$$

The expression $\langle \ell, v \rangle$ is linear in each variable when the other is held fixed. If ℓ is fixed we get a linear functional $v \mapsto \ell(v)$ on V, but if we fix v the map $\ell \mapsto \langle \ell, v \rangle$ is a linear map from $V^* \to \mathbb{K}$, and hence is an element $j(v) \in V^{**} = (V^*)^*$, the "double dual" of V.

3.10. Lemma. *If* $\dim(V) < \infty$ *the map* $j : V \to V^{**}$ *is linear and a bijection. It is a "natural" isomorphism (defined without reference to any coordinate system) that allows us to identify* V^{**} *with* V.

Proof: For any $\ell \in V^*$ we have

$$\langle j(v_1 + v_2), \ell \rangle = \langle \ell, v_1 + v_2 \rangle = \langle \ell, v_1 \rangle + \langle \ell, v_2 \rangle = \langle j(v_1), \ell \rangle + \langle j(v_2), \ell \rangle$$

and similarly

$$\langle j(\lambda \cdot v), \ell \rangle = \langle \ell, \lambda v \rangle = \lambda \cdot \langle \ell, v \rangle = \langle \lambda \cdot j(v), \ell \rangle$$

Since these relations are true for all $\ell \in V^*$ we see that $j(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 j(v_1) + \lambda_2 j(v_2)$ in V^{**} and $j : V \to V^{**}$ is linear.

Finite dimensionality of V insures that $\dim(V^{**}) = \dim(V^*) = \dim(V)$, so j is a bijection $\Leftrightarrow$ j is onto $\Leftrightarrow$ j is one-to-one $\Leftrightarrow$ $\ker(j) = (0)$. But $j(v) = 0$ if and only if $0 = \langle j(v), \ell \rangle = \langle \ell, v \rangle$ for every $\ell \in V^*$. This forces $v = 0$ (and hence $\ker(j) = (0)$) because if $v \neq 0$ there is a functional $\ell \in V^*$ such that $\langle \ell, v \rangle \neq 0$. [In fact, we can extend $\{v\}$ to a basis $\{v, v_2, ..., v_n\}$ of V. Then, if we form the dual basis $\{v^*, v_2^*, ..., v_n^*\}$ in V^* we have $\langle v^*, v \rangle = 1$.] $\square$

There is, on the other hand, no *natural* (basis-independent) isomorphism from V to V^*. The spaces V and V^* are isomorphic because they have equal dimension, so there are many un-natural bijective linear maps between them. (We can create such a map given any basis $\{e_i\} \subseteq V$ and any basis $\{f_j\} \subseteq V^*$ by sending $e_i \to f_i$.)

If we identify $V = V^{**}$ via the natural map j, then the dual basis $(\mathfrak{X}^*)^*$ gets identitfied with the original basis $\mathfrak{X}$ in V. [Details: In fact, the vector e_i^{**} in the dual basis to $\mathfrak{X}^*$ coincides with the image vector $j(e_i)$ because

$$\langle j(e_i), e_j^* \rangle = \langle e_j^*, e_i \rangle = \delta_{ij} \ ,$$

which is the defining property of the vectors $(e_i^*)^*$ in $\mathfrak{X}^{**}$. Hence $j(e_i) = e_i^{**}$.] By time-honored abuse of notation mathematicians often write "$\mathfrak{X}^{**} = \mathfrak{X}$" even though this is not strictly true.

Furthermore when we identify $V^{**} \cong V$, the "double transpose" $T^{\mathrm{tt}} = (T^{\mathrm{t}})^{\mathrm{t}}$ mapping $V^{**} \rightarrow V^{**}$ becomes the original operator T, allowing us to write

$$T^{\mathrm{tt}} = T \qquad \text{(again, by abuse of notation)}$$

The precise connection between T and T^{tt} is shown in the following commutative diagram

$$\begin{array}{ccc} V^{**} & \xrightarrow{(T^{\mathrm{t}})^{\mathrm{t}}} & V^{**} \\ j^{-1} \downarrow\uparrow j & & \uparrow j \\ V & \xrightarrow{T} & V \end{array} \qquad \big(\text{Diagram commutes:} \quad T^{\mathrm{tt}} = j \circ T \circ j^{-1}\big)$$

3.11. Exercise. If $|V| = \dim V < \infty$, $\mathfrak{X} = \{e_1, ..., e_n\}$ is a basis, and $T : V \rightarrow V$ a linear operator,

(a) Fill in the details needed to show that the diagram above commutes,

$$(T^{\mathrm{t}})^{\mathrm{t}} \circ j = j \circ T$$

(b) Prove the following useful fact relating matrix realizations of T and T^{tt}

$$[T^{\mathrm{tt}}]_{\mathfrak{X}^{**}\mathfrak{X}^{**}} = [T]_{\mathfrak{X}\mathfrak{X}}$$

for the bases $\mathfrak{X}$ and $\mathfrak{X}^{**} = j(\mathfrak{X})$.

For infinite dimensional spaces there is still a natural linear embedding $j : V \rightarrow V^{**}$. Although j is again one-to-one, it is not necessarily onto and there is a chain of distinct dual spaces V, V^* , V^{**}, V^{***}, ... When $\dim(V) < \infty$, this process terminates with $V^{**} \cong V$. For this reason finite dimensional vector spaces are said to be "*reflexive.*" (Some infinite dimensional space are reflexive too, but not many.)

Annhilators. Additional structure must imposed on a vector space in order to speak of "lengths" or "orthogonality" of vectors, or the "orthogonal complement" $W^{\perp}$ of some subspace. When $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, this is most often accomplished by imposing an "inner product" $B : V \times V \rightarrow \mathbb{K}$ on the space. However, in the absence of such extra structure there is still a natural notion of a "complementary subspace" to any subspace $W \subseteq V$; but this complement

$$W^{\circ} = \{\ell \in V^* : \langle \ell, w \rangle = 0 \text{ for all } w \in W\} \qquad (\text{the } \textbf{annihilator of } W)$$

lives in V^* rather than V. It is easily seen that W° is a vector subspace in V^*. Obviously $(0)^{\circ} = V^*$ and $V^{\circ} = (0)$ in V^*, and when W is a proper subspace in V the annihilator W° is a proper subspace of V^*, with $(0) \subsetneqq W^{\circ} \subsetneqq V^*$.

3.12. Lemma. *Let V be finite dimensional and $W \subsetneqq V$ a subspace. If $v_0 \in V$ lies outside of W there is a functional $\ell \in V^*$ such that $\langle \ell, W \rangle = 0$ so $\ell \in W^{\circ}$ but $\langle \ell, v_0 \rangle \neq 0$.*

If $W \neq V$, so $r < n = \dim(V)$, the idea is to start with a basis $\{e_1, ..., e_r\}$ for W. Given a vector $v_0 \notin W$, adjoin additional vectors $e_{r+1} = v_0$, e_{r+2}, ..., e_n to make a basis $\mathfrak{X}$ for V. The dual basis $\mathfrak{X}^*$ provides the answer. We leave the details as an exercise.

We list the basic properties of annihilators as a series of exercises, some of which are major theorems (hints provided). In proving one of these results you may use any prior exercise or theorem. In all cases we assume $\dim(V) < \infty$.

3.13. Exercise. Let W be a subspace and $\mathfrak{X} = \{e_1, \ldots, e_r, \ldots, e_n\}$ a basis for V such that $\{e_1, \ldots, e_r\}$ is a basis in W. If $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis, prove that $\{e_{r+1}^*, \ldots, e_n^*\}$ is a basis for the annihilator $W^\circ \subseteq V^*$.

3.14. Exercise. (Dimension Theorem for Annihilators). If W is a subspace in a finite dimensional vector space V, prove that

$$\dim_{\mathbb{K}}(W) + \dim_{\mathbb{K}}(W^\circ) = \dim_{\mathbb{K}}(V) \quad , \tag{28}$$

or in abbreviated form, $|W| + |W^\circ| = |V|$.

3.15. Lemma. If $T : V \to W$ is a linear operator,

(a) Prove that
$$K(T^{\mathrm{t}}) = R(T)^\circ \qquad \big(\text{annihilator of the range } R(T)\big)$$

(b) Is it also true that $R(T^{\mathrm{t}}) = K(T)^\circ$? If not, what goes wrong?

3.16. Exercise. If V is finite dimensional, $T : V \to V$ is linear, and W a subspace of V, prove that W is T-invariant if and only if its annihilator W° is invariant under the transpose T^{t}.
Hint: Implication ($\Rightarrow$) is easy; in the other direction think about dual bases.

3.17. Exercise. If $T : V \to W$ is linear operator between finite dimensional vector spaces, prove that $\mathrm{rank}(T^{\mathrm{t}}) = \mathrm{rank}(T)$.

Recall that the rank of any linear operator $T : V \to W$ is the dimension $|R(T)| = \dim(R(T))$ of its range. If $A \in \mathrm{M}(n \times m, \mathbb{K})$ we defined $L_A : \mathbb{K}^m \to \mathbb{K}^n$ via $L_A(v) = A \cdot v$, for $v \in \mathbb{K}^m$, and the **rank of the matrix** is $\mathrm{rk}(A) = \dim\big(R(L_A)\big)$. Furthermore, recall that the "column rank" of a matrix is the dimension of its column space: $\mathrm{colrank}(A) = \dim\big(\mathrm{Col}(A)\big)$, and similarly $\mathrm{rowrank}(A) = \dim\big(\mathrm{Row}(A)\big)$. It is important to know that these numbers, computed in entirely different ways, are always equal – i.e.

"*row rank* = *column rank*" = $\mathrm{rk}(A)$ for any matrix,

regardless of its shape. The following exercises address this issue.

3.18. Exercise. Let $T : V \to W$ be a linear map between finite dimensional spaces, with bases $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_j\}$. If $A = [T]_{\mathfrak{Y}\mathfrak{X}}$ prove that

(a) The range $R(L_A)$ is equal to column space $\mathrm{Col}(A)$, hence
$$\mathrm{rk}(A) = \mathrm{rank}(L_A) = \dim(R(L_A)) = \dim\big(\mathrm{Col}(A)\big) = \mathrm{colrank}(A)$$
for any $n \times m$ matrix.

(b) If $A = [T]_{\mathfrak{Y},\mathfrak{X}}$ then $\mathrm{rank}(T) = \mathrm{rank}(L_A) = \mathrm{colrank}(A)$

Hint: For (b) recall the commutative diagram Figure 2.3 of Chapter II; the vertical maps are isomorphisms and isomorphisms preserve dimensions of subspaces.

3.19. Exercise. If A is an $n \times m$ matrix,

(a) Prove that $\text{rk}\big(A^{\text{t}}\big) = \text{rank}\big(L_{A^{\text{t}}}\big)$ is equal to $\text{rank}\big((L_A)^{\text{t}}\big)$.

(b) This would follow if it were true that "$(L_A)^{\text{t}} = L_{A^{\text{t}}}$." Explain why this statement does not make sense.

Hint: Keep in mind the setting for this (and the next) Exercise. If $V = \mathbb{K}^m, W = \mathbb{K}^n$, and A is $n \times m$ we get a map $L_A : \mathbb{K}^m \to \mathbb{K}^n$. The transpose A^{t} is $m \times n$ and determines a linear map in the opposite direction:

$$V \xrightarrow{L_A} W \qquad V \xleftarrow{L_{A^{\text{t}}}} W \qquad V^* \xleftarrow{(L_A)^{\text{t}}} W^* \ ,$$

while the transpose $\big(L_A\big)^{\text{t}}$ maps $W^* \to V^*$.

Use the results of the previous exercises to prove the main result below.

3.20. Exercise. If A is an $n \times m$ matrix, prove that

Theorem: *For any $n \times m$ matrix,* $\text{rowrank}(A) = \text{colrank}(A)$

Chapter IV. Determinants.

IV.1 The Permutation Group $\mathbf{S}_n$.

The **permutation group** S_n consists of all bijections $\sigma : [1,n] \to [1,n]$ where $[1,n] = \{1,...,n\}$, with composition of operators

$$\sigma_1 \circ \sigma_2(k) = \sigma_1(\sigma_2(k)) \qquad \text{for } 1 \leq k \leq n$$

as the group operation. The identity element e is the identity map $\mathrm{id}_{[1,n]}$ such that $e(k) = k$, for all $k \in [1,n]$. We recall that a **group** is any set G equipped with a binary operation $(*)$ satisfying the following axioms:

1. Associativity: $x * (y * z) = (x * y) * z$;
2. Identity element: There is an $e \in G$ such that $e * x = x = x * e$, for all $x \in G$;
3. Inverses: Every $x \in G$ has a "two-sided inverse," an element $x^{-1} \in G$ such that $x^{-1} * x = x * x^{-1} = e$.

We do not assume that the system $(G, *)$ is commutative, with $x*y = y*x$; a group with this extra property is a *commutative group*, also referred to as an *abelian group.* Here are some examples of familiar groups.

1. The integers $(\mathbb{Z}, +)$ become a commutative group when equipped with $(+)$ as the group operation; multiplication $(\cdot)$ *does not* make $\mathbb{Z}$ a group. (Why?)
2. Any vector space equipped with its $(+)$ operation is a commutatve group, for instance $(\mathbb{K}^n, +)$;
3. The set $(\mathbb{C}^\times, \cdot) = \mathbb{C} \sim \{0\}$ of nonzero complex numbers equipped with complex multiplication $(\cdot)$ is a commutative group. So is the subset $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ (unit circle in the complex plane) because $|z|, |w| = 1 \Rightarrow |zw| = |z|\cdot|w| = 1$ and $|1/z| = 1/|z| = 1$.
4. General Linear Group. The set $\mathrm{GL}(n,\mathbb{K}) = \{A \in \mathrm{M}(n,\mathbb{K}) : \det(A) \neq 0\}$ of *invertible* $n \times n$ matrices is a group when equipped with matrix multiply as the group operation. It is noncommutative when $n \geq 2$. Validity of the group axioms for $(\mathrm{GL}, \cdot)$ follows because

   $$\det(AB) = \det(A)\cdot\det(B) \qquad \det(I) = 1 \qquad \det(A^{-1}) = \frac{1}{\det(A)} \quad ,$$

   and a matrix A has a two-sided inverse $\Leftrightarrow \det(A) \neq 0$.
   Special Linear Group. These properties of the determinant imply that the subset $\mathrm{SL}(n,\mathbb{K}) = \{A \in \mathrm{M}(n,\mathbb{K}) : \det(A) = 1\}$ equipped with matrix multiply is also a (noncommutative) group;
5. The set of permutations $(\mathrm{Per}(X), \circ)$, all bijections on a set X of n distinct objects, is also a group when equipped with composition $(\circ)$ as its product operation. No matter what the nature of the objects being permuted, we can restrict attention to permutations of the set of integers $[1,n]$ by labeling the original objects, and then we have the group S_n.

Permuations. The simplest permutations are the k-**cycles**.

1.1. Definition. *An ordered list* $(i_1, ..., i_k)$ *of* k *distinct indices in* $[1,n] = \{1,...,n\}$ *determines a* k-**cycle** *in* S_n*, the permutation that acts in the following way on the set* $X = [1,n]$.

$$(29)\quad \sigma \textit{ maps } \begin{cases} i_1 \to i_2 \to \ldots \to i_k \to i_1 & (\textit{a one-step "cyclic shift" of list entries}) \\ j \to j & \textit{for all } j \textit{ not in the list } \{i_1, \ldots, i_k\} \end{cases}$$

A 1-cycle (k) *is just the identity map* id_X *so we seldom indicate them explicitly, though it is permissible and sometimes quite useful to do so. The* **support** *of a* k*-cycle is the set of entries* $\mathrm{supp}(\sigma) = \{i_1, \ldots, i_k\}$*, in no particular order. The support of a one-cycle* (k) *is the one-point set* $\{k\}$.

The order of the entries in the symbol $\sigma = (i_1, \ldots, i_k)$ matters, but cycle notation is ambiguous: k different symbols

$$(i_1, \ldots, i_k) = (i_2, \ldots, i_k, i_1) = (i_3, \ldots, i_k, i_1, i_2) = \ldots = (i_k, i_1, \ldots, i_{k-1})$$

obtained by "cyclic shifts" of the list entries in σ; all describe the same operation in S_n. Thus a k-cycle might best be descibed by a "cyclic list" of the sort shown below, rather than a linearly ordered list, but such diagrams are a bit cumbersome for the printed page. If we change the cyclic order of the indices we get a new operator. Thus $(1,2,3) = (2,3,1) = (3,1,2) \neq (1,3,2)$ because $(1,2,3)$ sends $1 \to 2$ while $(1,3,2)$ sends $1 \to 3$.

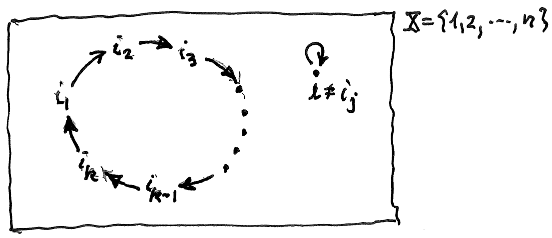

Figure 4.1. Action of the k-cycle $\sigma = (i_1, , \ldots, i_k)$ on $X = \{1, 2, \ldots n\}$. Points ℓ not in the "support set" $\mathrm{supp}(\sigma) = \{i_1, \ldots, i_k\}$ remain fixed; those in the support set are shifted one step clockwise in this cyclically ordered list. This σ is a "1-shift." (A 2-shift would move points 2 steps in the cyclic order, sending $i_1 \to i_3$ *to*$\ldots$ etc.

One (cumbersome) way to describe general elements $\sigma \in S_n$ employs a data arrray to show where each $k \in [1,n]$ ends up:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ j_1 & j_2 & j_3 & \ldots & j_n \end{pmatrix}$$

More efficient notation is afforded by the fact that every permutation σ can be uniquely written as a product of cycles with *disjoint* supports, which means that the factors commute.

1.2. Exercise. If $\sigma = (i_1, ..., i_r)$, $\tau = (j_1, .., j_s)$ act on *disjoint* sets of indices, show that these operators commute. This is no longer true if the sets of indices overlap. Check this by computing the effect of the following products $\sigma\tau(k) = \sigma(\tau(k))$ of permutations in S_5.

1. $(1,2,3)(2,4)$;

2. $(2,4)(1,2,3)$.

Is either product a cycle?

Thus the order of factors in a product of cycles is irrelevant if the cycles are disjoint.

The product of two cycles $\sigma\tau = \sigma \circ \tau$ is a composition of operators, so the action of $\sigma\tau = \sigma \circ \tau$ on an element $k \in [1, n]$ is evaluated by feeding k into the product *from the right* as below. Taking $\sigma = (1, 2)$, and $\tau = (1, 2, 3)$ in S_5 we have

$$\sigma\tau : k \xrightarrow{(1,2,3)} (1,2,3)\cdot k \xrightarrow{(1,2)} (1,2)\cdot\big((1,2,3)\cdot k\big) = \big((1,2)(1,2,3)\cdot k\big)$$

To determine the net effect we track what happens to each k:

$$\begin{array}{ccccc c c} & & \text{Action} & & & & \text{Net Effect} \\ 1 & \xrightarrow{(1,2,3)} & 2 & \xrightarrow{(1,2,3)} & 1 & & 1 \to 1 \\ 2 & \longrightarrow & 3 & \longrightarrow & 3 & & 2 \to 3 \\ 3 & \longrightarrow & 1 & \longrightarrow & 2 & & 3 \to 2 \\ 4 & \longrightarrow & 4 & \longrightarrow & 4 & & 4 \to 4 \\ 5 & \longrightarrow & 5 & \longrightarrow & 5 & & 5 \to 5 \end{array}$$

Thus the product $(1, 2)(1, 2, 3)$ is equal to $(2, 3) = (1)(2, 3)(4)(5)$, when we include redundant 1-cycles. On the other hand $(1, 2, 3)(1, 2) = (1, 3)$ which shows that cycles need not commute if their supports overlap. As another example we have

$$(1, 2, 3, 4)^2 = (1, 3)(2, 4)$$

which shows that a power σ^k of a cycle need not be a cycle, although it is a product of disjoint cycles. We cite without proof the fundamental cycle decomposition theorem.

1.3. Theorem (Cycle Decomposition of Permutations). *Every $\sigma \in S_n$ is a product of disjoint cycles. This decomposition is unique (up to order of the commuting factors) if we include the 1-cycles needed to account for all indices $k \in [1, n]$.*

1.4. Exercise. Write

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 5 & 1 & 3 \end{pmatrix}$$

as a product of disjoint commuting cycles.
Hint: Start by tracking $1 \to 2 \to 4 \to \ldots$ until a cycle is completed; then feed σ the first integer not included in the previous cycle, etc.

1.5. Exercise. Evaluate the net action of the following products of cycles

1. $(1, 2)(1, 3)$ in S_3;
2. $(1, 2)(1, 3)$ in S_6 ;
3. $(1, 2)(1, 2, 3, 4, 5)$ in S_5;
4. $(1, 2, 3, 4, 5)(1, 2)$ in S_5;
5. $(1, 2)^2$ in S_5;
6. $(1, 2, 3)^2$ in S_5.

Write each as a product of disjoint cycles.

1.6. Exercise. Determine the inverses σ^{-1} of the following elements in S_5

1. $(1, 2)$;
2. $(1, 2, 3)$;
3. Any 2-cycle (i_1, i_2) with $i_1 \neq i_2$;
4. Any k-cycle $(i_1, ..., i_k)$.

1.7. Exercise. Evaluate the following products in S_n as products of disjoint cycles

1. $(1, 5)(1, 4)(1, 3)(1, 2)$;

2. $(1,2)(1,3)(1,4)(1,5)$;

3. $(1,k)(1,2,...,k-1)$.

1.8. Exercise. The **order** $o(\sigma)$ of a permutation σ is the smallest integer $m \geq 1$ such that $\sigma^m = \sigma \cdot \ldots \cdot \sigma = e$.

1. Prove that every k-cycle has order $o(\sigma) = k$.

2. Verify that the r^{th} power σ^r of a k-cycle $\sigma = (i_1, \ldots, i_k)$ is an "r-shift" that moves every entry clockwise r steps in the cyclically ordered list of Figure 4.1.

3. If σ is a 6-cycle its square $\sigma^2 = \sigma \circ \sigma$ is a cyclic 2-shift of the entries $(i_1, \ldots, i_6)$. What is the order of this element in S_n?

Hint: By relabeling, it suffices to consider the standard 6-cycle $(1,2,3,4,5,6)$ in answering (3.)

The only element in S_n of order 1 is the identity e; two-cycles have order 2. As noted above, in (2.) the powers σ^r of a k-cycle need not be cycles (but sometimes they are).

Parity of a Permutation.

In a different direction we note that the 2-cycles (i,j) generate the entire group S_n in the sense that every $\sigma \in S_n$ can be written as a product $\sigma = \tau_1 \cdot \ldots \cdot \tau_r$ of 2-cycles. However these factors are not necessarily disjoint and need not commute, and such decompositions are far from unique since we have, for example,

$$e = (1,2)^2 = (1,2)^4 = (1,3)^2 \ \textit{etc.}.$$

Nevertheless an important aspect of such factorizations is unique, namely its **parity**

$$sgn(\sigma) = (-1)^r$$

where $r = \#(\text{2-cycles in the factorization } \sigma = \tau_1, \ldots, \tau_r)$. That means the elements $\sigma \in S_n$ fall into two disjoint classes: *even permutations* that can be written as a product of an even number of 2-cycles, and *odd permutations*. It is not obvious that all 2-cycle decompositions of a given permutation have the same parity. We prove that next, and then show how to compute $sgn(\sigma)$ effectively.

We first observe that a decomposition into 2-cycles always exists. By Theorem 1.3 it suffices to show that any k-cycle can be so decomposed. For 1-cycles this is obvious since $(k) = e = (1,2)\cdot(1,2)$. When $k > 1$ it is easy to check that

$$(1,2,...,k) = (1,k) \cdot \ldots \cdot (1,3)(1,2)$$

(with $k-1$ factors)

1.9. Exercise. Verify the preceding factorization of the cycle $(1,2,\ldots,k)$. Then by relabeling deduce that $(i_1,...,i_k) = (i_1,i_k)(i_1,i_{k-1}) \cdot \ldots \cdot (i_1,i_2)$ for any k-cycle.
Note: This is an example of "proof by relabeling."

Once we verify that the parity is well defined, this tell us how to recognize the parity of any k-cycle

$$sgn(i_1,i_2,...,i_k) = (-1)^{k-1} \qquad \text{for all } k > 0 \tag{30}$$

1.10. Theorem (Parity). *All decompositions $\sigma = \tau_1 \cdot \ldots \cdot \tau_r$ of a permutation as a product of 2-cycles have the same parity $sgn(\sigma) = (-1)^r$.*

Proof: The group S_n acts on the space of polynomials $\mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1,...,x_n]$ by permuting the variables

$$(\sigma \cdot f)(x_1,...,x_n) = f(x_{\sigma(1)},...,x_{\sigma(n)})$$

For instance $(1,2,3)\cdot f(x_1,x_2,x_3,x_4,x_5) = f(x_2,x_3,x_1,x_4,x_5)$. This is a "covariant group action" in the sense that

$$(\sigma\tau)\cdot f = \sigma\cdot\big(\tau\cdot f\big) \qquad \text{and} \qquad e\cdot f = f$$

for all f and all $\sigma,\tau \in S_n$. The notation makes this a bit tricky to prove; one way to convince yourself is to write

$$\begin{aligned}
\sigma\cdot(\tau\cdot f)(x_1,\ldots,x_n) &= \tau\cdot f\big(x_{\sigma(1)},\ldots,x_{\sigma(n)}\big) \\
&= \tau\cdot f(w_1,\ldots,w_n)\Big|_{w_1=x_{\sigma(1)},\ldots,w_n=x_{\sigma(n)}} \\
&= f(w_{\tau(1)},\ldots,w_{\tau(n)})\Big|_{w_k=x_{\sigma(k)}} \\
&= f\big(x_{\sigma(\tau(1))},\ldots,x_{\sigma(\tau(n))}\big) \\
&= f\big(x_{(\sigma\tau)(1)},\ldots,x_{(\sigma\tau)(n)}\big) \;=\; (\sigma\tau)\cdot f(x_1,\ldots,x_n)
\end{aligned}$$

Now consider the polynomial in n unknowns $\phi \in \mathbb{K}[x_1,...,x_n]$ given by

$$\phi(x_1,...,x_n) = \prod_{i<j}(x_i - x_j).$$

We claim that $\sigma\cdot\phi = (-1)\cdot\phi$ for any 2-cycle $\sigma = (i,j)$; by "covariance" it follows that $\sigma\cdot\phi = (-1)^r\phi$ if σ is a product $\tau_1\cdot\ldots\cdot\tau_r$ of r two-cycles. Since the definition of $\sigma\cdot\phi$ makes no reference to 2-cycle decompositions we will conclude that $(-1)^r$ must be the same for all such decompositions of σ, completing the proof.

To show that $\tau\cdot\phi = (-1)\phi$ for a 2-cycle (i,j) we may assume $i<j$. Note that the terms $x_k - x_\ell$ $(k<\ell)$ not involving i or j are unaffected when we switch $x_i \leftrightarrow x_j$. The remaining terms are of three types.

Case 1: Terms involving both i and j. The only such term is $x_i - x_j$ which becomes

$$\sigma\cdot(x_i - x_j) = x_j - x_i = (-1)(x_i - x_j)\ ,$$

suffering a change of sign.

Case 2: Terms involving only i. The possibilities (for $k \neq j$) are listed below

Terms	$x_k - x_i$ $1 \le k < i$	$x_i - x_k$ $i < k < j$	$x_i - x_k$ $j \le k \le n$
#(Terms)	$i-1$	$j-i-1$	$n-j$
Effect of $x_i \leftrightarrow x_j$ on sign of term	No change (since $k<i<j$)	$x_i - x_k \to x_j - x_k$ $= (-1)(x_k - x_j)$	No change (since $i<j<k$)

Case 3: Terms involving only j. These are (for $k \neq i$).

Terms	$x_k - x_j$ $1 \le k < i$	$x_k - x_j$ $i < k < j$	$x_j - x_k$ $j < k \le n$
#(Terms)	$j-1$	$j-i-1$	$n-j$
Effect of $x_i \leftrightarrow x_j$ on sign of term	No change (since $i<j<k$)	$x_k - x_j \to x_k - x_i$ $= (-1)(x_i - x_k)$	No change (since $i<j<k$)

The effect of switching $x_i \leftrightarrow x_j$ is to permute the terms in $\prod_{k<l}(x_k - x_l)$ changing the sign of some, so the product gets multiplied by $+1$ or -1. Counting the number of sign changes in all cases we see that

$$(-1)^{\#(\text{changes})} = (-1)^{1+\text{even}} = -1$$

as claimed. $\square$

1.11. Corollary. *The parity map $sgn : S_n \rightarrow \{\pm 1\}$, defined by $sgn(\sigma) = (-1)^r$ if σ can be written as a product of r two-cycles, has the following algebraic properties*

1. $sgn(e) = +1$;
2. $sgn(\sigma\tau) = sgn(\sigma) \cdot sgn(\tau)$;
3. $sgn(\sigma^{-1}) = \big(sgn(\sigma)\big)^{-1} = sgn(\sigma) \qquad (\textit{since } sgn = \pm 1)$.

Proof: Obviously $sgn(e) = 1$ since we may write $e = (1,2)^2$. If $\sigma = c_1 \cdot \ldots \cdot c_r$ and $\tau = c'_1 \cdot \ldots \cdot c'_s$ where c_i, c'_j are 2-cycles, then $\sigma\tau = c_1 \cdot \ldots \cdot c_r c'_1 \cdot \ldots \cdot c'_s$ is a product of $r + s$ cycles, proving (2.). The third property follows because

$$1 = sgn(e) = sgn(\sigma\sigma^{-1}) = sgn(\sigma) \cdot sgn(\sigma^{-1})$$

since the only values of sgn are ± 1. $\square$

IV.2 Determinants.

The previous digression about the permutation group S_n is needed to formulate the natural definition of $\det(A)$ for an $n \times n$ matrix $A \in \mathrm{M}(n, \mathbb{K})$, or of $\det(T)$ for a linear operator $T : V \rightarrow V$ on a finite dimensional vector space.

Any discussion that formulates this definition in terms of "expansion by minors" is confusing the natural definition of *det* with a commonly use *algorithm* for computing its value. Here is the real definition:

2.1. Definition. *If $A \in \mathrm{M}(n, \mathbb{K})$, we define its* **determinant** *to be*

$$\det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} = \sum_{\sigma \in S_n} sgn(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma(i)} \tag{31}$$

The products in this sum are obtained by taking $\sigma \in S_n$ and using it to select one entry from each row, taking each entry from a different column. Thus each σ determines a "template" for selecting matrix entries that are to be multiplied together (the product then weighted by the signature $sgn(\sigma)$ of the permutation). The idea is illustrated in Figure 4.2.

Many properties can be read directly from definition but the all-important multiplicative property $\det(AB) = \det(A) \cdot \det(B)$ is tricky no matter what definition we start from. We begin with several easy properties:

2.2. Theorem. *If $A \in \mathrm{M}(n, \mathbb{K})$ and $c \in \mathbb{K}$ we have*

1. $\det(I_{n\times n}) = 1$;
2. $\det(cA) = c^n \cdot \det(A)$ if A is $n \times n$;
3. $\det(A^{\mathrm{t}}) = \det(A)$;
4. *When $\mathbb{K} = \mathbb{C}$ we have $\det\big(\overline{A}\big) = \overline{\big(\det(A)\big)}$ where $\bar{z}$ = the complex conjugate of $z = x + iy$.*

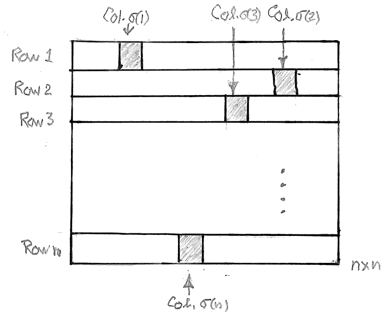

Figure 4.2. A permutation $\sigma \in S_n$ determines a "template" for selecting matrix entries by marking the address $(i, \sigma(i))$ – the one in Row_i, $\text{Column}_{\sigma(i)}$. Each row contains exactly one marked spot, and likewise for each column.

5. *If* A *is "upper triangular," so*

$$A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix},$$

then $\det(A) = \prod_{k=1}^{n} a_{kk}$ *is the product of the diagonal entries.*

Proof: Assertions (1.), (2.), (4.) are all trivial; we leave their proof to the reader. In (5.) the typical product $\pm a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)}$ in the definition of $\det(A)$ will equal 0 if any factor is zero. But unless $\sigma(k) = k$ for all k, there will be some row such that $\sigma(k) > k$ and some other row such that $\sigma(\ell) < \ell$. The resulting template includes a matrix entry below the diagonal, making the product for this template zero. The only permutation contributing a term to the sum (31) is $\sigma = e$, and that term is equal to $a_{11} \cdot \ldots \cdot a_{nn}$ as in (5.)

For (3.) we note that

$$\det(A^{\mathrm{t}}) = \sum_{\sigma \in S_n} sgn(\sigma)(b_{1,\sigma(1)} \cdot \ldots \cdot b_{n,\sigma(n)})$$

if $B = A^{\mathrm{t}} = [b_{ij}]$. By definition of A^{t}, $b_{ij} = a_{ji}$ so the typical term becomes

$$b_{1,\sigma(1)} \cdot \ldots \cdot b_{n,\sigma(n)} = a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n}$$

However, we may write $a_{\sigma(j),j} = a_{\sigma(j),\sigma^{-1}(\sigma(j))}$ for each j, and then

$$\det(A^{\mathrm{t}}) = \sum_{\sigma \in S_n} sgn(\sigma) b_{1,\sigma(1)} \cdot \ldots \cdot b_{n,\sigma(n)} = \sum_{\sigma \in S_n} sgn(\sigma) a_{\sigma(1),1} \cdot \ldots \cdot a_{\sigma(n),n}$$

Note that $\prod_i a_{\sigma(i),i} = \prod_{i=1}^{n} a_{\sigma(i),\sigma^{-1}(\sigma(i))}$, so if we replace the dummy index i in the product with $j = \sigma(i)$ the product becomes $\prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}$ and

$$\det(A^{\mathrm{t}}) = \sum_{\sigma \in S_n} sgn(\sigma) \cdot \prod_{j=1}^{n} a_{j,\sigma^{-1}(j)}.$$

Next, write $\tau = \sigma^{-1}$. The τ run through all of S_n as σ runs through S_n because S_n is a group. (This is our first encounter with the "group" property of S_n.) Furthermore

$sgn(\tau) = sgn(\sigma^{-1})$ so that

$$\det(A^{\mathrm{t}}) = \sum_{\tau \in S_n} sgn(\tau) \cdot \prod_{j=1}^{n} a_{j,\tau(j)} = \det(A) \quad \Box$$

The following observation will play a pivotal role in further discussion of determinants.

2.3. Lemma. *If B is obtained from A by interchanging two rows (or two columns) then*

$$\det(B) = (-1) \cdot \det(A) \ .$$

Proof: We do the case of column interchange. If $A = [a_{ij}]$ then $B = [b_{ij}]$ with $b_{ij} = a_{i\tau(j)}$; i.e. $\mathrm{Col}_j(B) = \mathrm{Col}_{\tau(j)}(A)$, for $1 \leq j \leq n$, where τ is the two-cycle $\tau = (k, \ell)$ that switches the column indices when we interchange $\mathrm{Col}_\ell(A) \leftrightarrow \mathrm{Col}_k(A)$. Then for any $\sigma \in S_m$, we have

$$b_{1,\sigma(1)} \cdot \ldots \cdot b_{n,\sigma(n)} = a_{1,\tau\sigma(1)} \cdot \ldots \cdot a_{n,\tau\sigma(n)}$$

But S_n is a *group* so $\tau S_n = S_n$ and the elements $\tau\sigma$ run through all of S_n as σ runs through S_n; furthermore, because τ is a 2-cycle we have $sgn(\tau) = -1$ and $sgn(\tau\sigma) = sgn(\tau) sgn(\sigma) = (-1) \cdot sgn(\sigma)$. Thus

$$\begin{aligned}
\det(B) \quad &= \quad \sum_{\sigma \in S_n} sgn(\sigma) \cdot \prod_{i=1}^{n} b_{i,\sigma(i)} \;=\; \sum_{\sigma \in S_n} sgn(\sigma) \cdot \prod_{i=1}^{n} a_{i,\tau\sigma(i)} \\
&= \quad \sum_{\sigma \in S_n} sgn(\tau)\, sgn(\tau\sigma) \cdot \prod_{i=1}^{n} a_{i,\tau\sigma(i)} \\
&= \quad sgn(\tau) \cdot \sum_{\mu \in S_n} sgn(\mu) \cdot \prod_{i=1}^{n} a_{i,\mu(i)} \;=\; (-1) \cdot \det(A) \qquad \Box
\end{aligned}$$

2.4. Exercise. Use the previous results to show that $\det(A) = 0$ if either:

1. A has two identical rows (or columns);
2. A has a row (or column) consisting entirely of zeros.

Recall the definition of the "elementary row operations" on a matrix A.

- Type I: $R_i \leftrightarrow R_j$: interchange Row_i and Row_j;
- Type II: $R_i \to \lambda \cdot R_i$: multiply Row_i by λ ($\lambda \in \mathbb{K}$);
- Type III: $R_i \to R_i + \lambda R_j$: Add to Row_i any scalar multiple of a different row R_j (leaving Row_j unaltered).

The effect of the first two operations on the determinant of a square matrix is easy to evaluate. We have just seen that Type I operations cause a sign change.

2.5. Exercise. Prove that if B has $R_i(B) = \lambda \cdot R_i(A)$ with all other rows unchanged, then $\det(B) = \lambda \cdot \det(A)$.

To deal with Type III operations we first observe that the map $det : \mathrm{M}(n, \mathbb{K}) \to \mathbb{K}$ is a **multilinear function** of the rows or columns of A.

2.6. Lemma. *If the* i^{th} *row of a matrix* A *is decomposed as a linear combination* $R_i = aR_i' + bR_i''$ *of two other rows of the same length, then*

$$\det(A) = \begin{pmatrix} R_1 \\ \vdots \\ aR_i' + bR_i'' \\ \vdots \\ R_n \end{pmatrix} = a \cdot \det \begin{pmatrix} R_1 \\ \vdots \\ R_i' \\ \vdots \\ R_n \end{pmatrix} + b \cdot \det \begin{pmatrix} R_1 \\ \vdots \\ R_i'' \\ \vdots \\ R_n \end{pmatrix} = a \cdot \det(A') + b \cdot \det(A'')$$

In other words $\det(A)$ is a multilinear function of its rows: If we vary only R_i holding the other rows fixed, the determinant is a linear function of R_i.

Proof: If $R_i' = (x_1, \ldots, x_n)$ and $R_i'' = (y_1, \ldots, y_n)$, then $A_{ij} = ax_j + by_j$ and

$$\begin{aligned} \det(A) &= \sum_{\sigma \in S_n} sgn(\sigma) \cdot \big(a_{1,\sigma(1)} \cdot \ldots \cdot (ax_{\sigma(i)} + by_{\sigma(i)}) \cdot \ldots \cdot a_{n,\sigma(n)}\big) \\ &= a \cdot \sum_{\sigma \in S_n} sgn(\sigma) \cdot \big(a_{1,\sigma(1)} \cdot \ldots \cdot x_{\sigma(i)} \cdot \ldots \cdot a_{n,\sigma(n)}\big) \\ &\qquad + b \cdot \sum_{\sigma \in S_n} sgn(\sigma) \cdot \big(a_{1,\sigma(1)} \cdot \ldots \cdot y_{\sigma(i)} \cdot \ldots \cdot a_{n,\sigma(n)}\big) \\ &= a \cdot \det(A') + b \det(A'') \end{aligned}$$

as claimed. $\square$

2.7. Corollary. *If* B *is obtained from* A *by a* Type III *row operation* $R_i \rightarrow R_i + cR_j$ $(j \neq i)$ *then* $Row_i(B) = R_i + cR_j$ *and*

$$\det(B) = \det \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} + c \cdot \det \begin{pmatrix} R_1 \\ \vdots \\ R_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{pmatrix} = \det(A) + 0 = \det(A)$$

because the second matrix has a repeated row.

Row Operations, Determinants, and Inverses.

Every row operation on an $n \times m$ matrix A can be implemented by multiplying A on the left by a suitable $n \times n$ "elementary matrix" E; the corresponding column operation is achieved by multiplying A on the right by the transpose E^{t}.

- Type I. $(Row_i) \rightarrow \lambda \cdot (Row_i)$: is equivalent to sending A to $E_{\text{I}}A$ where

$$E_{\text{I}} = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & \lambda & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix}$$

  Obviously $\det(E_{\text{I}}) = \lambda$ and

$$\det(E_{\text{I}}A) = \lambda \cdot \det(A) = \det(E_{\text{I}}) \cdot \det(A)$$

- Type II. $(Row_i) \leftrightarrow (Row_j)$: Now the result is achieved using the matrix

$$E_{\mathrm{II}} = \begin{array}{c} \begin{array}{ccccccc} & & \mathrm{Col}_i & & Col_j & & \end{array} \\ \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \boxed{0} & \cdots & \boxed{1} & & \\ & & \vdots & \ddots & \vdots & & \\ & & \boxed{1} & \cdots & \boxed{0} & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \end{array}$$

  Since E_{II} is $I_{n\times n}$ with two rows interchanged, $\det(E_{\mathrm{II}}) = -1$ and

$$\det(E_{\mathrm{II}}A) = (-1)\cdot\det(A) = \det(E_{\mathrm{II}})\cdot\det(A)$$

- Type III. $(Row_i) \rightarrow (Row_i) + \lambda(Row_j)$, with $j \neq i$. Assuming $i < j$, the appropriate matrix is

$$E_{\mathrm{III}} = \begin{array}{c} \begin{array}{ccccccc} & & \mathrm{Col}_i & & \mathrm{Col}_j & & \end{array} \\ \begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & \boxed{1} & \cdots & \boxed{\lambda} & & \\ & & \vdots & \ddots & \vdots & & \\ & & \boxed{0} & \cdots & \boxed{1} & & \\ & & & & & \ddots & \\ 0 & & & & & & 1 \end{pmatrix} \end{array}$$

  and we have $\det(E_{\mathrm{III}}) = 1$. But then by Lemma 2.6 we also have

$$\det(E_{\mathrm{III}}A) = \det(A) = \det(E_{\mathrm{III}})\cdot\det(A)$$

This proves:

2.8. Lemma. *If E is any $(n\times n)$ elementary matrix then*

$$\det(EA) = \det(E)\cdot\det(A)$$

for any $n\times n$ matrix A.

This allows us to compute determinants using row operations, exploiting the fact that $\det(A)$ can be calculated by inspection if A is upper triangular. First observe that the effect of a sequence of row operations is to map $A \mapsto E_m \cdot \ldots \cdot E_1 \cdot A$ (echelon form), but then

$$\det(E_m \cdot \ldots \cdot E_1 A) = \det(E_m)\cdot\det(E_{m-1}\cdot \ldots \cdot E_1 \cdot A) = \Big(\prod_{i=1}^{m}\det(E_i)\Big)\cdot\det(A)$$

Thus

$$\det(A) = \Big(\prod_{i=1}^{m}\det(E_i)^{-1}\Big)\cdot\det(E_1\cdot\ldots\cdot E_m A)$$

and calculating $\det(A)$ reduces to calculating the upper triangular row reduced form, whose determinant can be read by inspection. (You also have to keep track of the row operations used, and their determinants.)

Computing Inverses. Suitably chosen row operations will put an $n \times n$ matrix into echelon form; if we only allow elementary operations of Type II or Type III we can achieve nearly the same result, except that the pivot entries contain nonzero scalars λ_i rather than "1"s, as shown in Figure 4.3. Next recall that $\mathrm{M}(n, \mathbb{K})$ and the space of linear operators $\mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^n)$ are isomorphic as *associative algebras* under the correspondence

$$A \mapsto L_A \qquad \big(L_A(\mathbf{x}) = A{\cdot}\mathbf{x} = ((n \times n){\cdot}(n \times 1) \text{ matrix product} \big) \ ,$$

as we showed in the discussion surrounding Exercise 4.12 of Chapter II. That means the following statements are equivalent.

(32)

1. *A matrix inverse* A^{-1} *exists in* $\mathrm{M}(n, \mathbb{K})$;
2. $L_A : \mathbb{K}^n \to \mathbb{K}^n$ *is an invertible linear operator*;
3. $\ker(L_A) = (0)$;
4. *The matrix equation* $AX = 0$ *has only the trivial solution* $X = 0_{n\times 1}$.

We say that a matrix is **nonsingular** if any of these conditions holds; otherwise it is **singular**.

2.9. Exercise. If A, B are square matrices prove that

1. The product AB is *singular* if at least one of the factors is singular.
2. The product AB is *nonsingular* if both factors are nonsingular.

With this in mind we can deduce useful facts about matrix inverses from the preceding discussion of row operations and determinants.

2.10. Proposition. *The following statements regarding an* $n \times n$ *matrix are equivalent.*

1. $\det(A) \neq 0$;
2. *A has a multiplicative inverse* A^{-1} *in* $\mathrm{M}(n, \mathbb{K})$;
3. *The multiplication operator* $L_A : \mathbb{K}^n \to \mathbb{K}^n$ *is an invertible* (*bijective*) *linear operator on coordinate space.*

Proof: We already know (2.) $\Leftrightarrow$ (3.). Row operations of Type II and III reduce A to one of the two "modified echelon forms" A' (see Figure 4.3(a–b)), in which the step corners contain nonzero scalars $\lambda_1, \ldots, \lambda_r$ that need not equal 1, and $r = \mathrm{rank}(A)$. Obviously if there are columns that do not meet a step-corner, as in 4.3(a), then the product of diagonal entries $\det(A)$ is zero; at the same time, the matrix equations $A'X = 0$ and $AX = 0$ will have nontrivial solutions, so the left multiplication operator $L_A : \mathbb{K}^n \to \mathbb{K}^n$ fails to be invertible (because $\ker(L_A) \neq (0)$) and a matrix inverse A^{-1} fails to exist. The situation in Figure 4.3(b) is better: since Type II and Type III operations can only change $\det(A)$ by a $\pm$ sign, $\det(A) = \pm\det(A') = \pm\prod_{i=1}^n \lambda_i$ is nonzero. Concurrently, $AX = 0$ has only the trivial solution, L_A is an invertible linear operator on $\mathbb{K}^n$, and a matrix inverse A^{-1} exists. $\square$

To summarize, we have proved the following result (and a little more).

2.11. Theorem. *If* $A \in \mathrm{M}(n, \mathbb{K})$ *then* A^{-1} *exists if and only if* Type II *and* Type II *row operations yield a modified echelon form that is upper triangular, with all diagonal*

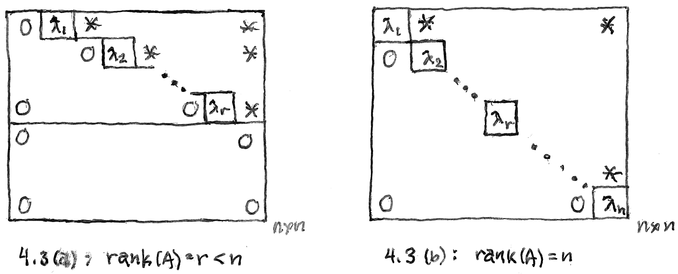

Figure 4.3. Row operations of Type II and III reduce an $n \times n$ matrix A to one of the two "modified echelon forms" A' shown in 4.3(a)–4.3(b); in both the step corners contain nonzero scalars $\lambda_1, \ldots, \lambda_r$ that need not $= 1$, and $r = \mathrm{rank}(A)$ with $r = n$ in 4.3(b).

If there are columns that do not meet a step-corner as in 4.3(a), then some diagonal entries in in A' are zero and $\det(A) = \pm\det(A') = 0$. In the situation of 4.3(b) $\det(A) = \pm\det(A') = \pm(\lambda_1 \cdot \ldots \cdot \lambda_n)$ because Type II and III elementary operations have determinant $= \pm 1$. In this case $\det(A)$ is nonzero and its value can be determined by inspection, except for a $(\pm)$ sign.

entries nonzero.

$$E_m \cdot \ldots \cdot E_1 A = \begin{pmatrix} \boxed{\lambda_1} & 0 & & & * \\ 0 & \boxed{\lambda_2} & & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ 0 & & & 0 & \boxed{\lambda_n} \end{pmatrix} \qquad (\lambda_i \neq 0)$$

Then the determinant is

$$\det(A) = \prod_{k=1}^{m} \det(E_k)^{-1} \cdot \prod_{i=1}^{n} \lambda_i$$

The factor $\prod_{k=0}^{n} \det(E_k)^{-1}$ *attributed to the row operations can only be* ± 1 *since no* Type I *operations are involved. On the other hand, if the modified echelon form contains columns that do not meet a setp corner, then* $\det(A) = 0$ *and* A^{-1} *does not exist.*

The basic definition (31) of the determinant is computationally very costly. Below we will give an algorithm ("*expansion by minors*") which is often useful in studying the algebraic properties of determinants, but it is still pretty costly compared to the row reduction method developed above. To illustrate:

n = Matrix Size	Expansion by Minors		Row Reduction	
	Adds	Multiplies	Adds	Multiplies
2	1	2	1	3
4	23	40	14	23
5	119	205	30	45
10	3.6×10^6	6.2×10^6	285	339

The technique used above also yields a fairly efficient algorithm for computing A^{-1} (which at the same time determines whether A is in fact invertible). Allowing all three types of row operations, an invertible matrix can be driven into its *reduced echelon form*, which is just the identity matrix $I_{n\times n}$. In this case

$$E_m \cdot \ldots \cdot E_1 \cdot A = I_{n\times n} \qquad \text{and} \qquad A^{-1} = E_1^{-1} \cdot \ldots \cdot E_m^{-1} \cdot I_{n\times n} \tag{33}$$

Each inverse E_k^{-1} is easily computed; it is just another elementary matrix of the same type as E_k. This can be codified as an explicit algorithm:

> **The Gauss-Seidel Algorithm.** *Starting with the augmented* $n \times 2n$ *matrix* $[A : I_{n\times n}]$, *perform row operations to put* A *into "reduced" echelon form* (*upper triangular with zeros above all step corners*). *If* $\text{rank}(A) < n$ *and* A *is not invertible this will be evident – not all columns include a step-corner – and the algorithm reports that* $\det(A) = 0$ *and* A *is not invertible. Otherwise, every column is a pivot column and the reduced echelon form of* A *is just the identity matrix. Applying the same operations to the entire augmented matrix transforms* $[A : I_{n\times n}] \rightarrow [I_{n\times n} : B]$ *in which* $B = A^{-1}$. (*Why?*)

Another consequence of the preceding discussion is the very important multiplicative property of determinants.

2.12. Theorem (Multiplicative Property). *If* $A,\ B \in \mathrm{M}(n, \mathbb{K})$ *then*

$$\det(AB) = \det(A) \cdot \det(B)$$

Proof: If A is singular then AB is singular (Exercise 2.9) so $\det(A) = 0$ and $\det(AB) = 0$ by Proposition 2.10. Thus

$$\det(AB) = 0 = 0{\cdot}\det(B) = \det(A){\cdot}\det(B) \quad ,$$

and similarly if B is the singular factor.

Otherwise A and B are nonsingular and so is AB, so we can find elementary matrices such that $E_m{\cdot}\ldots{\cdot}E_1 A = I_{n\times n}$, which implies $A = E_1^{-1}{\cdot}\ldots{\cdot}E_m^{-1}$. By repeated application of Lemma 2.8 we see that

$$\det(A) = \prod_{i=1}^{m} \det(E_i^{-1})$$

and

$$\begin{aligned}
\det(AB) &= \det(E_1^{-1}) \cdot \det\left(E_2^{-1} \cdot \ldots \cdot E_m^{-1} B\right) \\
&= \textstyle\prod_i \det(E_i^{-1}) \cdot \det(B) = \det(A) \cdot \det(B). \qquad \Box
\end{aligned}$$

2.13. Exercise. If $A \in \mathrm{M}(n, \mathbb{K})$ is invertible then $\det(A^{-1}) = \det(A)^{-1}$. If $A, B \in \mathrm{M}(n, \mathbb{K})$ and S is an invertible matrix such that $B = SAS^{-1}$ then $\det(B) = \det(A)$.

Thus $\det(A)$ is a "*similarity invariant*"– it has constant value for all matrices in a similarity class. We will encounter several other similarity invariants of matrices in the following discussion.

2.14. Exercise. Explain why $\text{rank}(A)$ of an $n \times n$ matrix is a similarity invariant.

2.15. Exercise. An $n \times n$ matrix A is said to be **orthogonal** if $A^{\mathrm{t}} A = I_{n\times n}$. Prove that

1. $A^{\mathrm{t}} A = I \Rightarrow AA^{\mathrm{t}} = I$, so A is orthogonal $\Leftrightarrow A^{\mathrm{t}} = A^{-1}$ (two-sided inverse).

2. $\det(A) = \pm 1$ for any orthogonal matrix, over any field.

Hint: Recall the comments posted in (32). For (1.) it suffices to show $A^{\mathrm{t}}A = I \Rightarrow$ the operator $L_A : \mathbb{K}^n \to \mathbb{K}^n$ is one-to-one.

2.16. Exercise. Use Type II and III row operations to find the determinant of the following matrix.

$$A = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 4 \end{pmatrix}$$

2.17. Exercise. Use Type II and III row operations to show that $\det(A) = -16i$ for the following matrix in $\mathrm{M}(4, \mathbb{C})$, where $i = \sqrt{-1}$.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}$$

2.18. Exercise. Apply the Gauss-Seidel algorithm to find A^{-1} for the matrices

$$\text{(i)}\quad A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 8 & 4 \\ 0 & 4 & 7 \end{pmatrix} \qquad\qquad \text{(ii)}\quad A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 4 & 1 \\ 0 & 4 & 2 \end{pmatrix}$$

2.19. Exercise. Consider the set of matrices H_n of the form

$$A = \begin{pmatrix} 1 & x_1 & \cdots & x_n & z \\ 0 & 1 & & & y_n \\ & & \ddots & & \vdots \\ & & & 1 & y_1 \\ 0 & & & & 1 \end{pmatrix}$$

with x_i, y_j, z in $\mathbb{K}$. When $\mathbb{K} = \mathbb{R}$ this is the n-dimensional *Heisenberg group* of quantum mechanics.

1. Prove that H_n is closed under matrix product.

2. Prove that the inverse A^{-1} of any matrix in H_n is also in H_n (compute it explicitly in terms of the parameters x_i, y_j, z).

Since the identity matrix is also in H_n, that means H_n is a *matrix group* contained in $\mathrm{GL}(n + 2, \mathbb{K})$.

2.20. Exercise. For $n \geq 2$ let

$$A = \begin{pmatrix} 0 & 1 & & & & 0 \\ 1 & 0 & 1 & & & \\ 0 & 1 & 0 & 1 & & \\ & & & \ddots & & \\ & & & 1 & 0 & 1 \\ 0 & & & & 1 & 0 \end{pmatrix}$$

Use row operations to

1. Calculate $\det(A)$.

2. Calculate the inverse A^{-1} if it exists.

Note: The outcome will depend on whether n is even or odd.

2.21. Exercise. Given a diagonal matrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with distinct entries, find an invertible matrix S such that conjugation $D \mapsto SDS^{-1}$ interchanges the i^{th} and j^{th} diagonal entries ($i \neq j$):

$$S \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \boxed{\lambda_i} & & & \\ & & & \ddots & & \\ & & & & \boxed{\lambda_j} & \\ & & & & & \ddots \\ 0 & & & & & \lambda_n \end{pmatrix} S^{-1} = \begin{pmatrix} \lambda_1 & & & & & 0 \\ & \ddots & & & & \\ & & \boxed{\lambda_j} & & & \\ & & & \ddots & & \\ & & & & \boxed{\lambda_i} & \\ & & & & & \ddots \\ 0 & & & & & \lambda_n \end{pmatrix}$$

Hint: Think row *and* column operations on D. Note that if E_{II} is a Type II elementary matrix then $E^{-1} = E = E^{\mathrm{t}}$, and right multiplication by E^{t} effects the corresponding column operation.

Determinants of Matrices vs Determinants of Linear Operators.

A determinant $\det(T)$ can be unambiguously assigned to any linear operator $T : V \to V$ on a finite dimensional space. Given a basis $\mathfrak{X} = \{e_i\}$ in V, we get a matrix $[T]_{\mathfrak{X}\mathfrak{X}}$ and could entertain the idea of assigning

$$\det(T) = \det\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big) \quad , \tag{34}$$

but for this to make sense the outcome must be independent of the choice of basis. This actually works. If $\mathfrak{Y}$ is any other basis we know there is an invertible matrix $S = [\mathrm{id}_{\mathrm{V}}]_{\mathfrak{Y}\mathfrak{X}}$ such that $[T]_{\mathfrak{Y}\mathfrak{Y}} = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1}$, and then by Theorem 2.12

$$\begin{aligned} \det\Big([T]_{\mathfrak{Y}\mathfrak{Y}}\Big) &= \det(S) \cdot \det\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big) \cdot \det(S^{-1}) \\ &= \det(SS^{-1})\cdot\det\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big) \;=\; \det(I_{n\times n}) \cdot \det\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big) \\ &= \det\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big) \end{aligned}$$

as required. Thus the determinant (34) of a linear operator is well defined.

The **trace** $\mathrm{Tr}(T)$ is another well-defined attribute of an operator $T : V \to V$ when $\dim(V) < \infty$. Recall Exercise 4.19 of Chapter II: For $n \times n$ matrices the trace $\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}$ is a linear operator $\mathrm{Tr} : \mathrm{M}(n, \mathbb{K}) \to \mathbb{K}$ such that $\mathrm{Tr}(I_{n\times n}) = n$ and $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. If $\mathfrak{X}$, $\mathfrak{Y}$ are bases for V, we get

$$\mathrm{Tr}\Big([T]_{\mathfrak{Y}\mathfrak{Y}}\Big) = \mathrm{Tr}\Big(S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1}\Big) = \mathrm{Tr}\Big(S^{-1}S \cdot [T]_{\mathfrak{X}\mathfrak{X}}\Big) = \mathrm{Tr}\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big)$$

Thus

$$\mathrm{Tr}(T) = \mathrm{Tr}\Big([T]_{\mathfrak{X}\mathfrak{X}}\Big) \tag{35}$$

determines a well-defined trace on operators. Note, however, that if $T : V \to W$ with $V \neq W$, there is no natural way to assign a "determinant" or "trace" to T, even if $\dim(V) = \dim(W)$. The problem is philosophical: there is no natural way to say that a basis $\mathfrak{X}$ in V is the "same as" another basis $\mathfrak{Y}$ in W.

The operator trace has the same algebraic properties as the matrix trace.

2.22. Exercise. If $A, B : V \to V$ are linear operators on a finite dimensional space V, prove that

1. $\mathrm{Tr} : \mathrm{Hom}_{\mathbb{K}}(V, V) \to \mathbb{K}$ is a $\mathbb{K}$-linear map between vector spaces:
$$\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B) \qquad \text{and} \qquad \mathrm{Tr}(\lambda \cdot A) = \lambda \cdot \mathrm{Tr}(A)$$
2. $\mathrm{Tr}(\mathrm{id}_V) = n \cdot \dim(V)$;
3. $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ (composition product of operators);
4. If S is an invertible operator and $B = SAS^{-1}$ then $\mathrm{Tr}(B) = \mathrm{Tr}(A)$.

The last statement shows that Tr is a similarity invariant for linear operators; so is the determinant *det*.

2.23. Exercise. If $T : V \to V$ is a linear operator on a finite dimensional space prove that
$$\mathrm{Tr}(T) = \mathrm{Tr}(T^{\mathrm{t}}) \qquad \text{and} \qquad \det(T) = \det(T^{\mathrm{t}})$$

Note: A conceptual issue arises here: T maps $V \to V$ while the transpose $T^{\mathrm{t}} : V^* \to V^*$ acts on an entirely different vector space! But if you take a basis $\mathfrak{X}$ in V and the dual basis $\mathfrak{X}^*$ in V^* the definitions (34) and (35) still have something useful to say.

2.24. Exercise. Let $P : V \to V$ be a projection (associated with some direct sum decomposition $V = E \oplus F$) that projects vectors onto E along F. Prove that $\mathrm{Tr}(P) = \dim_{\mathbb{K}}(E)$.

Expansion by Minors and Cramer's Rule. The following result allows a recursive computation of an $n \times n$ determinant once we can compute $(n-1) \times (n-1)$ determinants. Although it is useful for determining algebraic properties of determinants, and is handy for small matrices, it is prohibitively expensive in computing time for large n. This expansion is keyed to a particular row (or column) of A and involves an $(n-1) \times (n-1)$ determinant (the "minors" of the title) for each row entry.

2.25. Theorem (Cramer's Rule). *For any row $1 \leq i \leq n$, we can write*
$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \cdot \det\left(\tilde{A}_{ij}\right)$$
where $\tilde{A}_{ij}$ = the $(n-1) \times (n-1)$ submatrix obtained by deleting Row_i and Col_j from A. Similarly, for any column $1 \leq j \leq n$ we have
$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \cdot \det\left(\tilde{A}_{ij}\right)$$

Proof: Since $\det(A) = \det(A^{\mathrm{t}})$, it is enough to prove the result for expansion along a row. Each term in the sum
$$\det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \cdot \left(a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}\right)$$
contains just one term from $Row_i(A) = (a_{i1}, \ldots, a_{in})$, so by gathering together terms we may write
$$\det(A) = a_{i1}a_{i1}^* + \ldots + a_{in}a_{in}^*$$
in which a_{ij}^* involves no entry from $Row_i(A)$.

Our task is to show $a_{ij}^* = (-1)^{i+j} \det\left(\tilde{A}_{ij}\right)$. One approach is to reduce to the case when $i = j = n$. In that special situation, we get
$$a_{nn}a_{nn}^* = \sum_{\sigma \in S_n'} sgn(\sigma) \cdot \left(a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}\right)$$

where $S_n' \subseteq S_n$ is the subgroup of permutations such that $\sigma(n) = n$ (the subgroup that "stabilizes" the element "n" in $X = \{1, 2, \ldots, n\}$).

2.26. Exercise. If $\tilde{\sigma} \in S_{n-1}$ is regarded as the permutation $\sigma \in S_n' \subseteq S_n$ such that $\sigma(n) = n$ and $\sigma(k) = \tilde{\sigma}(k)$ for $1 \leq k \leq n-1$, show that $sgn(\tilde{\sigma}) = sgn(\sigma)$.

In view of this the sum $\sum_{\sigma \in S_n'}(...)$ becomes $\sum_{\tilde{\sigma} \in S_{n-1}}(...)$. Thus

$$a_{nn}^* = (-1)^{n+n} \det\left(\tilde{A}_{nn}\right) = \det(A_{nn})$$

Now consider any i and j. Interchange $Row_i(A)$ with successive adjacent rows ("flips") until it is at the bottom. This does not affect the value of $\det(\tilde{A}_{ij})$ because the relative positions of the other rows and columns are not affected; however each flip switched the sign of a_{ij} in the formula, and there are $n - i$ such changes. Similarly we may move $Col_j(A)$ to the n^{th} column, incurring $n - j$ sign changes. Thus

$$a_{ij}^* = (-1)^{n-i+n-j} \det(\tilde{A}_{ij}) = (-1)^{i+j} \det(\tilde{A}_{ij})$$

for all i and j, proving the theorem. $\square$

We post the following formula for A^{-1} without proof (cf Schaums, p 267-68). If matrix $A \in \mathrm{M}(n, \mathbb{K})$ is invertible we have

$$A^{-1} = \frac{1}{\det(A)} \cdot \left(\mathrm{Cof}(A)\right)^{\mathrm{t}} \tag{36}$$

where the $n \times n$ "*cofactor matrix*" $\mathrm{Cof}(A)$ has i, j entry $= (-1)^{i+j}\tilde{A}_{ij}$, and $\tilde{A}_{ij} =$ determinant of the $(n-1) \times (n-1)$ submatrix obtained by deleting (Row_i) and (Col_j) from A.

Chapter VI. Inner Product Spaces.

VI.1. Basic Definitions and Examples.

In Calculus you encountered Euclidean coordinate spaces $\mathbb{R}^n$ equipped with additional structure: an **inner product** $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.

Euclidean Inner Product: $B(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i$

which is often abbreviated to $B(x, y) = (x, y)$. Associated with it we have the **Euclidean norm**

$$\|\mathbf{x}\| = \sum_{i=1}^{n} |x_i|^2 = (\mathbf{x}, \mathbf{x})^{1/2}$$

which represents the "length" of a vector, and a **distance function**

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

which gives the Euclidean distance from $\mathbf{x}$ to $\mathbf{y}$. Note that $\mathbf{y} = \mathbf{x} + (\mathbf{y} - \mathbf{x})$.

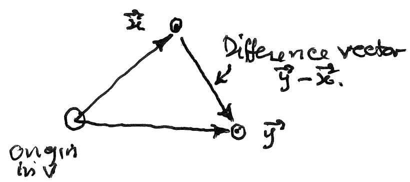

Figure 6.1. The distance between points $\mathbf{x}$, $\mathbf{y}$ in an inner product space is interpreted as the norm (length) $\|\mathbf{y} - \mathbf{x}\|$ of the difference vector $\Delta\mathbf{x} = \mathbf{y} - \mathbf{x}$.

This inner product on $\mathbb{R}^n$ has the following geometric interpretation

$$(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\| \cdot \|\mathbf{x}\| \cdot \cos\big(\theta(\mathbf{x}, \mathbf{y})\big)$$

where θ is the angle between x and y, measured in the plane $M = \mathbb{R}\text{-span}\{\mathbf{x}, \mathbf{y}\}$, the 2-dimensional subspace in $\mathbb{R}^n$ spanned by $\mathbf{x}$ and $\mathbf{y}$. **Orthogonality** of two vectors is then interpreted to mean $(\mathbf{x}, \mathbf{y}) = 0$; the zero vector is orthogonal to everybody, by definition. These notions of *length, distance*, and *orthogonality* do not exist in unadorned vector spaces.

We now generalize the notion of inner product to arbitrary vector spaces, even if they are infinite-dimensional.

1.1. Definition. *If V is a vector space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, an* **inner product** *is a map $B : V \times V \to \mathbb{K}$ taking ordered pairs of vectors to scalars $B(v_1, v_2) \in \mathbb{K}$ with the following properties*

1. Separate Additivity in each Entry. *B is additive in each input if the other input is held fixed:*
   - $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$
   - $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2)$.

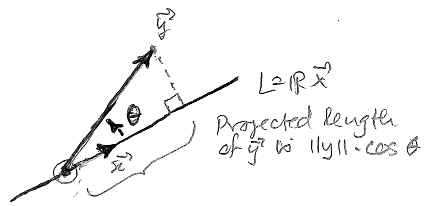

Figure 6.2. Geometric interpretation of the inner product $(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}\|\,\|\mathbf{y}\|\cdot\cos(\theta(\mathbf{x}, \mathbf{y}))$ in $\mathbb{R}^n$. The projected length of a vector $\mathbf{y}$ onto the line $L = \mathbb{R}\mathbf{x}$ is $\|\mathbf{y}\|\cdot\cos(\theta)$. The angle $\theta(\mathbf{x}, \mathbf{y})$ is measured within the two-dimensional subspace $M = \mathbb{R}\text{-span}\{\mathbf{x}, \mathbf{y}\}$. Vectors are *orthogonal* when $\cos\theta = 0$, so $(\mathbf{x}, \mathbf{y}) = 0$. The zero vector is orthogonal to everybody.

for v, v_i, w, w_i in V.

2. Positive Definite. *For all $v \in V$,*

$$B(v, v) \geq 0 \quad \text{and} \quad B(v, v) = 0 \text{ if and only if } v = 0$$

3. Hermitian Symetric. *For all $v, w \in V$,*

$$B(v, w) = \overline{B(w, v)} \qquad \text{when inputs are interchanged.}$$

Conjugation does nothing for $x \in \mathbb{R}$ ($\overline{x} = x$ for $x \in \mathbb{R}$), so an inner product on a real vector space is simply symmetric, with $B(w, v) = B(v, w)$.

4. Hermitian. *For $\lambda \in \mathbb{K}$, $v, w \in V$,*
   - $B(\lambda v, w) = \lambda B(v, w)$ *and,*
   - $B(v, \lambda w) = \bar{\lambda} B(v, w)$.

An inner product on a real vector space is just a **bilinear map** *– one that is $\mathbb{R}$-linear in each input when the other is held fixed – because conjugation does nothing in $\mathbb{R}$.*

The Euclidean inner product in $\mathbb{R}^n$ is a special case of the standard Euclidean inner product in complex coordinate space $V = \mathbb{C}^n$,

$$(\mathbf{z}, \mathbf{w}) = \sum_{j=1}^{n} z_j \overline{w_j} \quad ,$$

which is easily seen to have properties (1.)–(4.) The corresponding Euclidean norm and distance functions on $\mathbb{C}^n$ are then

$$\|\mathbf{z}\| = (\mathbf{z}, \mathbf{z})^{1/2} = \Big[\, \sum_{j=1}^{n} |z_j|^2 \,\Big]^{1/2} \qquad \text{and} \qquad d(\mathbf{z}, \mathbf{w}) = \|\mathbf{z} - \mathbf{w}\| = \Big[\, \sum_{j=1}^{n} |z_j - w_j|^2 \,\Big]^{1/2}$$

Again, properties (1.) - (4.) are easily verified.

For an arbitrary inner product B we define the corresponding norm and distance functions

$$\|v\|_B = B(v, v)^{1/2} \qquad d_B(v_1, v_2) = \|v_1 - v_2\|_B$$

which are no longer given by such formulas.

1.2. Example. Here are two important examples of inner product spaces.

1. On $V = \mathbb{C}^n$ (For $\mathbb{R}^n$) we can define "nonstandard" inner products by assigning different positive weights $\alpha_j > 0$ to each coordinate direction, taking

$$B_\alpha(\mathbf{z}, \mathbf{w}) = \sum_{j=1}^{n} \alpha_j \cdot z_j \overline{w_j} \qquad \text{with norm} \qquad \|\mathbf{z}\|_\alpha = \Big[\sum_{j=1}^{n} \alpha_j \cdot |z_j|^2\Big]^{1/2}$$

   This is easily seen to be an inner product. Thus the standard Euclidean inner product on $\mathbb{R}^n$ or $\mathbb{C}^n$, for which $\alpha_1 = \ldots = \alpha_n = 1$, is part of a much larger family.

2. The space $\mathcal{C}[a,b]$ of continuous complex-valued functions $f : [a,b] \to \mathbb{C}$ becomes an inner product space if we define

$$(f,h)_2 = \int_a^b f(t)\overline{h(t)}\, dt \qquad \text{(Riemann integral)}$$

   The corresponding "**L^2-norm**" of a function is then

$$\|f\|_2 = \Big[\int_a^b |f(t)|^2\, dt\Big]^{1/2};$$

   the inner product axioms follow from simple properties of the Riemann integral. This infinite-dimensional inner product space arises in many applications, particularly Fourier analysis. □

1.3. Exercise. Verify that both inner products in the last example actually satisfy the inner product axioms. In particular, explain why the L^2-inner product $(f,h)_2$ has $\|f\|_2 > 0$ when f is not the zero function ($f(t) \equiv 0$ for all t).

We now take up the basic properties common to all inner product spaces.

1.4. Theorem. *On any inner product space V the associated norm has the following properties*

1. $\|x\| \geq 0$*;*
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$ (*and in particular,* $\|-x\| = \|x\|$ *);*
3. (Triangle Inequality) *For $x, y \in V$, $\|x \pm y\| \leq \|x\| + \|y\|$.*

Proof: The first two are obvious. The third is important because it implies that the distance function $d_B(x,y) = \|x - y\|$ satisfies the "*geometric triangle inequality*"

$$d_B(x,y) \leq d_B(x,z) + d_B(z,y), \qquad \text{for all } x,y,z \in V$$

as indicated in Figure 6.3. This follows directlly from (3.) because

$$d_B(x,y) = \|x - y\| = \|(x-z) + (z-y)\| \leq \|x - z\| + \|z - y\| = d_B(x,z) + d_B(z,y)$$

The version of (3.) involving a $(-)$ sign follows from that featuring a $(+)$ because $v - w = v + (-w)$ and $\|-w\| = \|w\|$.

The proof of (3.) is based on an equally important inequality:

1.5. Lemma (Schwartz Inequality). *If B is an inner product on a real or complex vector space then*

$$|B(x,y)| \leq \|x\|_B \cdot \|y\|_B$$

for all $x, y \in V$.

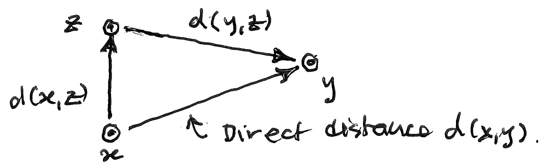

Figure 6.3. The meaning of the Triangle Inequality: direct distance from $\mathbf{x}$ to $\mathbf{y}$ is always $\leq$ the sum of distances $d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ to any third vector $\mathbf{z} \in V$.

Proof: For all real t we have $\phi(t) = \|x + ty\|_B^2 \geq 0$. By the axioms governing B we can rewrite $\phi(t)$ as

$$\begin{aligned}
\phi(t) &= B(x + ty, x + ty) \\
&= B(x, x) + B(ty, x) + B(x, ty) + B(ty, ty) \\
&= \|x\|_B^2 + t\,B(x, y) + t\,\overline{B(x, y)} + t^2\|y\|_B^2 \\
&= \|x\|_B^2 + 2t\,\mathrm{Re}(B(x, y)) + t^2\|y\|_B^2
\end{aligned}$$

because $B(tx, y) = tB(x, y)$ and $B(x, ty) = tB(x, y)$ (since $t \in \mathbb{R}$), and $z + \overline{z} = 2\,\mathrm{Re}(z) = 2x$ for $z = x + iy$ in $\mathbb{C}$. Now $\phi : \mathbb{R} \to \mathbb{R}$ is a quadratic function whose minimum value occurs at t_0 where

$$\frac{d\phi}{dt}(t_0) = 2t_0\|y\|_B^2 + \mathrm{Re}(B(x, y)) = 0$$

or

$$t_0 = \frac{-\mathrm{Re}(B(x, y))}{2\|y\|_B^2}$$

Inserting this into ϕ we find the actual minimum value of ϕ:

$$0 \leq \min\{\phi(t) : t \in \mathbb{R}\} = \frac{\|x\|_B^2 \cdot \|y\|_B^2 - 2|\mathrm{Re}(B(x, y))|^2 + |\mathrm{Re}(B(x, y))|^2}{\|y\|_B^2}$$

Thus

$$0 \leq \|x\|_B^2 \cdot \|y\|_B^2 - |\mathrm{Re}(B(x, y))|^2$$

which in turn implies

$$|\mathrm{Re}\,B(x, y)| \leq \|x\|_B \cdot \|y\|_B \qquad \text{for all } x, y \in V.$$

If we replace $x \mapsto e^{i\theta}x$ this does not change $\|x\|$ since $|e^{i\theta}| = |\cos(\theta) + i\sin(\theta)| = 1$ for real θ; in the inner product on the left we have $B(e^{i\theta}x, y) = e^{i\theta}B(x, y)$. We may now take $\theta \in \mathbb{R}$ so that $e^{i\theta} \cdot B(x, y) = |B(x, y)|$. For this particular choice of θ we get

$$\begin{aligned}
0 \leq |\mathrm{Re}(B(e^{i\theta}x, y))| &= |\mathrm{Re}(e^{i\theta}B(x, y))| \\
&= \mathrm{Re}(|B(x, y)|) = |B(x, y)| \leq \|x\|_B \cdot \|y\|_B \quad .
\end{aligned}$$

That proves the Schwartz inequality. $\square$

Proof (Triangle Inequality): The algebra is easier if we prove the (equivalent) inequality obtained when we square both sides:

$$\begin{aligned}
0 \;\leq\; \|x + y\|^2 &\leq \big(\|x\| + \|y\|\big)^2 \\
&= \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2
\end{aligned}$$

In proving the Schwartz inequality we saw that

$$\|x+y\|^2 = (x+y\,,\,x+y) = \|x\|^2 + 2\,\mathrm{Re}(x,y) + \|y\|^2$$

so our proof is finished if we can show $2\,\mathrm{Re}(x,y) \leq 2\|x\|\cdot\|y\|$. But

$$\mathrm{Re}(z) \leq |\,\mathrm{Re}(z)| \leq |z| \qquad \text{for all } z \in \mathbb{C}$$

and then the Schwartz inequality yields

$$\mathrm{Re}\big(B(x,y)\big) \leq |B(x,y)| \leq \|x\|_B \cdot \|y\|_B$$

as desired. $\square$

1.6. Example. On $V = \mathrm{M}(n,\mathbb{K})$ we define the **Hilbert-Schmidt** inner product and norm for matrices:

$$(A,B)_{\mathrm{HS}} = \mathrm{Tr}(B^*A) \qquad \text{and} \qquad \|A\|^2_{\mathrm{HS}} = \sum_{i,j=1} |a_{ij}|^2 = \mathrm{Tr}(A^*A) \tag{44}$$

It is easily verified that this is an inner product. First note that the trace map from $\mathrm{M}(n,\mathbb{K}) \rightarrow \mathbb{K}$

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}$$

is a complex linear map and $\mathrm{Tr}(\,\overline{A}\,) = \overline{\mathrm{Tr}(A)}$; then observe that

$$\|A\|_2^2 = (A,A)_{\mathrm{HS}} = \sum_{i,j=1}^{n} |a_{ij}|^2 \text{ is } > 0 \text{ unless } A \text{ is the zero matrix.}$$

Alternatively, consider what happens when we identify $\mathrm{M}(n,\mathbb{C}) \cong \mathbb{C}^{n^2}$ as complex vector spaces. The Hilbert-Schmidt norm becomes the usual Euclidean norm on $\mathbb{C}^{n^2}$, and likewise for the inner products; obviously $(A,B)_{\mathrm{HS}}$ is then an inner product on matrix space.

The norm $\|A\|_{\mathrm{HS}}$ and the sup-norm $\|A\|_\infty$ discussed in Chapter V are different ways to measure the "size" of a matrix; the HS-norm turns out to be particularly well adapted to applications in statistics, starting with "least-squares regression" and moving on into "analysis of variance." Each of these norms determines a notion of matrix convergence $A_n \rightarrow A$ as $n \rightarrow \infty$ in $\mathrm{M}(N,\mathbb{C})$.

$$\|\cdot\|_2\text{-Convergence:} \qquad \|A_n - A\|_{\mathrm{HS}} = \Big[\sum_{i,j} |a_{ij}^{(n)} - a_{ij}|^2\Big]^{1/2} \rightarrow 0 \text{ as } n \rightarrow \infty$$

$$\|\cdot\|_\infty\text{-Convergence:} \qquad \|A_n - A\|_\infty = \max_{i,j}\{\,|a_{ij}^{(n)} - a_{ij}|\,\} \rightarrow 0 \text{ as } n \rightarrow \infty$$

However, despite their differences both norms *determine the same notion of matrix convergence.*

$$A_n \rightarrow A \text{ in } \|\cdot\|_2\text{-norm} \;\Leftrightarrow\; A_n \rightarrow A \text{ in } \|\cdot\|_\infty\text{-norm}$$

The reason is explained in the next exercise. $\square$

1.7. Exercise. Show that there exist bounds $M_2, M_\infty > 0$ such that the $\|\cdot\|_2$ and $\|\cdot\|_\infty$ norms mutually dominate each other

$$\|x\|_2 \leq M_\infty\,\|x\|_\infty \qquad \text{and} \qquad \|x\|_\infty \leq M_2\,\|x\|_2$$

for all $x \in \mathbb{C}^n$. Explain why this leads to the conclusion that $A_n \rightarrow A$ in $\|\cdot\|_2$-norm if and only if $A_n \rightarrow A$ in $\|\cdot\|_\infty$-norm.
Hint: The Schwartz inequality might be helpful in one direction.

The **polarization identities** below show that inner products over $\mathbb{R}$ or $\mathbb{C}$ can be reconstructed if we only know the norms of vectors in V. Over $\mathbb{C}$ we have

$$(45) \quad B(x,y) = \frac{1}{4}\sum_{k=0}^{3}\frac{1}{i^k}B(x+i^k y\,,\,x+i^k y) = \frac{1}{4}\sum_{k=0}^{3}\frac{1}{i^k}\,\|x+i^k y\|^2, \qquad \text{where } i=\sqrt{-1}$$

Over $\mathbb{R}$ we only need 2 terms:

$$B(x,y) = \frac{1}{4}\Big(B(x+y\,,\,x+y) + (-1)B(x-y\,,\,x-y)\Big)$$

1.8. Exercise. Expand

$$(x+i^k y\,,\,x+i^k y) = \|x+i^k y\|^2$$

to verify the polarization identities.

Orthonormal Bases in Inner Product Spaces.

A set $\mathfrak{X} = \{e_i : i \in I\}$ of vectors is **orthogonal** if $(e_i, e_j) = 0$ for $i \neq j$; it is **orthonormal** if

$$(e_i, e_j) = \delta_{ij} \quad \text{(Kronecker delta)} \qquad \text{for all } i,j \in I\ .$$

An orthonormal set can be infinite (in infinite dimensional inner product spaces), and all vectors in it are nonzero; an orthogonal family could have $v_i = 0$ for some indices since $(v,0) = 0$ for any v. The set $\mathfrak{X}$ is an **orthonormal basis** (ON basis) if it is orthonormal and V is spanned by $\{\mathfrak{X}\}$.

1.9. Proposition. *Orthonormal sets have the following properties.*

1. *Orthonormal sets are independent;*

2. *If $\mathfrak{X} = \{e_i : i \in I\}$ is a finite orthonormal set and v is in $M = \mathbb{K}$-span$\{\mathfrak{X}\}$ then by* (1.) *$\mathfrak{X}$ is a basis for M and the expansion of any v in M with respect to this basis is just*
$$v = \sum_{i\in I}(v,e_i)\,e_i$$
(*Finiteness of $\mathfrak{X}$ required for $\sum_{i\in I}(\ldots)$ to make sense; otherwise the right side is an infinite series*).

In particular if $\mathfrak{X} = \{e_1, ..., e_n\}$ is an orthonormal basis for a finite-dimansional inner product space V, the coefficients in the expansion

$$v = \sum_{i=1}^{n}(v,e_i)\,e_i, \qquad \textit{for every } v \in V$$

are easily computed by taking inner products.

Proof: For (1.), if a finite sum $\sum_i c_i e_i$ equals 0 we have

$$0 = (v,e_k) = \sum_i c_i(e_i,e_k) = \sum_i c_i\delta_{ik} = c_k$$

for each k, so the e_i are independent. Part (2.) is an immediate consequence of (1.): we know $\{e_i\}$ is a basis, and if $v = \sum_i c_i e_i$ is its expansion the inner product with a typical basis vector is

$$(v, e_k) = \sum_i c_i (e_i, e_k) = \sum_i c_i \delta_{ik} = c_k \ . \quad \square$$

1.10. Corollary. *If vectors $\{v_1, ..., v_n\}$ are nonzero, orthogonal, and a vector basis in V, then the renormalized vectors*

$$e_i = \frac{v_i}{\|v_i\|} \qquad \text{for } 1 \leq i \leq n$$

are an orthonormal basis. $\square$

Entries in the matrix $[T]_{\mathfrak{Y}\mathfrak{X}}$ of a linear operator are easily computed by taking inner products if the bases are orthonormal (but not for arbitrary bases).

1.11. Exercise. Let $T : V \rightarrow W$ be a linear operator between finite-dimensional inner product spaces and let $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_i\}$ be orthonormal bases. Prove that the entries in $[T]_{\mathfrak{Y}\mathfrak{X}}$ are given by

$$T_{ij} = \big(T(e_j), f_i\big)_W = \overline{\big(f_i, T(e_j)\big)_W}$$

for $1 \leq i \leq \dim(W),\ 1 \leq j \leq \dim(V)$.

The fundamental fact about ON bases is that the coefficients in $v = \sum_{k=1}^n (v, e_i)\, e_i$ determine the norm $\|v\|$ via a generalization of Pythagoras' Formula for $\mathbb{R}^n$,

$$\textsc{Pythagoras:} \qquad \text{If } \mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i \qquad \text{then} \qquad \|\mathbf{x}\|^2 = \sum_{i=1}^n |x_i|^2$$

We start by proving a fundamental inequality.

1.12. Theorem (Bessel's Inequality). *Let $\mathfrak{X} = \{e_1, \ldots, e_m\}$ be any finite orthonormal set in an inner product space V (possibly infinite-dimensional). Then*

$$\sum_{i=1}^n |(v, e_i)|^2 \leq \|v\|^2 \qquad \text{for all } v \in V \tag{46}$$

Furthermore, if $v' = v - \sum_{i=1}^n (v, e_i)\, e_i$, this vector is orthogonal to each e_j and hence is orthogonal to all the vectors in the linear span $M = \mathbb{K}$-span$\{\mathfrak{X}\}$.

Note: The inequality (46) becomes an equality if $\mathfrak{X}$ is an orthonormal basis for V because then $v' = 0$.

Proof: Since inner products are conjugate bilinear, we have

$$\begin{aligned}
0 \ \leq\ \|v'\|^2 \ &= \ (v', v') \ = \ \Big(v - \sum_{i=1}^m (v, e_i)\, e_i\, ,\, v - \sum_{j=1}^m (v, e_j)\, e_j \ \Big) \\
&= \ (v, v) - \big(\sum_i (v, e_i)\, e_i\, ,\, v\big) - \big(v\, ,\, \sum_j (v, e_j)\, e_j\big) + \big(\sum_i (v, e_i)\, e_i\, ,\, \sum_j (v, e_j)\, e_j\,\big) \\
&= \ \|v\|^2 - \sum_i (v, e_i) \cdot (e_i, v) - \sum_j \overline{(v, e_j)} \cdot (v, e_j) + \sum_{i,j} (v, e_i) \cdot \overline{(v, e_j)} \cdot (e_i, e_j) \\
&= \ \|v\|^2 - \sum_i |(v, e_i)|^2 - \sum_j |(v, e_j)|^2 + \sum_i |(v, e_i)|^2 \quad (\text{since } (e_k, v) = \overline{(v, e_k)}\ ) \\
&= \ \|v\|^2 - \sum_i |(v, e_i)|^2
\end{aligned}$$

Therefore
$$\sum_{i=1} |(v,e_i)|^2 \leq \|v\|^2$$
as required.

The second statement now follows easily because
$$\begin{aligned}(v',e_k) &= \Big(v - \sum_j (v,e_j)\, e_j\,,\, e_k\Big) \;=\; (v,e_k) - \sum_j (v,e_j)\cdot(e_j,e_k) \\ &= (v,e_k)-(v,e_k) = 0 \qquad \text{for all } k\end{aligned}$$
Furthermore, if $w = \sum_{k=1}^m c_k e_k$ is any vector in M we also have
$$(v',w) = \sum_k c_k (v',e_k) = 0\ ,$$
so v' is orthogonal to M as claimed. $\Box$

1.13. Corollary (Pythagoras). *If $\mathfrak{X}$ is an orthonormal basis in a finite dimensional inner product space, then*
$$\|v\|^2 = \sum_{i=1}^m |(v,e_i)|^2$$
(*sum of squares of the coefficients in the basis expansion $v = \sum_i (v,e_i)\, e_i$*).

1.14. Theorem. *Orthonormal bases exist in any finite dimensional inner product space.*

Proof: We argue by induction on $n = \dim(V)$; the result is trivial if $n = 1$ (any vector of length 1 is an orthonormal basis). If $\dim(V) = n+1$, let v_0 be any nonzero vector. The linear functional $\ell_0 : v \to (v, v_0)$ is nonzero, and as in Example 1.3 of Chapter III its kernel
$$M = \{v : (v,v_0) = 0\} = \big(\mathbb{K}v_0\big)^{\perp}$$
is a hyperplane of dimension $\dim(V) - 1 = n$. By the induction hypothesis there is an ON basis $\mathfrak{X}_0 = \{e_1, , \ldots, e_n\}$ in M, and every vector in M is orthogonal to v_0. If we rescale v_0 and adjoin $e_{n+1} = v_0/\|v_0\|$ to $\mathfrak{X}_0$ the enlarged set $\mathfrak{X} = \{e_1,\ldots,e_n,e_{n+1}\}$ is obviously orthonormal; it is also a basis for V. [By Lemma 4.4 of Chapter III, $\mathfrak{X}$ is a basis for $W = \mathbb{K}\text{-span}\{\mathfrak{X}\} \subseteq V$, and since $\dim(W) = |\mathfrak{X}| = n+1 = \dim(V)$ we must have $W = V$.] $\Box$

VI.2. Orthogonal Complements and Projections.

If M is a subspace of a (possibly infinite-dimensional) inner product space V, its **orthogonal complement** $M^{\perp}$ is the set of vectors orthogonal to every vector in M,
$$M^{\perp} = \{\, v \in V : (v,m) = 0, \text{ for all } m \in M\,\} = \{\, v : (v,M) = \{0\}\,\}\ .$$
Obviously $\{0\}^{\perp} = V$ and $V^{\perp} = \{0\}$ from the Axioms for inner product.

2.1. Exercise. Show that $M^{\perp}$ is again a subspace of V, and that
$$M_1 \subseteq M_2 \ \Rightarrow\ M_2^{\perp} \subseteq M_1^{\perp}\ .$$

2.2. Proposition. *If M is a finite dimensional subspace of a (possibly infinite-dimensional) inner product space V, then*

1. *$M \cap M^{\perp} = \{0\}$ and $M + M^{\perp} = V$, so we have a direct sum decomposition $V = M \oplus M^{\perp}$.*

2. *If* $\dim(V) < \infty$ *we also have* $(M^{\perp})^{\perp} = M$*; if* $|V| = \infty$ *we can only say that* $M \subseteq (M^{\perp})^{\perp}$.

Proof: If $v \in M \cap M^{\perp}$ then $\|v\|^2 = (v, v) = 0$ so $v = 0$ and $M \cap M^{\perp} = \{0\}$. Now let $\{e_1, ..., e_n\}$ be an orthonormal basis for M. If $v \in V$ write

$$v = \Big(v - \sum_{i=0}^{m} (v, e_i)\, e_i\Big) + \sum_{i=1}^{m} (v, e_i)\, e_i \;=\; v_{\perp} + v_{\|}$$

in which $v_{\perp}$ is orthogonal to M and $v_{\|}$ is the component of v "parallel to" the subspace M (because it lies in M). Then for all $v \in V$ we have

$$(v, v_{\perp}) \;=\; (v_{\perp} + v_{\|}\,,\, v_{\perp}) \;=\; (v_{\perp}, v_{\perp}) + (v_{\|}, v_{\perp}) = \|v_{\perp}\|^2 + 0 \;=\; \|v_{\perp}\|^2$$

If $v \in (M^{\perp})^{\perp}$, so $(v, v_{\perp}) = 0$, we conclude that $\|v_{\perp}\| = 0$ and hence $v = v_{\perp} + v_{\|} = 0 + v_{\|}$ is in M. That proves the reverse inclusion $M^{\perp\perp} \subseteq M$. $\square$

The situation is illustrated in Figure 6.4.

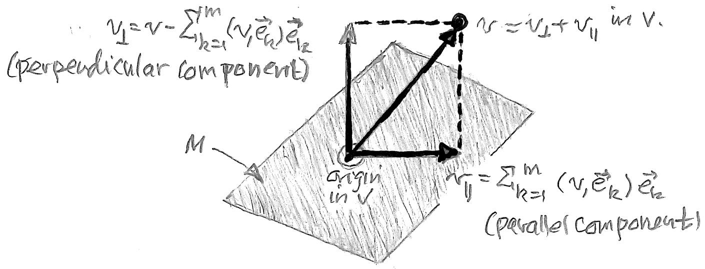

Figure 6.4. Given an ON basis $\{e_i, \ldots, e_m\}$ in a finite dimensional subspace $M \subseteq V$, the vector $v_{\|} = \sum_{k=1}^{m} (v, e_k)\, e_k$ is in M and $v_{\perp} = v - v_{\|}$ is orthogonal to M. These are the components of $v \in V$ "parallel to$M$" and "perpendicular to M," with $v = v_{\perp} + v_{\|}$.

Orthogonal Projections on Inner Product Spaces.

If an inner product space is a direct sum $V = V_1 \oplus \ldots \oplus V_r$ we call this an **orthogonal direct sum** if the subspaces are mutually orthogonal.

$$(V_i, V_j) = 0 \qquad \text{if } i \neq j$$

We indicate this by writing $V = V_1 \dot{\oplus} \ldots \dot{\oplus} V_r = \dot{\bigoplus}_{i=1}^{r} V_i$. The decomposition $V = M \dot{\oplus} M^{\perp}$ of Proposition 2.2 was an orthogonal decomposition.

In equation Exercise 3.5 of Chapter II we defined the linear projection operators $P_i : V \to V$ associated with an ordinary direct sum decomposition $V = V_1 \oplus \ldots \oplus V_r$, and showed that such operators are precisely the linear operators that have the *idempotent property* $P^2 = P$. In fact there is a bijective correspondence

$$(\textit{idempotent linear operators}) \;\longleftrightarrow\; (\textit{direct sum decompositions } V = R \oplus K)\ ,$$

described in Proposition 3.7 of Chapter II, and reprised below.

Theorem. *If a linear operator* $P : V \to V$ *is idempotent operator, so* $P^2 = P$*, there is a direct sum decomposition* $V = R \oplus P$ *such that* P *projects* V *onto* R *along* K*. In particular,*

$$R = R(P) = \text{range}(P) \qquad \textit{and} \qquad K = K(P) = \ker(P)$$

Furthermore $Q = I - P$ is also idempotent and

$$R(Q) = K(P) \qquad \textit{and} \qquad K(Q) = R(P)$$

When V is an inner product space we will see that the projections associated with an *orthogonal* direct sum $V = E\dot{\oplus}F$ have special properties. They are also easy to compute using the inner product. (Compare what follows with the calculations in Example 3.6 of Chapter II, of projections associated with an ordinary direct sum decomposition $V = E \oplus F$ in a space without inner product.)

Projections associated with an orthogonal direct sum decomposition $V = V_1\dot{\oplus}\dots\dot{\oplus}V_r$ are called **orthogonal projections**.

2.3. Lemma. *If $V = E\dot{\oplus}F$ is an orthogonal direct sum decomposition of a finite dimensional inner product space, then*

$$E^{\perp} = F \quad \text{and} \quad F^{\perp} = E \qquad\qquad E^{\perp\perp} = E \quad \text{and} \quad F^{\perp\perp} = F$$

Proof: The argument for F is the same as that for E. We proved that $E^{\perp\perp} = E$ in Proposition 2.2 and we know that $E \subseteq F^{\perp}$ by definition; based on this we will prove the reverse inequality $E \supseteq F^{\perp}$.

Since $|V| < \infty$ we have $V = F \oplus F^{\perp}$, so that $|V| = |F| + |F^{\perp}|$; since $V = E \oplus F$ we also have $|V| = |F| + |E|$. Therefore $|E| = |F^{\perp}|$. But $E \subseteq F^{\perp}$ in an orthogonal direct sum $E\dot{\oplus}F$, so we conclude that $E = F^{\perp}$. $\square$

2.4. Exercise. Let $V = V_1\dot{\oplus}\dots\dot{\oplus}V_r$ be an orthogonal direct sum decomposition of an inner product space (not necessarily finite dimensional).

(a) If W_i is the linear span $\sum_{j\neq i} V_j$, prove that $W_i \perp V_i$ for each i, and $V = V_i\dot{\oplus}W_i$.

(b) If $v = v_1 + \dots + v_r$ is the unique decomposition into pairwise orthogonal vectors $v_i \in V_i$, prove that $\|v\|^2 = \sum_i \|v_i\|^2$.

The identity (2.) is yet another version of Pythagoras' formula.

2.5. Exercise. In a finite dimensional inner product space, prove that the *Parseval formula*

$$(v, w) = \sum_{i=1}^{n} (v, e_i)\cdot(e_i, w)$$

holds for every orthonormal basis $\{e_1, \dots, e_n\}$.

The Gram-Schmidt Construction. We now show how any independent set of vectors $\{v_1, \dots, v_n\}$ in an inner product space can be modified to obtain an orthonormal set of vectors $\{e_1, \dots, e_n\}$ with the same linear span. This **Gram-Schmidt construction** is recursive, and at each step we have

1. $e_k \in \mathbb{K}\text{-span}\{v_1, ..., v_k\}$
2. $M_k = \mathbb{K}\text{-span}\{e_1, ..., e_k\}$ is equal to $\mathbb{K}\text{-span}\{v_1, .., v_k\}$ for each $1 \leq k \leq n$.

The result is an orthonormal basis $\{e_1, ..., e_n\}$ for $M = \mathbb{K}\text{-span}\{v_1, .., v_n\}$ (and for all of V if the $\{v_i\}$ span V). The construction procedes inductively by constructing two sequences of vectors $\{u_i\}$ and $\{e_i\}$.

Step 1: Take

$$u_1 = v_1 \qquad \text{and} \qquad e_1 = \frac{v_1}{\|v_1\|}$$

Conditions (1.) and (2.) obviously hold and $\mathbb{K}\cdot v_1 = \mathbb{K} \cdot u_1 = \mathbb{K}\cdot e_1$.

STEP 2: Define

$$u_2 = v_2 - (v_2|e_1)\cdot e_1 \qquad \text{and} \qquad e_2 = \frac{u_2}{\|u_2\|}.$$

Obviously $u_2 \in \mathbb{K}\text{-span}\{v_1, v_2\}$ and $u_2 \neq 0$ because $v_2 \notin \mathbb{K}v_1 = \mathbb{K}e_1 = M_1$; thus e_2 is well defined. Furthermore

1. $u_2 \perp M_1$ because

$$(u_2, e_1) = \big(v_2 - (v_2, e_1)e_1\,,\, e_1\big) = (v_2, e_1) - (v_2, e_1)\cdot(e_1, e_1) = 0 \;\Rightarrow\; e_2 \perp M_1$$

   hence $\{e_1, e_2\}$ is an orthonormal set of vectors;

2. $M_2 = \mathbb{K}\text{-span}\{e_1, e_2\} = \mathbb{K}u_2 + \mathbb{K}e_1 = \mathbb{K}v_2 + \mathbb{K}e_1 = \mathbb{K}v_2 + \mathbb{K}v_1 = \mathbb{K}\text{-span}\{v_1, v_2\}$.

If $n = 2$ we're done; otherwise continue with

STEP 3: Define

$$u_3 = v_3 - \sum_{i=1}^{2}(v_3, e_i)\cdot e_i = v_3 - \sum_{i=1}^{2}\frac{(v_3, u_i)}{\|u_i\|^2}\,u_i$$

Then $u_3 \neq 0$ because the sum is in $\mathbb{K}\text{-span}\{v_1, v_2\}$ and the v_i are independent; thus $e_3 = \frac{u_3}{\|u_3\|}$ is well defined. We have $u_3 \perp M_2$ because

$$\begin{aligned}
(u_3, e_1) &= \Big(v_3 - \sum_{i=1}^{2}(v_3, e_i)e_i\,,\, e_1\Big) \\
&= (v_3, e_1) - \sum_{i=1}^{2}(v_3, e_i)\cdot(e_i, e_1) \\
&= (v_3, e_1) - (v_3, e_1) = 0\ ,
\end{aligned}$$

and similarly $(u_3, e_2) = 0$, hence $e_3 \perp M_2 = \mathbb{K}\text{-span}\{e_1, e_2\}$. Finally,

$$\begin{aligned}
\mathbb{K}\text{-span}\{e_1, e_2, e_3\} &= \mathbb{K}u_3 + \mathbb{K}\text{-}\{e_1, e_2\} \;=\; \mathbb{K}v_3 + \mathbb{K}\text{-}\{e_1, e_2\} \\
&= \mathbb{K}v_3 + \mathbb{K}\text{-}\{v_1, v_2\} \;=\; \mathbb{K}\text{-}\{v_1, v_2, v_3\}
\end{aligned}$$

At the k^{th} step we have produced orthonormal vectors $\{e_1, ..., e_k\}$ with $\mathbb{K}\text{-span}\{e_1, ..., e_k\} = \mathbb{K}\text{-span}\{v_1, ..., v_k\} = M_k$. Now for the induction step:

STEP $k + 1$: Define

$$u_{k+1} = v_{k+1} - \sum_{i=1}^{k}(v_{k+1}, e_i)\,e_i = v_{k+1} - \sum_{i=1}^{k}\frac{(v_{k+1}, u_i)}{\|u_i\|^2}\,u_i$$

and

$$e_{k+1} = \frac{u_{k+1}}{\|u_{k+1}\|}.$$

Again $u_{k+1} \neq 0$ because $v_{k+1} \notin M_k = \mathbb{K}\text{-span}\{v_1, ..., v_k\} = \mathbb{K}\text{-span}\{e_1, ..., e_k\}$, so e_{k+1} is well defined. Furthermore $u_{k+1} \perp M_k$ because

$$\begin{aligned}
(u_{k+1}, e_j) &= \Big(v_{k+1} - \sum_{i=1}^{k}(v_{k+1}, e_i)\,e_i\,,\, e_j\Big) \\
&= (v_{k+1}, e_j) - \sum_{i=1}^{k}(v_{k+1}, e_i)\cdot(e_i, e_j) \\
&= (v_{k+1}, e_j) - (v_{k+1}, e_j) = 0
\end{aligned}$$

hence also $e_{k+1} \perp M_k$. Then

$$\begin{aligned} \mathbb{K}\text{-}\{e_1, ..., e_{k+1}\} &= \mathbb{K}u_{k+1} + \mathbb{K}\text{-}\{e_1, ..., e_k\} = \mathbb{K}v_{k+1} + \mathbb{K}\text{-}\{e_1, ..., e_k\} \\ &= \mathbb{K}\text{-}\{v_1, ..., v_{k+1}\} \ . \end{aligned}$$

By induction, $\{e_1, ..., e_n\}$ has the properties claimed. $\square$

Note that the outcome of Step(k+1) depends only on the $\{e_1, ..., e_k\}$ and the new vector v_{k+1}; the original vectors $\{v_1, ..., v_k\}$ play no further role in the inductive process.

2.6. Example. The standard inner product in $\mathcal{C}[-1, 1]$ is the L^2 inner product

$$(f, h)_2 = \int_{-1}^{1} f(t)\overline{h(t)}\, dt$$

for functions $f : [-1, 1] \to \mathbb{C}$. Regarding $v_1 = \mathbb{1}$, $v_2 = x$, $v_3 = x^2$ as functions from $[-1, 1] \to \mathbb{C}$, these vectors are independent. Find the orthonormal set $\{e_1, e_2, e_3\}$ produced by the Gram-Schmidt process.

Solution: We have $u_1 = v_1 = \mathbb{1}$ and since $\|u_1\|^2 = \int_{-1}^{1} \mathbb{1}\, dx = 2$, we get $e_1 = \frac{1}{\sqrt{2}} \cdot \mathbb{1}$. At the next step

$$u_2 = v_2 - (v_2, e_1)\, e_1 = v_2 - \frac{(v_2, u_1)}{\|u_1\|^2}\, u_1 = x - \frac{\int_{-1}^{1} x \cdot \mathbb{1}\, dx}{\|u_1\|^2} \cdot \mathbb{1} = x - 0 = x$$

and

$$\|u_2\|^2 = \int_{-1}^{1} x^2\, dx = 2\int_{0}^{1} x^2\, dx = 2\left[\frac{1}{3}x^3\Big|_0^1\right] = \frac{2}{3}$$

The second basis vector is

$$e_2 = \frac{u_2}{\|u_2\|} = \sqrt{\frac{3}{2}} \cdot x$$

At the next step:

$$\begin{aligned} u_3 &= v_3 - \big((v_3|e_1)e_1 + (v_3, e_2)e_2\big) \\ &= v_3 - \Big(\frac{(v_3, u_2)}{\|u_2\|^2} \cdot u_2 + \frac{(v_3, u_1)}{\|u_1\|^2} \cdot u_1\Big) \\ &= x^2 \ - \ \frac{\int_{-1}^{1} x^2 \cdot x\, dx}{\frac{2}{3}} \cdot x \ - \ \frac{\int_{-1}^{1} x^2 \cdot \mathbb{1}\, dx}{2} \cdot \mathbb{1} \\ &= x^2 - 0 - \frac{1}{3}\mathbb{1} \ = \ x^2 - \frac{1}{3} \end{aligned}$$

Then

$$\begin{aligned} \|u_3\|^2 &= \int_{-1}^{1} |u_3(x)|^2\, dx \ = \ \int_{-1}^{1} \big(x^2 - \frac{1}{3}\big)^2\, dx \\ &= \int_{-1}^{1} \Big(x^4 - \frac{2}{3}x^2 + \frac{1}{9}\Big)dx \\ &= 2 \cdot \left[\frac{x^5}{5} - \frac{2}{9}x^3 + \frac{1}{9}x\,\Big|_0^1\right] \ = \ \frac{8}{45} \end{aligned}$$

and the third orthonormal basis vector is

$$e_3 = \frac{u_3}{\|u_3\|} = \sqrt{\frac{45}{8}}\Big(x^2 - \frac{1}{3}\Big) = \sqrt{\frac{5}{8}}(3x^2 - 1) \quad \square$$

If we extend the original list to include $v_4 = x^4$ we may compute e_4 knowing only e_1, e_2, e_3 (or u_1, u_2, u_3) and v_4; there is no need to repeat the previous calculations!

2.7. Exercise. Find u_4 and e_4 in the above situation.

This process can be continued indefinitely to produce the orthonormal family of **Legendre polynomials** $e_1(t), e_2(t), \ldots, e_n(t) \ldots$ in the space of polynomials $\mathbb{C}[x]$ restricted to the interval $[-1, 1]$. (This is also true for $\mathbb{R}[x]$ restricted to $[-1, 1]$ since the Legendre polynomials all have real coefficients.) Clearly the $(n + 1)$-dimensional subspace M_n obtained by restricting the space of polynomials of degree $\leq n$

$$\mathcal{P}_n = \mathbb{K}\text{-span}\{e_1, ..., e_{n+1}\} = \mathbb{K}\text{-span}\{1, x, \ldots, x^n\}$$

to the interval (so $M_n = \mathcal{P}_n|[-1, 1]$) has $\{e_1, \ldots, e_{n+1}\}$ as an ON basis with respect to the usual inner product on $\mathcal{C}[-1, 1]$

$$(f, h)_2 = \int_{-1}^{1} f(t)\overline{h(t)}\, dt\ .$$

Restricting the full set of Legendre polynomials $e_1(t), \ldots, e_{n+1}(t), \ldots$ to $[-1, 1]$ yields an orthonormal set of vectors in the infinite-dimensional inner product space $\mathcal{C}[-1, 1]$. The orthogonal projection $P_n : \mathcal{C}[-1, 1] \to M_n \subseteq \mathcal{C}[-1, 1]$ associated with the orthogonal direct sum decomposition $V = M_n \oplus (M_n)^\perp$ (in which $\dim(M_n)^\perp = \infty$) is given by the explicit formula

$$\begin{aligned} P_n f(t) &= \sum_{k=1}^{n+1} (f, e_k)\, e_k(t) \qquad (-1 \leq t \leq 1) \\ &= \sum_{k=1}^{n+1} \Big(\int_{-1}^{1} f(x)\overline{e_k(x)}\, dx \Big) \cdot e_k(t) \\ &= \sum_{k=0}^{n} c_k\, t^k \qquad (c_k \in \mathbb{C}) \end{aligned}$$

for any continuous function on $[-1, 1]$. The projected image $P_n f(t)$ is a polynomial of degree $\leq n$ even though $f(t)$ is continuous and need not be differentiable.

A standard result from analysis shows that the partial sums of the infinite series $\sum_{k=0}^{\infty} c_k\, t^k$ converge in the L^2-norm to the original function $f(t)$ throughout the interval $-1 \leq t \leq 1$,

$$\|f - P_n f\|_2 = \Big[\int_{-1}^{1} |f(t) - P_n f(t)|^2\, dt \Big]^{1/2} \to 0 \quad \text{as } n \to \infty$$

for all $f \in \mathcal{C}[-1, 1]$.

It must be noted that this series expansion of $f(t) \sim \sum_{k=0}^{\infty} c_k\, t^k$ is not at all the same thing as a Taylor series expansion about $t = 0$, which in any case would not make sense because $f(t)$ is only assumed continuous (the derivatives used to compute Taylor coefficients might not exist!) In fact, convergence of this series in the L^2-norm is much more robust than convergence of Taylor series, which is why it is so useful in applications.

Fourier Series Expansions. The *complex trig polynomials* $E_n(t) = e^{2\pi i n t}$ $(n \in \mathbb{Z})$ are periodic complex-valued functions on $\mathbb{R}$; each has period $\Delta t = 1$ since

$$e^{2\pi i n(t+1)} = e^{2\pi i n t} \cdot e^{2\pi i n} = e^{2\pi i n t} \qquad \text{for all } t \in \mathbb{R} \text{ and } n \in \mathbb{Z}.$$

If $e_n(t)$ is the restriction of $E_n(t)$ to the "period-interval" $I = [0,1]$ we get an ON family of vectors with respect to the usual inner product $(f,h) = \int_0^1 f(t)\overline{h(t)}\,dt$ on $\mathcal{C}[0,1]$, because

$$\begin{aligned} \|e_n\|^2 &= \int_0^1 |e_n(t)|^2\,dt \;=\; \int_0^1 1\,dt = 1 \\ (e_m, e_n) &= \int_0^1 e_m(t)\overline{e_n(t)}\,dt \;=\; \int_0^1 e^{2\pi i(m-n)t}dt \\ &= \left[\frac{e^{2\pi i(m-n)t}}{2\pi i(m-n)}\Big|_0^1\right] = 0 \qquad \text{if} \;\; m \neq n. \end{aligned}$$

Thus $\{e_n : n \in \mathbb{Z}\}$ is an orthonormal family in $\mathcal{C}[0,1]$.

For $N \geq 0$ let $M_N = \mathbb{K}\text{-span}\{e_k : -N \leq k \leq N\}$. For f in this subspace we have the basis expansion:

$$f = \sum_{k=-N}^{N} (f, e_k)\, e_k = \sum_{k=-N}^{N} c_k e^{2\pi i k t}$$

where c_k is the k^{th} **Fourier coefficient**

$$c_k = (f, e_k) = \int_0^1 f(t) e^{-2\pi i k t} dt. \tag{47}$$

By Bessel's inequality:

$$\|f\|_2^2 = \int_0^1 |f(t)|^2 dt \geq \sum_{k=-N}^{N} |c_k|^2 = \sum_{k=-N}^{N} |(f, e_k)|^2$$

and this is true for $N = 1, 2, \ldots$. The projection P_N of $\mathcal{C}[0,1]$ onto M_N along $M_N^{\perp}$ is then given by

$$P_N f(t) = \sum_{k=-N}^{N} c_k e_k(t) = \sum_{k=-N}^{N} (f, e_k)\, e^{2\pi i k t}, \qquad N = 0, 1, 2, \ldots$$

because $P_N(f) \in M_N$ by definition, and $\big(f - P_N f, e_k\big) = 0$ for $-N \leq k \leq N$.

The **Fourier series** of a continuous (or bounded Riemann integrable) complex-valued function $f : [0,1] \to \mathbb{C}$ is the infinite series

$$f \sim \sum_{k \in \mathbb{Z}} (f, e_k) \cdot e^{2\pi i k t} \tag{48}$$

whose coefficients $c_k = (f, e_n)$ are the Fourier coefficients defined in (47).

It is not immediately clear when this series converges, but when convergence is suitably interpreted it can be proved that the series does converge, and to the initial function $f(t)$. This expansion has proved to be extremely useful in applications. Its significance is best described as follows.

If t is regarded as a time variable, and $F(t)$ is some sort of periodic "signal" or "waveform" such that $F(t+1) = F(t)$ for all t, then F is completely determined by its restriction $f = F|\,[0,1]$ to the basic period interval $0 \leq t \leq 1$. The Fourier series expansion of f on this interval can in turn be regarded as a representation of the original waveform as a "superposition," with suitable weights, of the basic periodic waveforms $E_n(t) = e^{2\pi i n t}$ $(t \in \mathbb{R})$.

$$F(t) \sim \sum_{n=-\infty}^{+\infty} c_n \cdot E_n(t) \qquad \text{for all } t \in \mathbb{R}$$

For instance, this implies that any periodic sound wave $F(t)$ with period $\Delta t = 1$ can be reconstructed by superposing scalar multiples of the "pure tones" $E_n(t)$, which have frequencies $\omega_n = n$ cycles per second. This is precisely how sound synthesizers work. It is remarkable, that the correct "weight" assigned to each pure tone is the Fourier coefficient $c_n = (f, e_n)$; even more remarkable is the fact that *complex-valued* weights $c_k \in \mathbb{C}$ must be allowed, even if the signal is real-valued, because the functions $E_n(t) = \cos(2\pi nt) + i\sin(2\pi nt)$ are complex-valued.

If f is piecewise differentiable the infinite series (48) converges (except at points of discontinuity) to the original periodic function $f(t)$. Furthermore the following results can be proved for any continuous (or Riemann integrable) function on $[0, 1]$.

THEOREM. *If $f(t)$ is bounded and Riemann integrable for $0 \le t \le 1$, then*

1. L^2-NORM CONVERGENCE: *The partial sums of the Fourier series* (48) *converge to $f(t)$ in the* L^2*-norm.*

$$\|f - \sum_{k=-N}^{N} (f, e_k)\, e_k\|_2 \to 0 \ \textit{as}\ N \to \infty$$

2. EXTENDED BESSEL: $\|f\|^2 = \int_0^1 |f(t)|^2\,dt$ *is equal to* $\sum_{k\in\mathbb{Z}} |(f, e_k)|^2$.

The norm $\|f - h\|_2 = \Big[\int |f - h|^2\,dt\Big]^{1/2}$ is often referred to as the "RMS = *Root Mean Square*" distance between f and h.

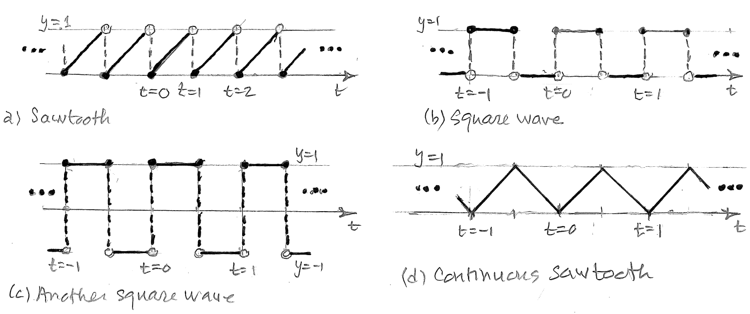

Figure 6.5. Various waveforms with period $\Delta t = 1$, whose Fourier transforms can be computed by Calculus methods.

2.8. Example. Let

$$f(t) = \begin{cases} t & \text{for } 0 \le t < 1 \\ 0 & \text{for } t = 1 \end{cases}$$

This is the restriction to $[0, 1]$ of the periodic "sawtooth" waveform in Figure 6.5(a). Find its Fourier series.

Solution: If $k \neq 0$ integration by parts yields

$$\begin{aligned} c_k &= \int_0^1 te^{-2\pi ikt}\,dt \\ &= \left[\frac{-1}{2\pi ik}e^{-2\pi ikt}\cdot t\Big|_0^1\right] - \int_0^1 \frac{-1}{2i\pi ik}e^{-2\pi ikt}\,dt \\ &= \frac{-1}{2\pi ik} + \frac{1}{2\pi ik}(e_k, e_0) \qquad \text{(where } e_0(t) \equiv 1 \text{ for all } t\text{)} \\ &= \frac{-1}{2\pi ik} \qquad \text{if } k \neq 0\ . \end{aligned}$$

For $k = 0$ we get a different result:

$$c_0 = \int_0^1 t\,dt = \tfrac{1}{2}$$

By Bessel's Inequality we have

$$\begin{aligned} \|f\|_2^2 &= \int_0^1 |f(t)|^2\,dt \ = \int_0^1 t^2\,dt = \tfrac{1}{3} \quad \text{(by direct calculation)} \\ &\geq \sum_{k=-N}^{N} |(f, e_k)|^2 \ = \sum_{k=-N}^{N} |c_k|^2 \\ &= \frac{1}{4} + \sum_{k\neq 0, -N\leq k\leq N} \frac{1}{4\pi^2 k^2} \end{aligned}$$

for any $N = 1, 2, ...$ If we multiply both sides by $4\pi^2$, then for all N we get

$$\begin{aligned} \frac{4}{3}\pi^2 &\geq \sum_{0<|k|\leq N} \frac{1}{k^2} + \pi^2 \\ \frac{1}{3}\pi^2 &\geq 2\cdot\sum_{k=1}^{N}\frac{1}{k^2} \\ \frac{\pi^2}{6} &\geq \sum_{k=1}^{N}\frac{1}{k^2} \quad \text{for all } N = 1, 2, \ldots \ \Rightarrow\ \frac{\pi^2}{6} \geq \sum_{k=1}^{\infty}\frac{1}{k^2} \end{aligned}$$

(the infinite series converges by the Integral Test). Once we know that $\|f\|^2 = \sum_{k\in\mathbb{Z}} |c_k|^2$ we get the famed formula

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}$$

The Fourier series associated with the sawtooth function $f(t)$ is

$$f(t) \sim \sum_{k=-\infty}^{\infty} (f, e_k)\, e_k(t) = \frac{1}{2}\cdot \mathbb{1} + \sum_{k\neq 0} \frac{-1}{2\pi ik}e^{2\pi ikt}\ ,$$

which converges pointwise for all $t \in \mathbb{R}$ except the "jump points" $t \in \mathbb{Z}$, where the series converges to the middle value $\frac{1}{2}$. $\square$

2.9. Exercise. Compute the Fourier transforms of the periodic functions whose graphs are shown in Figure 6.5 (b) – (d).

A Geometry Problem. The following result provides further insight into the meaning of the projection $P_N(v) = \sum_{i=1}^N (v, e_i)\, e_i$ where $\{e_i\}$ is an orthonormal family in an inner product space V.

2.10. Theorem. *If $\{e_1, \ldots, e_n\}$ is an orthonormal family in an inner product space, and $P_M(v) = \sum_{i=1}^n (v, e_i)\, e_i$ the projection of v onto $M = \mathbb{K}$-span$\{e_1, ..., e_n\}$ along $M^\perp$, then the image $P_M(v)$ is the point in M closest to v,*

$$\|P_M(v) - v\| = \min\{\, \|u - v\| : u \in M\}$$

for any $v \in V$. In particular the minimum is achieved at the unique point $P_M(v) \in M$.

Proof: Write $v = v_\parallel + v_\perp$ where $v_\parallel = P_M(v) = \sum_{i=1}^N (v, e_i)\, e_i$ and $v_\perp = v - \sum_i (v, e_i)\, e_i$. Obviously $v_\parallel \perp v_\perp$ and if z is any point in M we have $(v_\parallel - z) \in M$ and $(v - v_\parallel) \perp M$, so by Pythagoras

$$\begin{aligned} \|v - z\|^2 &= \|(v - v_\parallel) + (v_\parallel - z)\|^2 \\ &= \|v - v_\parallel\|^2 + \|v_\parallel - z\|^2 \end{aligned}$$

Thus

$$\|v - z\|^2 \geq \|v - v_\parallel\|^2$$

for all $z \in M$, so $\|v - z\|^2$ is minimized at $z = v_\parallel = \sum_{i=1}^N (v, e_i)\, e_i$. Figure 6.6 shows why the formula $\|v\|^2 = \|v_\parallel\|^2 + \|v_\perp\|^2$ really is equivalent to Pythagora's formula for right triangle (see the shaded triangle). $\square$

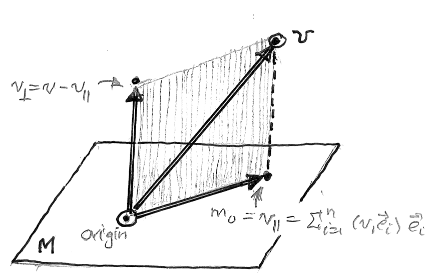

Figure 6.6. If M is a finite dimensional subspace of inner product space V and $v \in V$, the unique point in M closest to v is $m_0 = v_\parallel = \sum_i (v, e_i)\, e_i$, and the minimized distance is $\|v - m_{)}\|$. The shaded plane is spanned by the orthogonal vectors $v_\parallel$ and $v_\perp$ and we have $\|v\|^2 = \|v_\parallel\|^2 + \|v_\perp\|^2$ (Pythagoras' formula).

V.3. Adjoints and Orthonormal Decompositions.

Let V be a finite dimensional inner product space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Recall that a linear operator $T : V \to V$ is *diagonalizable* if there is a basis $\{e_1, \ldots, e_n\}$ of eigenvectors (so $T(e_i) = \mu_i e_i$ for some $\mu_i \in \mathbb{K}$). We have seen that this happens if and only if $V = \bigoplus_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$ where

$$\begin{aligned} \mathrm{sp}(T) &= (\text{the distinct eigenvalues of } T \text{ in } \mathbb{K}) = \{\lambda \in \mathbb{K} : E_\lambda(T) \neq (0)\} \\ E_\lambda(T) &= \{v \in V : (T - \lambda I)v = 0\} = \ker(T - \lambda I) \end{aligned}$$

We say T is **orthogonally diagonalizable** if there is an orthonormal basis $\{e_1, \ldots, e_n\}$ of eigenvectors, so $T(e_i) = \mu_i e_i$ for some $\mu_i \in \mathbb{K}$.

3.1. Lemma. *A linear operator $T : V \to V$ on a finite dimensional inner product space is orthogonally diagonalizable if and only if the eigenspaces span V and are pairwise* *orthogonal, so $E_\lambda(T) \perp E_\mu(T)$ for $\lambda \neq \nu$ in* $\mathrm{sp}(T)$.

Proof ($\Leftarrow$): is easy. We have seen that the span $W = \sum_{\lambda\in\mathrm{sp}(T)} E_\lambda(T)$ is a direct sum whether or not $W = V$. If $W = V$ and the E_λ are orthogonal then we have an orthogonal direct sum decomposition $V = \dot{\bigoplus}_\lambda E_\lambda(T)$. Taking an orthonormal basis in each E_λ we get a diagonalizing orthonormal basis for all of V.

Proof ($\Rightarrow$): If $\mathfrak{X} = \{e_1, \ldots, e_n\}$ is a diagonalizing orthonormal basis with $T(e_i) = \mu_i e_i$, each μ_i is an eigenvalue. Define

$$\mathrm{sp}' = \{\lambda \in \mathrm{sp}(T) : \lambda = \mu_i \text{ for some } i\,\} \subseteq \mathrm{sp}(T)$$

and for $\lambda \in \mathrm{sp}(T)$ let

$$M_\lambda = \sum \{\mathbb{K}e_i : \mu_i = \lambda\} \subseteq E_\lambda(T)$$

(which will $= (0)$ if λ does not appear among the scalars μ_i). Obviously $|M_\lambda| \leq |E_\lambda|$; furthermore, each e_i lies in some eigenspace E_λ, so

$$V = \mathbb{K}\text{-span}\{e_1, \ldots, e_n\} \subseteq \sum_{\lambda\in\mathrm{sp}(T)} E_\lambda \subseteq V$$

and these subspaces coincide. Thus

$$|V| = \sum_{\lambda\in\mathrm{sp}(T)} |E_\lambda| \geq \sum_{\lambda\in\mathrm{sp}'} |E_\lambda| \geq \sum_{\lambda\in\mathrm{sp}'} |M_\lambda| \;\geq\; |V|$$

and all sums are equal. (The last inequality holds because $\sum_{\lambda\in\mathrm{sp}'} M_\lambda \supseteq \sum_{j=1}^n \mathbb{K}e_j = V$.)

Now if $\mathrm{sp}(T) \neq \mathrm{sp}'$ the first inequality would be strict, and if $M_\lambda \overset{\subseteq}{\neq} E_\lambda$ the second the second would be strict, both impossible. We conclude that $|M_\lambda| = |E_\lambda(T)|$ so $M_\lambda = E_\lambda(T)$. But the M_λ are mutually orthogonal by definition, so the eigenspaces E_λ are pairwise orthogonal as desired. $\square$

Simple examples (discussed later) show that a linear operator on an inner product space can be diagonalizable in the ordinary sense but fail to be orthogonally diagonalizable. To explore this distinction further we need additional background, particularly the definition of *adjoints* of linear operators.

Dual Spaces of Inner Product Spaces. There is a natural identification of any finite dimensional inner product space V with its dual space V^*. It is implemented by a map $J : V \to V^*$ where $J(v) =$ the functional $\ell_v \in V^*$ such that

$$\langle \ell_v, x\rangle = (x, v) \qquad \text{for all } x \in V\ .$$

Each map ℓ_v is a linear functional because the inner product $(*, *)$ is $\mathbb{K}$-linear in its left hand entry (but conjugate linear in the right hand entry unless $\mathbb{K} = \mathbb{R}$). The map J is one-to-one because

$$J(v_1) = J(v_2) \;\Rightarrow\; 0 = \langle \ell_{v_1}, x\rangle - \langle \ell_{v_2}, x\rangle = (x, v_1) - (x, v_2) = (x, v_1 - v_2)$$

for all $x \in V$. Taking $x = v_1 - v_2$, we get $0 = \|v_1 - v_2\|^2$ which implies $v_1 - v_2 = 0$ and $v_1 = v_2$ by positive definiteness of the inner product. To see J is also surjective we invoke:

3.2. Lemma. *If V is finite dimensional inner product space, $\{e_1, \ldots, e_n\}$ an orthonormal basis, and $\ell \in V^*$, then*

$$\ell = J(v_0) \qquad \textit{where} \qquad v_0 = \sum_{i=1}^n \overline{\langle \ell, e_i\rangle}\, e_i$$

(*proving J surjective*).

Proof: For any $x \in V$ we have $x = \sum_i (x, e_i)\, e_i$. Hence by conjugate-linearity of $(*,*)$

$$\begin{aligned}
\langle J(v_0), x\rangle &= (x, v_0) = \Big(\sum_i x_i e_i\,,\, \sum_j \overline{\langle \ell, e_j\rangle}\, e_j\Big) = \sum_{i,j} x_i \langle \ell, e_j\rangle \cdot (e_i, e_j) \\
&= \sum_i x_i \langle \ell, e_i\rangle = \Big\langle \ell, \sum_i x_i e_i \Big\rangle = \ell(x) \qquad \text{for all } x \in V.
\end{aligned}$$

Therefore $J(v_0) = \ell$ as elements of V^*. $\square$

3.3. Exercise. Prove that $J : V \to V^*$ is a *conjugate linear* bijection: it is additive, with $J(v + v') = J(v) + J(v')$ for all $v, v' \in V$, but $J(\lambda v) = \overline{\lambda} J(v)$ for $v \in V$, $\lambda \in \mathbb{C}$.

The Adjoint Operator $\mathbf{T^*}$. If $T : V \to W$ is a linear operator between finite dimensional vector spaces we showed that there is a natural transpose $T^{\mathrm{t}} : W^* \to V^*$. Since $V \cong V^*$ for inner product spaces, it follows that there is a natural *adjoint operator* $T^* : V \to W$ between the original vector spaces, rather than their duals.

3.4. Theorem (Adjoint Operator). *Let V, W be finite dimensional inner product spaces and $T : V \to W$ a $\mathbb{K}$-linear operator. Then there is a unique $\mathbb{K}$-linear* **adjoint operator** $T^* : W \to V$ *such that*

$$(T(v), w)_W = (v, T^*(w))_V \qquad \text{for all } v \in V,\ w \in W \quad , \tag{49}$$

or equivalently $(T^*(w), v)_V = (w, T(v))_W$ owing to Hermitian symmetry of the inner product.

Proof: We define $T^*(w)$ for $w \in W$ using our observations about dual spaces. Given $w \in W$, we get a well defined linear functional ϕ_w on V if define

$$\langle\, \phi_w, v\rangle = (\, T(v), w)_W$$

(w is fixed; the variable is v).

Obviously $\phi_w \in V^*$ because $(*,*)_W$ is linear in its left-hand entry. By the previous discussion there is a unique vector in V, which we label $T^*(w)$, such that $J(T^*(w)) = \phi_w$ in V^*, hence

$$(T(x), w)_W = \langle \phi_w, x\rangle = \langle J(T^*(w)), x\rangle = (x, T^*(w))_V$$

We obtain a well defined map $T^* : W \to V$.

Once we know a map T^* satisfying (49) exists, it is easy to use these scalar identities to verify that T^* is a linear operator, and verify its important properties. For linearity we first observe that two vectors v_1, v_2 are equal in V if and only if $(v_1, x) = (v_2, x)$, for all $x \in V$ because the inner product is positive definite.

Then $T^*(w_1 + w_2) = T^*(w_1) + T^*(w_2)$ in V follows: for all $v \in V$ we have

$$\begin{aligned}
(T^*(w_1 + w_2), v)_V &= (w_1 + w_2, T(v))_W = (w_1, T(v))_W + (w_2, T(v))_W \\
&= (T^*(w_1), v)_V + (T^*(w_2, v)_V \qquad \text{(definition of } T^*(w_k)) \\
&= (T^*(w_1) + T^*(w_2), v)_V \quad \text{(linearity of } (*|*) \text{ in first entry)}
\end{aligned}$$

Similarly, $T^*(\lambda w) = \lambda T^*(w)$, for all $\lambda \in \mathbb{K}$, $w \in W$ (check that λ comes forward instead of $\overline{\lambda}$). $\square$

Note: A general philosophy regarding calculations with adjoints: Don't look at $T^*(v)$; look at $(T^*(v), w)$ instead, for all $v \in V, w \in W$.

3.5. Lemma. *On an inner product space $(T^*)^* = T$ as linear maps from $V \to W$.*

Proof: It suffices to check the scalar identities $(T^{**}(v), w)_W = (T(v), w)_W$, for all $v \in V$, $w \in W$. But by definition,

$$\big(T^{**}(v), w\big)_W = \big(v, T^*(w)\big)_V = (T(v), w)_W$$

Done. $\square$

The adjoint $T^* : W \to V$ of a linear operator $T : V \to W$ between inner product space is analogous to the transpose $T^{\mathrm{t}} : W^* \to V^*$. In fact, if V, W are inner product spaces and we identify $V = V^*$, $W = W^*$ via the maps $J_V : V \to V^*$, $J_W : W \to W^*$ then T^* *becomes* the transpose $T^{\mathrm{t}} : W^* \to V^*$ in the sense that the following diagram commutes:

$$\begin{array}{ccc} W & \xrightarrow{T^*} & V \\ J_W \downarrow & & \downarrow J_V \\ W^* & \xrightarrow{T^{\mathrm{t}}} & V^* \end{array}$$

That is ,

$$T^t \circ J_W = J_V \circ T^* \qquad (\text{ or } T^* = J_V^{-1} \circ T^{\mathrm{t}} \circ J_W)$$

3.6. Exercise. Prove this last identity from the definitions.

Furthermore, as remarked earlier, when V is just a vector space, there is a natural identification of $V \cong V^{**}$

$$j : V \to V^{**} \qquad \langle j(v), \ell\rangle = \ell(v) \quad \text{for all } \ell \in V^*, v \in V$$

We remarked that under this identification of $V^{**} \cong V$ we have $T^{\mathrm{tt}} = T$ for any linear operator $T : V \to W$, in the sense that the following diagram commutes

$$\begin{array}{ccc} V^{**} & \xrightarrow{T^{\mathrm{tt}}} & W^{**} \\ j_V \uparrow & & \uparrow j_W \\ V & \xrightarrow{T} & W \end{array}$$

If V, W are inner product spaces, we may actually identify $V \simeq V^*$ (something that cannot be done in any natural way in the absence of the extra structure an inner product provides). Then we may identify $V \cong V^* \cong V^{**} \cong V^{***} \cong ...$ and $W \cong W^* \cong W^{**} \cong W^{***} \cong ...$; when we do, T^{t} becomes T^* and T^{tt} becomes $T^{**} = T$.

3.7. Exercise (Basic Properties of Adjoints). Use (49) to prove:

(a) $I^* = I$ and $(\lambda I)^* = \overline{\lambda} I$,

(b) $(T_1 + T_2)^* = T_1^* + T_2^*$,

(c) $(\lambda T)^* = \bar{\lambda} T^*$ (conjugate-linearity)

3.8. Exercise. Given linear operators $V \xrightarrow{S} W \xrightarrow{T} Z$ between finite dimensional inner product spaces, prove that

$$(T \circ S)^* = S^* \circ T^* : Z \to V \ .$$

Note the reversal of order when we take adjoints.

3.9. Exercise. If $A \in \mathrm{M}(n, \mathbb{C})$ and $(A^*)_{ij} = \overline{A_{ji}}$ is the usual adjoint matrix, consider the operator $L_A : \mathbb{C}^n \to \mathbb{C}^n$ such that $L_A(\mathbf{z}) = A \cdot \mathbf{z}$. If $\mathbb{C}^n$ is given the standard inner product prove that

(a) If $\mathfrak{X} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard orthonormal basis then $[L_A]_{\mathfrak{X}\mathfrak{X}} = A$.

(b) $\big(L_A\big)^* = L_{A^*}$ as operators on $\mathbb{C}^n$.

3.10. Example (Self-Adjointness of Orthogonal Projections). On an unadorned vector space V the "idempotent" relation $P^2 = P$ identifies the linear operators that are projections associated with an ordinary direct sum decomposition $V = M \oplus N$. The same is true of an inner product space, but if we only know $P = P^2$ the subspaces M, N are not necessarily orthogonal. We now show that an idempotent operator P on an inner product space corresponds to an *orthogonal* direct sum decomposition $V = M \dot{\oplus} N$ if and only if it is **self-adjoint** ($P^* = P$), so that

$$P^2 = P = P^* \tag{50}$$

Discussion: If $M \perp N$ it is fairly easy to verify (Exercise 3.11) that the associated projection P_M of V onto $M = \mathrm{range}(P_M)$ along $N = \ker(P_M)$ is self-adjoint. If $v, w \in V$, let us indicate the components by writing $v = v_M + v_N$, $w = w_M + w_N$. With (49) in mind, self-adjointness of P_M emerges from the following calculation.

$$\begin{aligned}
\big(v, P_M^*(w)\big) &= \big(P_M(v), w\big) = (v_M, w_M + w_N) \quad (\text{definition of } P_M(v) = v_M) \\
&= (v_M, w_M) \qquad\qquad (\text{since } w_N \perp w_M) \\
&= (v_M + v_N, w_M) = (v, w_M) = \big(v, P_M(w)\big)
\end{aligned}$$

Since the is true for all $v \in V$ we get $P_M^*(w) = P_M(w)$ for all w, whence $P_M^* = P_M$ as operators.

For the converse we must prove: If the projection P_M associated with an ordinary direct sum decomposition $V = M \oplus N$ is self-adjoint, so that $P_M^* = P_M$, then the subspaces must be orthogonal. We leave this proof as an exercise. $\square$

3.11. Exercise. If $P : V \rightarrow V$ is a linear operator on a vector space such that $P^2 = P$ it is the projection operator associated with the decomposition

$$V = R \oplus K \qquad \text{where} \qquad R = \mathrm{range}(P),\ K = \ker(P)$$

If V is an inner product space prove that the subspaces must be orthogonal ($R \perp K$) if the projection is self-adjoint, so $P^2 = P = P^*$. $\square$

Matrix realizations of adjoints are easily computed, provided we restrict attention to *orthonormal* bases in both V and W. With respect to arbitrary bases the computation of $[T^*]_{\mathfrak{X}\mathfrak{Y}}$ can be quite a mess.

3.12. Proposition. *Let $T : V \rightarrow W$ be a linear operator between finite dimensional inner product spaces and let $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_j\}$ be orthonormal bases in V, W. Then*

$$[T^*]_{\mathfrak{X}\mathfrak{Y}} = \big([T]_{\mathfrak{Y}\mathfrak{X}}\big)^* \qquad (\textit{taking matrix adjoint on the right}) \tag{51}$$

where A^* is the usual $m \times n$ "adjoint matrix," the conjugate-transpose of A such that $(A^*)_{ij} = \overline{A_{ji}}$ for $A \in \mathrm{M}(n \times m, \mathbb{K})$.

Proof: By definition, the entries of $[T]_{\mathfrak{Y}\mathfrak{X}}$ are determined by the vector identities

$$T(e_i) = \sum_{k=1}^{n} T_{ki}\, f_k \quad \text{which imply} \quad \big(T(e_i), f_j\big)_W = \sum_{k=1}^{n} T_{ki}\, (f_k, f_j)_W = T_{ji},$$

for all i, j. Hence

$$T^*(f_i) = \sum_{k=1}^{n} [T^*]_{ki} e_k \;\Rightarrow\; (T^*(f_i), e_j) = [T^*]_{ji}\ ,$$

from which we see that

$$\begin{aligned}
[T^*]_{ij} &= \big(T^*(f_j), e_i\big)_V \\
&= (f_j, T(e_i))_W \ = \ \overline{(T(e_i), f_j)_W} \ = \ \overline{[T]_{ji}} = \Big([T]^*\Big)_{ij}
\end{aligned}$$

where $(A^*)_{ij} = \overline{A_{ji}}$ for any matrix. $\square$

3.13. Exercise. Let V_N be the restrictions to $[0,1]$ of polynomials $f \in \mathbb{C}[x]$ having degree $\leq N$. Give this $(N+1)$-dimensional space of $\mathcal{C}[0,1]$ the usual L^2 inner product $(f,h)_2 = \int_0^1 f(t)\overline{h(t)}\,dt$ inherited from the larger space of continuous functions. Let $D: V_N \to V_N$ be the differentiation operator

$$D\big(a_0 + a_1 t + a_2 t^2 + \ldots + a_N t^N\big) = a_1 + 2a_2\,t + 3a_3\,t^2 + \ldots + N a_n\,t^{N-1}$$

(a) Is D one-to-one? Onto? What are range(D) and $\ker(D)$?

(b) Determine the matrix $[D]_{\mathfrak{X}\mathfrak{X}}$ with respect to the vector basis $\mathfrak{X} = \{1, x, x^2, \ldots, x^N\}$.

(c) Determine the eigenvalues of $D: V_N \to V_N$ and their multiplicities.

(d) Compute the L^2-inner product $(f,h)_2$ in terms of the coefficients a_k, b_k that determine f and h.

(e) Is D a self-adjoint operator? Skew-adjoint?

3.14. Exercise. If D^* is the adjoint of the differentiation operator $D: V_N \to V_N$, entries D^*_{ij} in its matrix $[D^*]_{\mathfrak{X}}$ with respect to the basis $\mathfrak{X} = \{1, x, x^2, \ldots, x^N\}$ are determined by the vector identities $D^*(x^i) = \sum_{k=0}^N D^*_{ki}\,x^k$. By definition of the adjoint D^* we have

$$\big(x^i\,,\,D(x^j)\big)_2 = \big(D^*(x^i)\,,\,x^j\big)_2 \ = \sum_{k=0}^{N} D^*_{ik}\,\big(x^k\,,\,x^j\big)_2 \qquad \text{for } 0 \leq i,j \leq N$$

and since $\mathfrak{X}$ is a basis these identities implicitly determine the D^*_{ij}. Compute explicit matrices B and C such that $[D^*]_{\mathfrak{X}} = C{\cdot}B^{-1}$. As in the preceding problem, $D(x^k) = kx^{k-1}$ and inner products in V_N are integrals

$$\big(f\,,\,h\big)_2 = \int_0^1 f(x){\cdot}h(x)\,dx$$

for polynomials $f,h \in V_N$.
Hint: Beware: The powers x^i are NOT an orthonormal basis, so you will have to use some algebraic brute force instead of (51). This could get complicated. For something more modest, just compute the action of D^* on the three-dimensional space $V = \mathbb{C}$-span$\{1, t, t^2\}$.

3.15. Exercise. Let $V = \mathcal{C}_c^\infty(\mathbb{R})$ be the space of real-valued functions $f(t)$ on the real line that have continuous derivatives $D^k f$ of all orders, and have "bounded support" – each f is zero off of some bounded interval (which is allowed to vary with f). Because all such functions are "zero near ∞" there is a well defined inner product

$$(f,h)_2 = \int_{-\infty}^{\infty} f(t)\overline{h(t)}\,dt$$

The derivative $Df = df/dt$ is a linear operator on this infinite dimensional space.

(a) Prove that the adjoint of D is *skew-adjoint*, with $D^* = -D$.

(b) Prove that the second derivative $D^2 = d^2/dt^2$ is *self-adjoint*.

Hint: Integration by parts.

Normal and Self-Adjoint Operators. Various classes of operators $T : V \to V$ can be defined on an finite dimensional inner product space.

1. SELF-ADJOINT: $T^* = T$
2. SKEW-ADJOINT: $T^* = -T$
3. UNITARY: $T^*T = I$ (which implies $TT^* = I$ because $T : V \to V$ is one-to-one $\Leftrightarrow$ onto $\Leftrightarrow$ bijective.) Thus "unitary" is equivalent to saying that $T^* = T^{-1}$, at least when V is finite dimensional. (In the infinite-dimensional case we need both identities $TT^* = T^*T = I$ to get $T^* = T^{-1}$.)
4. NORMAL: $T^*T = TT^*$ (T commutes with T^*)

The **spectrum** $\mathrm{sp}_{\mathbb{K}}(T) = \{\lambda \in \mathbb{K} : E_\lambda(T) \neq (0)\}$ of T is closely related to that of T^*.

3.16. Lemma. *On any inner product space*

$$\mathrm{sp}(T^*) = \overline{\mathrm{sp}(T)} = \{\,\overline{\lambda} : \lambda \in \mathrm{sp}(T)\}$$

Proof: If $(T - \lambda I)(v) = 0$ for some $v \neq 0$, then $0 = \det(T - \lambda I) = \det\big([T]_{\mathfrak{X}} - \lambda I_{n\times n}\big)$ for any basis $\mathfrak{X}$ in V. If $\mathfrak{X}$ is an orthonormal basis we get $[T^*]_{\mathfrak{X}} = [T]^*_{\mathfrak{X}} = \overline{[T]^t_{\mathfrak{X}}}$. Then

$$\begin{aligned}
\det\big([T^*]_{\mathfrak{X}} - \overline{\lambda} I_{n\times n}\big) &= \overline{\det\big([T]^{\mathrm{t}}_{\mathfrak{X}} - \lambda I_{n\times n}\big)} \;=\; \overline{\det\big([T]_{\mathfrak{X}} - \lambda I_{n\times n}\big)^{\mathrm{t}}} \\
&= \overline{\det\big([T]_{\mathfrak{X}} - \lambda I_{n\times n}\big)} \;=\; 0
\end{aligned}$$

because

$$\det(A^{\mathrm{t}}) = \det(A) \qquad \text{and} \qquad \det(\overline{A}) = \overline{\det(A)}\ .$$

Hence $\overline{\lambda} \in \mathrm{sp}(T^*)$. Since $T^{**} = T$, we get

$$\mathrm{sp}(T) = \mathrm{sp}(T^{**}) \subseteq \overline{\mathrm{sp}(T^*)} \subseteq \overline{\overline{\mathrm{sp}(T)}} = \mathrm{sp}(T) \quad \square$$

3.17. Exercise. If $A \in \mathrm{M}(n, \mathbb{K})$ prove that its matrix adjoint $(A^*)_{ij} = \overline{A_{ji}}$ has determinant

$$\det(A^*) = \overline{\det(A)}.$$

If $T : V \to V$ is a linear map on an inner product space, prove that $\det(T^*) = \overline{\det(T)}$.

3.18. Exercise. *If $T : V \to V$ is a linear map on an inner product space, show that the characteristic polynomial satisfies*

$$p_{T^*}(\lambda) = \overline{p_T(\overline{\lambda})} \qquad \textit{or equivalently} \qquad p_T(\overline{\lambda}) = \overline{p_T(\lambda)}$$

for all $\lambda \in \mathbb{K}$. In particular,

$$\mathrm{sp}_{\mathbb{K}}(T^*) = \overline{\mathrm{sp}_{\mathbb{K}}(T)} = \{\,\overline{\lambda} : \lambda \in \mathrm{sp}_{\mathbb{K}}(T)\}.$$

Proof: Since $I^* = I$ and $(\lambda I)^* = \overline{\lambda} I$ we get

$$\begin{aligned}
p_{T^*}(\lambda) &= \det(T^* - \lambda I) \;=\; \det\big(T^* - (\overline{\lambda} I)^*\big) \\
&= \det\big(\,(T - \overline{\lambda} I)^*\big) \;=\; \overline{\det(T - \overline{\lambda} I)} \;=\; \overline{p_T(\overline{\lambda})}
\end{aligned}$$

Recall that $\mu \in \mathrm{sp}_{\mathbb{K}}(T) \Leftrightarrow p_T(\mu) = 0$. $\Box$

VI.4. Diagonalization in Inner Product Spaces.

If M is a T-invariant subspace of inner product space V it does not follow that $T^*(M) \subseteq M$. The true relationship between invariance under T and under T^* is:

4.1. Exercise. If V is any inner product space and $T : V \to V$ a linear map, prove that

(a) A subspace $M \subseteq V$ is T-invariant (so $T(M) \subseteq M$) $\Rightarrow$ $M^\perp$ is T^*-invariant.

(b) If $\dim_{\mathbb{K}}(V) < \infty$ (so $M^{\perp\perp} = M$) then $T(M) \subseteq M \Leftrightarrow T^*(M^\perp) \subseteq M^\perp$.

4.2. Proposition. *If $T : V \to W$ is a linear map between finite dimensional inner product spaces, let $R(T) = \mathrm{range}(T)$, $K(T) = \ker(T)$. Then $T^* : W \to V$ and*

$$\begin{aligned} K(T^*) &= R(T)^\perp \text{ in } W \\ R(T^*) &= K(T)^\perp \text{ in } V \end{aligned}$$

In particular if T is self-adjoint then $\ker(T) \perp \mathrm{range}(T)$ *and we have an orthogonal direct sum decomposition $V = K(T) \dot{\oplus} R(T)$.*

Proof: If $w \in W$ then

$$\begin{aligned} T^*(w) = 0 \quad &\Leftrightarrow \quad \big(v, T^*(w)\big)_V = 0 \quad \text{ for all } v \in V \\ &\Leftrightarrow \quad 0 = (v, T^*(w))_V = (T(v), w)_W, \quad \text{ for all } v \in V \\ &\Leftrightarrow \quad w \perp R(T)\ . \end{aligned}$$

Hence $w \in K(T^*)$ if and only if $w \perp R(T)$. The second part follows because $T^{**} = T$ and $M^{\perp\perp} = M$ for any subspace. $\Box$

We will often invoke this result.

Orthogonal Diagonalization. Not all linear operators $T : V \to V$ are diagonalizable, let alone orthogonally diagonalizable, but if V is an inner product space we can always find a basis that at least puts it into *upper-triangular* form, which can be helpful. In fact, this can be achieved via an othonormal basis provided the characteristic polynomial splits into linear factors over $\mathbb{K}$ (always true if $\mathbb{K} = \mathbb{C}$).

4.3. Theorem (Schur Normal Form). *Let $T : V \to V$ be a linear operator on a finite dimensional inner product space over $K = \mathbb{R}$ or $\mathbb{C}$ such that $p_T(x) = \det(T - xI)$ splits over $\mathbb{K}$. Then there are scalars $\lambda_1, \ldots, \lambda_n$ and an orthonormal basis* $\mathfrak{X}$ in V such that

$$[T]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Proof: Work by induction on $n = \dim_{\mathbb{K}}(V)$; the case $n = 1$ is trivial. For $n > 1$, since p_T splits there is an eigenvalue λ in $\mathbb{K}$ and a vector $v_0 \neq 0$ such that $T(v_0) = \lambda v_0$. Then $\overline{\lambda}$ is an eigenvalue for T^*, so there is some $w_0 \neq 0$ such that $T^*(w_0) = \overline{\lambda} w_0$.

Let $M = \mathbb{K}w_0$; this one-dimensional space is T^*-invariant, so $M^\perp$ is invariant under $(T^*)^* = T$ and has dimension $n - 1$. Scale w_0 if necessary to make $\|w_0\| = 1$. By the Induction Hypothesis there there is an orthonormal basis $\mathfrak{X}_0 = \{e_1, ..., e_{n-1}\}$ in $M^\perp$ such that

$$[T|M^\perp]_{\mathfrak{X}_0} = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_{n-1} \end{pmatrix}$$

Then letting $e_n = w_0$ (norm = 1) we get an orthonormal basis for V such that $[T]_{\mathfrak{X}\mathfrak{X}}$ has the form:

$$[T]_{\mathfrak{X}\mathfrak{X}} = \left(\begin{array}{ccc|c} \lambda_1 & & * & c_1 \\ & \ddots & & \vdots \\ 0 & & \lambda_{n-1} & c_{n-1} \\ \hline 0 & & 0 & \lambda_n \end{array} \right)$$

where

$$T(e_n) = T(w_0) = \lambda_n e_n + \sum_{j=1}^{n-1} c_j \, e_j$$

(Remember: $M = \mathbb{K}w_0$ need not be invariant under T.) $\square$

4.4. Exercise. Explain why the diagonal entries in the Schur normal form must be the roots in $\mathbb{K}$ of the characteristic polynomial $p_T(x) = \det(T - xI)$, each counted according to its algebraic multiplicity.

Note: Nevertheless, it might not be possible to find an orthonormal basis such that all occurrences of a particular eigenvalue $\lambda \in \mathrm{sp}_{\mathbb{K}}(T)$ appear in a consecutive string $\lambda, \ldots, \lambda$ on the diagonal. $\square$

Recall that a linear operator $T : V \to V$ on an inner product space is *normal* if it commutes with its adjoint, so that $T^*T = TT^*$. We will eventually show that when $\mathbb{K} = \mathbb{C}$ (or when $\mathbb{K} = \mathbb{R}$ and the characteristic polynomial of T splits into linear factors: $p_T(x) = \prod_{i=1}^{n}(x - \alpha_i)$ with $\alpha_i \in K$), then T is orthogonally diagonalizable *if and only if* T is normal. Note carefully what this does not say: T might be (non-orthogonally) diagonalizable over $\mathbb{K} = \mathbb{C}$ even if T is not normal. This latter issue can only be resolved by determining the pattern of eigenspaces $E_\lambda(T)$ and demonstrating that they span all of V.

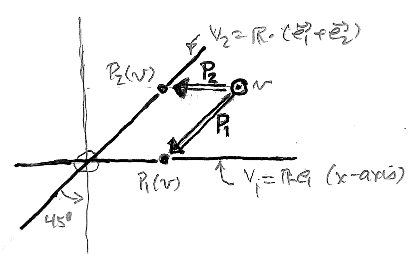

Figure 6.7. The (non-orthogonal) basis vectors $\mathbf{u}_1 = \mathbf{e}_1$ and $\mathbf{u}_2 = \mathbf{e}_1 + \mathbf{e}_2$ in Exercise 4.5.

4.5. Exercise. Let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard orthonormal basis vectors in $V = \mathbb{K}^2$, and consider the ordinary direct sum decomposition

$$V = V_1 \oplus V_2 = \mathbb{K}\mathbf{e}_1 \oplus \mathbb{K}(\mathbf{e}_1 + \mathbf{e}_2) = \mathbb{K}\,\mathbf{f}_1 \oplus \mathbb{K}\,\mathbf{f}_2 \qquad \text{where} \qquad \mathbf{f}_1 = \mathbf{e}_1, \mathbf{f}_2 = \mathbf{e}_1 + \mathbf{e}_2 \; .$$

These subspaces are not orthogonal with respect to the standard Euclidean inner product

$$\big(x_1\mathbf{e}_1 + x_2\mathbf{e}_1 \, , \, y_1\mathbf{e}_1 + y_2\mathbf{e}_2\big) \; = \; x_1\overline{y_1} + x_2\overline{y_2}$$

Define a $\mathbb{K}$-linear map $T : V \to V$, letting

$$T(\mathbf{e}_1) = 2\mathbf{e}_1 \qquad T(\mathbf{e}_1 + \mathbf{e}_2) = \tfrac{1}{2}(\mathbf{e}_1 + \mathbf{e}_2)$$

(see Figure 6.7). Then T is diagonalized by the basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ with $\mathbf{f}_1 = \mathbf{e}_1$ and $\mathbf{f}_2 = \mathbf{e}_1 + \mathbf{e}_2$ (which is obviously not orthonormal), with

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

(a) Determine the action of T on the orthonormal basis vectors $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ and find $[T]_{\mathfrak{X}\mathfrak{X}}$;

(b) Describe the operator T^* by determining its action on the standard orthonormal basis $\mathfrak{X}$, and find $[T^*]_{\mathfrak{X}\mathfrak{X}}$;

(c) Explain why T is not a normal operator on V. Explain why no orthonormal basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ in V can possibly diagonalize T.

Hint: The discussion is exactly the same for $\mathbb{K} = \mathbb{R}$ and $\mathbb{C}$, so assume $\mathbb{K} = \mathbb{R}$ if that makes you more comfortable.

Diagonalizing Self-Adjoint and Normal Operators. We now show that a linear operator $T : V \to V$ on a finite dimensional inner product space is orthogonally diagonalizable if and only if T is normal. First, we analyze the special case of self-adjoint operators ($T^* = T$), which motivates the more subtle proof needed for normal operators.

4.6. Theorem (Diagonalizing Self-Adjoint T). *On a finite dimensional inner product space any self-adjoint linear operator $T : V \to V$ is orthogonally diagonalizable.*

Proof: If $\mu, \lambda \in \mathrm{sp}_{\mathbb{K}}(T)$, we first observe that:

1. If $T = T^*$ all eigenvalues are real, so $\mathrm{sp}_{\mathbb{K}}(T) \subseteq \mathbb{R} + i0$.

   Proof: If $v \in E_\lambda(T)$, $v \neq 0$, we have

   $$\lambda \|v\|^2 = (Tv, v) = (v, T^*v) = (v, Tv) = (v, \lambda v) = \overline{\lambda} \|v^2\|^2$$

   which implies $\lambda = \overline{\lambda}$.

2. If $\lambda \neq \mu$ in $\mathrm{sp}(T)$ the eigenspaces $E_\lambda(T)$ and $E_\mu(T)$ must be orthogonal.

   Proof: If $v \in E_\lambda(T)$, $w \in E_\mu(T)$ then

   $$\lambda(v, w) = (Tv, w) = (v, T^*w) = (v, \mu w) = \overline{\mu}\,(v, w) = \mu\,(v, w)$$

   since eigenvalues are real when $T^* = T$. But $\mu \neq \lambda$, hence $(v, w) = 0$ and $E_\lambda(T) \perp E_\mu(T)$. Thus the linear span $E = \sum E_\lambda(T)$ (which is always a direct sum) is actually an orthogonal sum $E = \dot{\bigoplus}_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$.

3. If $T^* = T$ the span of the eigenspaces is all of V, hence T is orthogonally diagonalizable.

   Proof: If $\lambda \in \mathrm{sp}_{\mathbb{K}}(T)$, then $E_\lambda(T) \neq (0)$ and $M = E_\lambda(T)^\perp$ has $\dim(M) < \dim(V)$. By Exercise 4.1 the orthogonal complement is T^*-invariant, hence T-invariant because $T^* = T$. It is easy (see Exercise 4.7 below) to check that if $W \subseteq V$ is T-invariant and $T^* = T$ on V, then the restriction $T|_W : W \to W$ is self-adjoint on W if one equips W with the restricted inner product from V.

4.7. Exercise. If $T : V \to V$ is linear and $T^* = T$, prove that

$$(T|_W)^* = (T^*|_W)$$

for any T-invariant subspace $W \subseteq V$ equipped with the restricted inner product.

To complete our discussion we show that self-adjoint operators are orthogonally diagonalizable, arguing by induction on $n = \dim(V)$. This is clear if $\dim(V) = 1$, so assume it true whenever $\dim(V) \leq n$ and consider a space of dimension $n+1$. Since all eigenvalues (roots of the characteristic polynomial) are real there is a nontrivial eigenspace $M = E_\lambda(T)$, and if this is all of V we're done: $T = \lambda I$. Otherwise, M has lower dimension and by Exercise 4.7 it has an orthonormal basis that diagonalizes $T|_M$. But $V = M \dot{\oplus} M^\perp$ (an orthogonal direct sum), and $M = E_\lambda$ obviously has an orthonormal basis of eigenvectors. Combining these bases we get an orthonormal diagonalizing basis for all of V. $\square$

We now elaborate the basic properties of normal operators on an inner product space.

4.8. Proposition. *A normal linear operator $T : V \rightarrow V$ on a finite dimensional inner product space has the following properties.*

1. *If $T : V \rightarrow V$ is normal, $\|T(v)\| = \|T^*(v)\|$ for all $v \in V$.*
   Proof: *We have*

$$\begin{aligned} \|T(v)\|^2 &= (Tv, Tv) = (T^*T(v), v) = (TT^*(v), v) \\ &= (T^*v, T^*v) = \|T^*(v)\|^2 \end{aligned}$$

2. *For any $c \in \mathbb{K}$, $T - cI$ is also normal because $(T - cI)^* = T^* - \overline{c}I$ and cI commutes with all operators.*

3. *If $T(v) = \lambda v$ then for the same vector v we have $T^*(v) = \overline{\lambda}v$. In particular, $E_{\overline{\lambda}}(T^*) = E_\lambda(T)$. (This is a much stronger statement than our earlier observation that $\mathrm{sp}_{\mathbb{K}}(T^*) = \overline{\mathrm{sp}_{\mathbb{K}}(T)} = \{\,\overline{\lambda} : \lambda \in \mathrm{sp}_{\mathbb{K}}(T)\}$).*
   Proof: *$(T - \lambda)$ is also normal. Therefore if $v \in V$ and $T(v) = \lambda v$, we have*

$$T(v) = \lambda v \;\Rightarrow\; \big((T-\lambda)^*(T-\lambda)\,v\,,\,v\big) = \|(T-\lambda)\,v\|^2 = 0$$

   which implies that

$$0 = \big((T-\lambda)(T-\lambda)^*\,v\,,\,v\big) = \|(T^* - \overline{\lambda}I)\,v\|^2 \;\Rightarrow\; T^*(v) = \overline{\lambda}v$$

4. *If $\lambda \neq \mu$ in $\mathrm{sp}_{\mathbb{K}}(T)$, then $E_\lambda \perp E_\mu$.*
   Proof: *If v, w are in E_λ, E_μ then*

$$\lambda(v, w) = (\lambda v, w) = (Tv, w) = (v, T^*w) = \big(v, \overline{\mu}\,w\big) = \mu(v, w)$$

   since $T^*(w) = \overline{\mu}\,w$ if $T(w) = \mu\,w$. Therefore $(v, w) = 0$ if $\mu \neq \lambda$.

If $M = \sum_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$ for a normal operator T, it follows that this is a direct sum of orthogonal subspaces $M = \dot{\bigoplus}_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$, and that there is an orthonormal basis $\{e_1, \ldots, e_n\} \subseteq M$ consisting of eigenvectors.

4.9. Corollary. *If $T : V \rightarrow V$ is normal and $\mathbb{K} = \mathbb{C}$ (or if $\mathbb{K} = \mathbb{R}$ and the characteristic polynomial p_T splits over $\mathbb{R}$), there is a diagonalizing orthonormal basis $\{e_i\}$ and V is an orthogonal direct sum $\dot{\bigoplus}_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$.*

Proof: The characteristic polynomial $p_T(x) = \det(T - xI)$ splits in $\mathbb{C}[x]$, so there is an eigenvalue λ_0 such that $T(v_0) = \lambda_0 v_0$ for some $v_0 \neq 0$. The one-dimensional space $M = \mathbb{C}v_0$ is T-invariant, but is also T^*-invariant since $T^*(v_0) = \overline{\lambda_0}v_0$ by (3.). Then

$$T^*(M) \subseteq M \;\Rightarrow\; T^{**}(M^\perp) \;=\; T(M^\perp) \subseteq M^\perp\ .$$

We also have $T^*(M^\perp) \subseteq M^\perp$ because $T(M) \subseteq M \Leftrightarrow T^*(M^\perp) \subseteq M^\perp$. $\square$

4.10. Exercise. If N is a subspace in an inner product space that is invariant under both T and T^*, prove that $T|_N$ satisfies

$$(T|_N)^* = (T^*|_N)$$

Note: Here we do not assume $T^* = T$, which was assumed in Exercise 4.7.

Since $T|_{M^\perp}$ is again a normal operator with respect to the inner product $M^\perp$ inherits from the larger space V, but $\dim(M^\perp) < \dim(V)$, we may argue by induction to get an orthonormal basis of eigenvectors. $\square$

4.11. Theorem (Orthogonal Diagonalization). *Let $T : V \to V$ be a linear operator on a finite dimensional inner product space. Assume that the characteristic polynomial $p_T(x)$ splits over $\mathbb{K}$ (certainly true for $\mathbb{K} = \mathbb{C}$). There is an orthonormal basis that diagonalizes T if and only if T is normal: $T^*T = TT^*$*

Note: It follows that $V = \bigoplus_{\lambda \in \mathrm{sp}_{\mathbb{K}}(T)} E_\lambda(T)$; in particular, the eigenspaces are mutually orthogonal. Once the eigenspaces are determined it is easy to construct the diagonalizing orthonormal basis for T.

Proof: $(\Rightarrow)$ has just been done.

Proof: $(\Leftarrow)$. If there is an orthonormal basis $\mathfrak{X} = \{e_i\}$ that diagonalizes T then

$$[T]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

But $[T^*]_{\mathfrak{X}\mathfrak{X}}$ is the adjoint of the matrix $[T]_{\mathfrak{X}\mathfrak{X}}$,

$$[T^*]_{\mathfrak{X}\mathfrak{X}} = \overline{[T]^{\mathrm{t}}_{\mathfrak{X}\mathfrak{X}}} = \begin{pmatrix} \overline{\lambda_1} & & & 0 \\ & \overline{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \overline{\lambda_n} \end{pmatrix}$$

Obviously these diagonal matrices commute (all diagonal matrices do), so

$$[T^*T]_{\mathfrak{X}\mathfrak{X}} = [T^*]_{\mathfrak{X}\mathfrak{X}}[T]_{\mathfrak{X}\mathfrak{X}} = [T]_{\mathfrak{X}\mathfrak{X}}[T^*]_{\mathfrak{X}\mathfrak{X}} = [TT^*]_{\mathfrak{X}\mathfrak{X}}$$

which implies $T^*T = TT^*$ as operators on V. $\square$

4.12. Example. Let $L_A : \mathbb{C}^2 \to \mathbb{C}^2$ be the multiplication operator determined by

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$

so that $L_A(\mathbf{e}_1) = \mathbf{e}_1$ and $L_A(\mathbf{e}_1 + \mathbf{e}_2) = 2\cdot(\mathbf{e}_1 + \mathbf{e}_2)$, where $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ is the standard orthonormal basis. As we saw in Chapter 2, $[L_A]_{\mathfrak{X}\mathfrak{X}} = A$. But L_A is obviously diagonalizable with respect to the *non-orthonormal* basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ where $\mathbf{f}_1 = \mathbf{e}_1$, $\mathbf{f}_2 = \mathbf{e}_1 + \mathbf{e}_2$. The $\mathbf{f}_i$ are basis vectors for the (one-dimensional) eigenspaces of L_A, which are uniquely determined without any reference to the inner product in $V = \mathbb{C}^2$; if there were an orthonormal basis that diagonalized L_A the eigenspaces would be orthogonal. which they are not. This operator cannot be orthogonally diagonalized with respect to the standard inner product in $\mathbb{C}^2$. $\square$

4.13. Exercise. Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be L_A for the matrix

$$A = A^* = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

in M(2.$\mathbb{C}$). Determine the eigenvalues in $\mathbb{C}$ and the eigenspaces, and exhibit an orthonormal basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ that diagonalizes T.

4.14. Exercise. Prove that $|\lambda| = 1$ for all eigenvalues $\lambda \in \text{sp}(T)$ of a unitary operator (so λ lies on the unit circle if $\mathbb{K} = \mathbb{C}$, or $\lambda = \pm 1$ if $\mathbb{K} = \mathbb{R}$).

4.14A. Exercise. If P is a projection on a finite dimensional vector space (so $P^2 = P$),

(a) Explain why P is diagonalizable, over any field $\mathbb{K}$. What are the eigenvalues and eigenspaces?

(b) Give an explicit example of a projection operator on a finite dimensional inner product space that is not *orthogonally diagonalizable*.

4.14B. Exercise. If P is a projection operator (so $P^2 = P$) on a finite dimensional inner product space, prove that P is a normal operator $\Leftrightarrow$ $K(P) = \ker(P)$ and $R(P) = \text{range}(P)$ are orthogonal subspaces.
Note: ($\Rightarrow$) is trivial since $K(P) = E_{\lambda=0}(P)$ and $R(P) = E_{\lambda=1}(P)$.

4.14C. Exercise. A projection operator P (with $P^2 = P$) on an inner product space is fully determined once we know its kernel $K(P)$ and range $R(P)$, since $V = R(P) \oplus K(P)$. The adjoint P^* is also a projection operator because $(P^*P^*) = (PP)^* = P^*$.

(a) In an inner product space, how are $K(P)$ and $R(P)$ related to $K(P^*)$ and $R(P^*)$?

(b) For the non-orthogonal direct sum decomposition of Exercise VI-4.5 give explicit descriptions of the subspaces $K(P^*)$ and $R(P^*)$. (Find bases for each.)

If $T : V \to V$ is an arbitrary linear operator on an inner product space we showed in IV.3.16 that $\text{sp}(T^*)$ is equal to $\overline{\text{sp}(T)}$; in VI-3.48 we showed that

$$E_{\overline{\lambda}}(T^*) = E_\lambda(T) \qquad (\lambda \in \text{sp}(T))$$

for *normal* operators. Unfortunately the latter property is not true in general.

4.14D. Exercise. If $T : V \to V$ is a linear operator on an inner product space and $\lambda \in \text{sp}(T)$, prove that

(a) $E_{\overline{\lambda}}(T^*) = K\big(T^* - \overline{\lambda}I\big)$ is equal to $R(T - \lambda I)^\perp$.

(b) $\dim E_{\bar{\lambda}}(T^*) = \dim E_\lambda(T)$.

(c) T diagonalizable $\Rightarrow T^*$ is diagonalizable.

As the next example shows, $E_{\overline{\lambda}}(T^*) = K\big(T^* - \overline{\lambda}I\big)$ is not always equal to $E_\lambda(T)$ unless T is normal.

4.14E. Exercise. If $P : V \to V$ is an idempotent operator on a finite dimensional vector space (so $P^2 = P$), explain why P must be diagonalizable over any field. If $P \neq 0$ and $P \neq I$, what are its eigenvalues and its eigenspaces.

4.14F. Exercise. Let P be the projection operator on an inner product space V corresponding to a *non-orthogonal* direct sum decomposition $V = R(P) \oplus K(P)$. Its adjoint P^* is also a projection, onto $R(P^*)$ along $K(P^*)$.

(a) What are the eigenvalues and eigenspaces for P and P^*?

(b) For $\lambda = 1$, is $E_{\overline{\lambda}}(T^*) = K\big(T^* - \overline{\lambda}I\big)$ is equal to $E_\lambda(T)$?

Hint: See Exercise VI-4.14C and D.

Unitary Equivalence of Operators. We say that two operators T, T' on a vector space V are **similar**, written as $T' \sim T$, if there is an invertible linear operator S such that $T' = SAS^{-1}$; this means they are represented by the same matrix $[T']_{\mathfrak{Y}\mathfrak{Y}} = [T]_{\mathfrak{X}\mathfrak{X}}$ with respect to suitably chosen bases in V. We say T' is **unitarily equivalent to** T if there is a unitary operator U such that $T' = UTU^*(= UTU^{-1})$. This relation, denoted $T' \cong T$, is an RST equivalence relation between operators on an inner product space, but is more stringent than mere similarity. We now show $T' \cong T$ if and only if there are *orthonormal* bases $\mathfrak{X}, \mathfrak{Y}$ such that $[T']_{\mathfrak{Y}\mathfrak{Y}} = [T]_{\mathfrak{X}\mathfrak{X}}$.

4.15. Definition. *A* **linear isometry** *is a linear operator $U : V \to W$ between inner product spaces that preserve distances in mapping points from V into W,*

$$\|Uv - Uv'\|_W = \|U(v - v')\|_W = \|v - v'\|_V \ ; \tag{52}$$

in particular $\|U(v)\|_W = \|v\|_V$ for all $v \in V$. Isometries are one-to-one but need not be bijections unless $\dim V = \dim W$ *(see exercises below).*

A linear map $U : V \to W$ is **unitary** *if $U^*U = \mathrm{id}_V$ and $UU^* = \mathrm{id}_W$, which means U is invertible with $U^{-1} = U^*$ (hence* $\dim V = \dim W$*). Obviously the inverse map $U^{-1} : W \to V$ is also unitary. Unitary operators $U : V \to W$ are also isometries since*

$$\|Ux\|_W^2 = \big(Ux, Ux\big)_W = (x, U^*Ux)_V = \|x\|_V^2 \ ,$$

Thus unitary maps are precisely the bijective linear isometries from V to W.

If V is finite dimensional and we restrict attention to the case $V = W$, either of the conditions $UU^* = \mathrm{id}_V$ or $U^*U = \mathrm{id}_V$ implies U is invertible with $U^{-1} = U^*$ because

$$U \text{ one-to-one} \quad \Leftrightarrow \quad U \text{ is surjective} \quad \Leftrightarrow \quad U \text{ is bijective},$$

for any linear operator $U : V \to V$ when $\dim(V) < \infty$.

4.16. Exercise. If V, W are inner product spaces of the same *finite* dimension, explain why there must exist a bijective linear isometry $T : V \to W$. Is T unique? Is the adjoint $T^* : W \to V$ also an isometry?

4.17. Exercise. Let $V = \mathbb{C}^m, W = \mathbb{C}^n$ with the usual inner products. Exhibit examples of linear operators $U : V \to W$ such that

(a) $UU^* = \mathrm{id}_W$ but $U^*U \neq \mathrm{id}_V$.

(b) $U^*U = \mathrm{id}_V$ but $UU^* \neq \mathrm{id}_W$.

Note: This might not be possible for all choices of m, n (for instance $m = n$).

4.18. Exercise. If $m < n$ and the coordinate spaces $\mathbb{K}^m$, $\mathbb{K}^n$ are equipped with the standard inner products, consider the linear operator

$$T : \mathbb{K}^m \to \mathbb{K}^n \qquad T(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, , 0)$$

This is an isometry from $\mathbb{K}^m$ into $\mathbb{K}^n$, with trivial kernel $K(T) = (0)$ and range $R(T) = \mathbb{K}^m \times (\mathbf{0})$ in $\mathbb{K}^n = \mathbb{K}^m \oplus \mathbb{K}^{n-m}$.

(a) Provide an explicit description of the adjoint operator $T^* : \mathbb{K}^n \to \mathbb{K}^m$ and determine $K(T^*), R(T^*)$.

(b) Compute the matrices of $[T]$ and $[T^*]$ with respect to the standard orthonormal bases in $\mathbb{K}^m, \mathbb{K}^n$.

(c) How is the action of T^* related to the subspaces $K(T), R(T^*)$ in $\mathbb{K}^m$ and $R(T), K(T^*)$ in $\mathbb{K}^n$? Can you give a geometric description of this action?

Unitary operators can be described in several different ways, each with its own advantages in applications.

4.19. Theorem. *The statements below are equivalent for a linear operator $U : V \rightarrow W$ between finite dimensional inner product spaces.*

(a) $UU^* = \mathrm{id}_W$ *and* $U^*U = \mathrm{id}_V$ *(so* $U^* = U^{-1}$ *and* $\dim V = \dim W$*).*

(b) *U maps* SOME *orthonormal basis $\{e_i\}$ in V to an orthonormal basis $\{f_i = U(e_i)\}$ in W.*

(c) *U maps* EVERY *orthonormal basis $\{e_i\}$ in V to an orthonormal basis $\{f_i = U(e_i)\}$ in W.*

(d) *U is a surjective isometry, so distances are preserved:*

$$\|U(x) - U(y)\|_W = \|x - y\|_V \qquad \textit{for } x, y \in V$$

(*Then U is invertible and U^{-1} is also an isometry*).

(e) *U is a bijective map that preserves inner products, so that*

$$\big(U(x), U(y)\big)_W = (x, y)_V \qquad \textit{for all } x, y \in V.$$

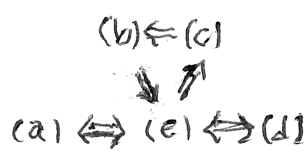

Figure 6.8. The pattern of implications in proving Theorem 4.19.

Proof: We prove the implications shown in Figure 6.8.

Proof: (d) ⇔ (e). Clearly (e) ⇒ (d). For the converse, (d) implies U preserves lengths of vectors, with $VertUx\|_W = \|x\|_V$ for all x. Then by the Polarization Identity for inner products

$$(x, y) = \frac{1}{4} \sum_{k=0}^{3} \frac{1}{i^k} \|x + i^k y\|^2 \ ,$$

so inner products are preserved, proving (d) ⇒ (e) when $\mathbb{K} = \mathbb{C}$; same argument but with only 2 terms if $\mathbb{K} = \mathbb{R}$.

Proof: (e) ⇒ (c) ⇒ (b). These are obvious since "orthonormal basis" is defined in terms of the inner product. For instance if (e) holds and $\mathfrak{X} = \{e_i\}$ is an orthonormal basis in V then $\mathfrak{Y} = \{f_i = U(e_j)\}$ is an orthonormal family in W because

$$(f_i, f_j)_W = \big(U(e_i), U(e_j)\big)_W = (e_i, U^*Ue_j)_V = (e_i, e_j)_V = \delta_{ij} \quad \text{(Kronecker delta).}$$

But

$$\mathbb{K}\text{-span}\{f_j = U(e_j)\} = U\big(\mathbb{K}\text{-span}\{e_j\}\big) = U(V) = W \ ,$$

so $\mathfrak{Y}$ spans W and therefore is a basis.

Proof (a) $\Leftrightarrow$ (e). We have

$$U^*U = \mathrm{id}_V \Leftrightarrow U^*Ux = x \ \text{ for all } x \Leftrightarrow (Ux, Uy)_W = (x, U^*Uy)_V = (x,y)_V$$

for all $x, y \in V$.

Proof: (b) $\Rightarrow$ (e). Given an orthonormal basis $\mathfrak{X} = \{e_i\}$ in V such that the vectors $\mathfrak{Y} = \{f_i = U(e_i)\}$ are an orthonormal basis in W, we may write $x, y \in V$ as $x = \sum_i (x, e_i)\, e_i$, etc. Then

$$U(x) = \sum_i (x, e_i)_V\, U(e_i) = \sum_i (x, e_i)_V\, f_i, \qquad \text{etc} \ ,$$

hence by orthonormality

$$\begin{aligned}(Ux, Uy)_W \ &= \ \Big(\sum_i (x,e_i)_V\, f_i\, , \sum_j (y,e_j)_V\, f_j\Big)_W \ = \ \sum_{i,j} (x,e_i)_V\, \overline{(y,e_j)}_V\, (f_i,f_j)_W \\ &= \ \sum_k (x,e_k)_V\, (e_k,y)_V \ = \ (x,y)_V \quad \Box\end{aligned}$$

Here we applied a formula worth remembering (*Parseval's identity*).

4.20. Lemma (Parseval). *If $x = \sum_i a_i e_i$, $y = \sum b_j e_j$ with respect to an orthonormal basis in a finite dimensional inner product space then $(x,y) = \sum_{k=1}^n a_k \overline{b_k}$. Equivalently, since $a_i = (x, e_i)$, ... etc, we have*

$$(x,y) = \sum_{k=1}^n (x, e_k)(e_k, y) \qquad \textit{for all } x, y$$

in any finite dimensional inner product space, since $\overline{(y, e_k)} = (e_k, y)$. $\quad \Box$

Unitary Operators vs Unitary Matrices.

4.21. Definition. *A matrix $A \in \mathrm{M}(n, \mathbb{K})$ is* **unitary** *if $AA^* = I$ (which holds $\Leftrightarrow AA^* = I \Leftrightarrow A^* = A^{-1}$), where A^* is the* **adjoint matrix** *such that $(A^*)_{ij} = \overline{A_{ji}}$. The set of all unitary matrices is a group since products and inverses of such matrices are again unitary. When $\mathbb{K} = \mathbb{C}$ this is the* **unitary group**

$$\mathrm{U}(n) = \{A \in \mathrm{M}(n, \mathbb{C}) : A^*A = I\} = \{A \in \mathrm{M}(n, \mathbb{C}) : A^* = A^{-1}\} \ .$$

But when $\mathbb{K} = \mathbb{R}$ and $A^* = A^{\mathrm{t}}$ (the transpose matrix), it goes by another name and is called the **orthogonal group***,*

$$\mathrm{O}(n) = \{A \in \mathrm{M}(n, \mathbb{R}) : A^{\mathrm{t}}A = I\} = \{A \in \mathrm{M}(n, \mathbb{R}) : A^{\mathrm{t}} = A^{-1}\}$$

Both groups lie within the **general linear group** *of nonsingular matrices $\mathrm{GL}(n, \mathbb{K}) = \{A : \det(A) \neq 0\}$, and both contain noteworthy subgroups*

$$\begin{aligned}\text{Special Unitary Group: } \ \mathrm{SU}(n) \ &= \ \{A : A^*A = I \textit{ and } \det(A) = +1\} \\ \text{Special Orthogonal Group: } \ \mathrm{SO}(n) \ &= \ \{A : A^{\mathrm{t}}A = I \textit{ and } \det(A) = +1\}\end{aligned}$$

The group SU(3), for instance, seems to be the symmetry group that governs the relations between electromagnetic forces and the weak and strong forces of nuclear physics. As we will see in the next section, SO(3) is the group of rotations in Euclidean space $\mathbb{R}^3$, by any angle about any oriented line through the origin (with a similar interpretation for $\mathrm{SO}(n)$ in higher dimensional spaces $\mathbb{R}^n$).

Given a matrix $A \in \mathrm{M}(n, \mathbb{K})$ it is important to know when the operator $L_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is unitary with respect to the standard inner product. The answer extends the list of condititions (a) – (e) of Theorem VI-4.19 describing when an operator is unitary, and is quite useful in calculations.

4.22. Proposition. *If $A \in \mathrm{M}(n, \mathbb{K})$ the following conditions are equivalent.*

1. *$L_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ is unitary;*
2. *A is a unitary matrix, so $A^*A = AA^* = I$ in $\mathrm{M}(n, \mathbb{K})$*
3. *The rows in A form an orthonormal basis in $\mathbb{K}^n$.*
4. *The columns in A form an orthonormal basis in $\mathbb{K}^n$.*

Proof: With respect to the standard basis $\mathfrak{X} = \{e_1, \ldots, e_n\}$ in $\mathbb{K}^n$ we know that $[L_A]_{\mathfrak{X}} = A$, but since $\mathfrak{X}$ is an orthonormal basis we also have $[(L_A)^*]_{\mathfrak{X}} = [L_A]^*_{\mathfrak{X}} = A^*$ (the adjoint matrix), by Exercise 3.12. Next observe that

$$L_{A^*} = \big(L_A\big)^* \qquad \text{as operators on } \mathbb{K}^n$$

(This may sound obvious, but it actually needs to be proved keeping in mind how the various "adjoints" are defined – see Exercise 4.24 below.) Then we get

$$\begin{aligned} A^*A = I \quad &\Leftrightarrow \quad \mathrm{id}_{\mathbb{K}^n} = [L_{A^*A}]_{\mathfrak{X}} = [L_{A^*}]_{\mathfrak{X}} \cdot [L_A]_{\mathfrak{X}} = [(L_A)^*]_{\mathfrak{X}} \cdot [L_A]_{\mathfrak{X}} \\ &\Leftrightarrow \quad (L_A)^* L_A = \mathrm{id}_{\mathbb{K}^n} \ \Leftrightarrow \ (L_A \text{ is a unitary operator}) \quad , \end{aligned}$$

proving (1.) $\Leftrightarrow$ (2.)

By definition of row-column matrix multiplication we have

$$\delta_{ij} = (AA^*)_{ij} = \sum_k A_{ik}(A^*)_{kj} = \sum_k A_{ik}\overline{A_{jk}} = \Big(\mathrm{Row}_i(A)\,,\,\mathrm{Row}_j(A)\Big)_{\mathbb{K}^n}$$

This says precisely that the rows are an orthonormal basis with respect to the standard inner product in $\mathbb{K}^n$. Thus (2.) $\Leftrightarrow$ (3.), and similarly $A^*A = I \Leftrightarrow$ the columns form an orthonormal basis in $\mathbb{K}^n$. $\square$

A similar criterion allows us to decide when a general linear operator is unitary.

4.23. Proposition. *A linear operator $T : V \rightarrow V$ on a finite dimensional inner product space is unitary $\Leftrightarrow$ its matrix $A = [T]_{\mathfrak{X}}$ with respect to any orthonormal basis is a unitary matrix (so $AA^* = A^*A = I$).*

Proof: For any orthonormal basis we have

$$I = [\mathrm{id}_V]_{\mathfrak{X}} = [T^*T]_{\mathfrak{X}} = [T^*]_{\mathfrak{X}}\,[T]_{\mathfrak{X}} = \big([T]_{\mathfrak{X}}\big)^*\,[T]_{\mathfrak{X}} = A^*A$$

and similarly $AA^* = I$, so A is a unitary matrix.

Conversely, if $A = [T]_{\mathfrak{X}}$ is a unitary matrix we have

$$\begin{aligned} (Te_i, Te_j) \quad &= \quad \Big(\sum_k A_{ki}\,e_k\,,\,\sum_\ell A_{\ell j}\,e_j\Big) \;=\; \sum_{k,\ell} A_{ki}\overline{A_{\ell j}}\,\delta_{k\ell} \\ &= \quad \sum_k A_{ki}(A^*)_{jk} \;=\; (AA^*)_{ji} \;=\; \delta_{ji} = (e_i, e_j) \end{aligned}$$

Thus T maps orthonormal basis $\mathfrak{X}$ to a new orthonormal basis $\mathfrak{Y} = \{T(e_i)\}$, and T is unitary by Theorem 4.19(c). $\square$

4.24. Exercise. Prove that $L_{A^*} = \big(L_A\big)^*$ when $\mathbb{K}^n$ is given the standard inner product.
Hint: Show that $(A^*\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A\mathbf{y})$ for the standard inner product.

This remains true when A in an $n \times m$ matrix, $L_A : \mathbb{K}^m \rightarrow \mathbb{K}^n$, and $(L_A)^* : \mathbb{K}^n \rightarrow \mathbb{K}^m$.

4.24A. Exercise. If $A \in \mathrm{M}(n, \mathbb{C})$ give a careful proof that $A^*A = I \Leftrightarrow AA^* = I$.

4.25. Exercise. Given two orthonormal bases $\{e_i\}$, $\{f_j\}$ in finite dimensional inner product spaces V, W of the same dimension, construct a unitary operator $U : V \rightarrow W$ such that $U(e_i) = f_i$ for all i.

Change of Orthonormal Basis.

If $T : V \rightarrow V$ is a linear operator on a finite dimensional inner product space, and we know its matrix $[T]_{\mathfrak{X}\mathfrak{X}}$ with respect to one orthonormal basis, what is its matrix realization with respect to a different orthonormal basis $\mathfrak{Y}$?

4.26. Definition. *Matrices $A, B \in \mathrm{M}(n, \mathbb{K})$ are* **unitarily equivalent***, indicated by writing $A \cong B$, if there is some unitary matrix $S \in \mathrm{M}(n, \mathbb{K})$ such that $B = SAS^* = SAS^{-1}$.* $\square$

4.27. Theorem (Change of Orthonormal Basis). *If $\mathfrak{X} = \{e_i\}$ and $\mathfrak{Y} = \{f_j\}$ are orthonormal bases in a finite dimensional inner product space and $T : V \rightarrow V$ is any linear operator, the corresponding matrices $A = [T]_{\mathfrak{X}\mathfrak{X}}$ and $B = [T]_{\mathfrak{Y}\mathfrak{Y}}$ are unitarily equivalent: there is some unitary matrix S such that*

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^* = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1} \qquad \textit{where } S = [\mathrm{id}_V]_{\mathfrak{Y}\mathfrak{X}} = [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}^{-1} = [\mathrm{id}_V]_{\mathfrak{X}\mathfrak{Y}}^{*} \tag{53}$$

The identity (53) *remains true if the transition matrix S is multiplied by any scalar such that $|\lambda|^2 = \lambda\overline{\lambda} = 1$.*

Proof: For *arbitrary* vector bases $\mathfrak{X}, \mathfrak{Y}$ in V we have $[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}^{-1}$ and

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} \cdot [T]_{\mathfrak{X}\mathfrak{X}} \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1} \tag{54}$$

where $S = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}$ is given by the vector identities $e_i = \mathrm{id}(e_i) = \sum_j S_{ji}f_j$. But we also have $e_i = \sum_j (e_i, f_j)\, f_j$, so $S_{ij} = (e_j, f_i)$, for $1 \leq i, j \leq n$.

The transition matrix S in (54) is unitary because $S_{ij} = (e_j, f_i) \Rightarrow$

$$\begin{aligned}
\big(\mathrm{Row}_i(S)\,,\, \mathrm{Row}_j(S)\big)_{\mathbb{K}^n} &= \sum_k S_{ik}\overline{S_{jk}} \;=\; \sum_k (e_k, f_i)\,\overline{(e_k, f_j)} \\
&= \sum_k (f_j, e_k)\,(e_k, f_i) \;=\; (f_j, f_i) = \delta_{ij}
\end{aligned}$$

by Parseval's identity. Then $S^* = S^{-1} = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}^{-1} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}$ by Theorem 4.22, and

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^* = S\,[T]_{\mathfrak{X}\mathfrak{X}}S^{-1} \quad \square$$

We conclude that the various matrix realizations of T with respect to orthonormal bases in V are related by unitary equivalence (similarity modulo a unitary matrix) rather than similarity modulo a matrix that is merely invertible. Unitary equivalence is therefore a more stringent condition on two matrices than similarity (as defined in Chapter V).

Elements U in the unitary group $\mathrm{U}(n)$ act on matrix space $X = \mathrm{M}(n, \mathbb{C})$ by conjugation, sending

$$A \mapsto UAU^{-1} = UAU^*\,.$$

This group action $\mathrm{U}(n) \times X \rightarrow X$ partitions X into disjoint orbits

$$\mathcal{O}_A = \mathrm{U}(n) \cdot A = \{UAU^* : U \in \mathrm{U}(n)\}\ ,$$

which are the the **unitary equivalence classes** in matrix space. There is a similar group action $\mathrm{O}(n) \times \mathrm{M}(n, \mathbb{R}) \to \mathrm{M}(m, \mathbb{R})$ of the orthogonal group on real matrices. Recall that the *similarity class* of an $n \times n$ matrix A is its orbit $\mathrm{GL}(n, \mathbb{K}) \cdot A = \{EAE^{-1} : E \in \mathrm{GL}(n, \mathbb{K})\}$ under the action of the general linear group $\mathrm{GL}(n, \mathbb{K}) = \{A : \det(A) \neq 0\}$, which is considerably larger than $\mathrm{U}(n)$ or $\mathrm{O}(n)$ and has larger orbits.

Diagonalization over $\mathbb{K} = \mathbb{C}$: A Summary. We recall that the spectra $\mathrm{sp}_{\mathbb{C}}(T)$ of operators over $\mathbb{C}$ and their adjoints have the following properties.

1. For any T, $\mathrm{sp}(T^*) = \overline{\mathrm{sp}(T)}$ and $\dim E_{\overline{\lambda}}(T^*) = \dim E_\lambda(T)$. But as we will see in 4.14E below, the $\bar{\lambda}$ eigenspace $E_{\bar{\lambda}}(T^*)$ is not always equal to $E_\lambda(T)$ unless T is normal.

2. If $T = T^*$ then T is orthogonally diagonalizable, and all eigenvalues are *real* because $T(v) = \lambda \Rightarrow$
$$\lambda\|v\|^2 = (T(v), v) = \big(v, T^*(v)\big) = (v, \lambda v) = \overline{\lambda}\|v\|^2$$

3. If T is unitary then all eigenvalues satisfy $|\lambda| = 1$ (they lie on the unit circle in $\mathbb{C}$), because
$$\begin{aligned} T(v) = \lambda \cdot v \quad &\Rightarrow \quad \|v\|^2 = (T^*Tv, v) = (Tv, Tv) = (\lambda v, \lambda v) = |\lambda|^2 \cdot \|v\|^2 \\ &\Rightarrow \quad |\lambda|^2 = 1 \text{ if } v \neq 0 \end{aligned}$$

4. If T is skew-adjoint, so $T^* = -T$, then all eigenvalues are pure imaginary because
$$\lambda\|v\|^2 = (Tv, v) = (v, T^*v) = (v, -T(v)) = (v, -\lambda v) = -\overline{\lambda}\|v\|^2$$
Consequently, $\overline{\lambda} = -\lambda$ and $\lambda \in 0 + i\mathbb{R}$ in $\mathbb{C}$.

5. A general normal operator is orthogonally diagonalizable, but there are no restrictions on the pattern of eigenvalues.

In Theorem 4.11 we proved the following necessary and sufficient condition for a linear operator on a complex inner product space to be diagonalizable.

4.28. Theorem (Orthogonal Diagonalization). *A linear operator $T : V \to V$ on a finite dimensional complex inner product space is orthogonally diagonalizable $\Leftrightarrow$ T is normal (so $T^*T = TT^*$).* $\square$

VI.5. Some Operators on Real Inner Product Spaces: Reflections, Rotations and Rigid Motions.

All this works over $\mathbb{K} = \mathbb{R}$ except that in this context unitary operators are referred to as **orthogonal transformations**. The corresponding matrices $A = [T]_{\mathfrak{X},\mathfrak{X}}$ with respect to orthonormal bases satisfy $A^{\mathrm{t}}A = I = AA^{\mathrm{t}}$, so $A^{\mathrm{t}} = A^{-1}$ in $\mathrm{M}(n, \mathbb{R})$. An orthogonal transformation might not have enough real eigenvalues to be diagonalizable, which happens $\Leftrightarrow$ the eigenspaces $E_\lambda(T)$ $(\lambda \in \mathbb{R})$ fail to span V. In fact there might not be any real eigenvalues at all. For example, if $R_\theta =$ (counterclockwise rotation about origin by θ radians) in $\mathbb{R}^2$, and if θ is not an integer multiple of π, then with respect to the standard $\mathbb{R}$-basis $\mathfrak{X} = \{e_1, e_2\}$ we have
$$[R_\theta]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$
whose complex eigenvalues are $e^{i\theta}$ and $e^{-i\theta}$; there are no real eigenvalues if $\theta \neq n\pi$, even though R_θ is a normal operator. (A rotation by $\theta \neq n\pi$ radians cannot send a vector

$v \neq 0$ to a scalar multiple of itself.)

The Group of Rigid Motions $\mathrm{M}(n)$**.** **Rigid motions** on $\mathbb{R}^n$ are the bijective maps $\rho : \mathbb{R}^n \to \mathbb{R}^n$ that preserve distances between points,

$$\|\rho(x) - \rho(y)\| = \|x - y\| \quad \text{for all } x, y \ .$$

We DO NOT assume ρ is linear. The rigid motions form a group $\mathrm{M}(n)$ under composition; it includes two important subgroups

1. TRANSLATIONS: Operators $T = \{t_b : b \in \mathbb{R}^n\}$ where

   $$t_b(x) = x + b \qquad \text{for all } x \in \mathbb{R}^n \quad (b \in \mathbb{R}^n \text{ fixed})$$

   Under the bijective map $\phi : \mathbb{R}^n \to T$ with $\phi(t) = t_b$ we have $\phi(s+t) = \phi(s) \circ \phi(t)$ and $\phi(0) = \mathrm{id}_{\mathbb{R}^n}$. Obviously translations are isometric mappings since

   $$\| t_b(x) - t_b(y)\| = \|(x+b) - (y-b)\| = \|x - y\| \qquad \text{for all } b \text{ and } x, y$$

   but they are NOT linear operators on $\mathbb{R}^n$ (unless $b = 0$) because the zero element does not remain fixed: $t_b(0) = b$.

2. LINEAR ISOMETRIES: Operators $H = \{L_A : A \in \mathrm{O}(n)\}$ where $L_A(x) = A{\cdot}x$ and A is any orthogonal real $n \times n$ matrix (so A is invertible with $A^{\mathrm{t}} = A^{-1}$).

Although rigid motions need not be linear operators, it is remarkable that they are nevertheless simple combinations of a linear isometry (an orthogonal linear mapping on $\mathbb{R}^n$) and a translation operator.

$$\rho(x) = \big(t_b \circ L_A\big)(x) = A{\cdot}x + b \qquad (b \in \mathbb{R}^n, A \in \mathrm{O}(n)) \tag{55}$$

for all $x \in \mathbb{R}^n$. In particular, any rigid motion $\rho : \mathbb{R}^n \to \mathbb{R}^n$ that leaves the origin fixed is *automatically linear.*

5.1 Proposition. *If $\rho : \mathbb{R}^n \to \mathbb{R}^n$ is a rigid motion that fixes the origin (so $\rho(\mathbf{0}) = \mathbf{0}$), then ρ is in fact a* LINEAR *operator on $\mathbb{R}^n$, $\rho = L_A$ for some $A \in \mathrm{O}(n)$. In general, every rigid motion is a composite of the form* (55).

Proof: The second statement is immediate from the first, for if ρ moves the origin to $b = \rho(\mathbf{0})$, the operation $t_{-b} \circ \rho$ is a rigid motion that fixes the origin, and $\rho = t_b \circ (t_{-b} \circ \rho)$.

To prove the first assertion, let $\{\mathbf{e}_j\}$ be the standard orthonormal basis in $\mathbb{R}^n$ written as column vectors and let $\mathbf{e}'_j = \rho(\mathbf{e}_j)$. Since $\rho(\mathbf{0}) = 0$ lengths are preserved because $\|\rho(\mathbf{x})\| = \|\rho(\mathbf{x}) - \rho(\mathbf{0})\| = \|\mathbf{x}\|$, and then inner products are also preserved because

$$\begin{aligned} -2\left(\rho(\mathbf{x}), \rho(\mathbf{y})\right) &= \|\rho(\mathbf{x}) - \rho(\mathbf{y})\|^2 - \|\rho(\mathbf{x})\|^2 - \|\rho(\mathbf{y})\|^2 \\ &= \|\mathbf{x} - \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2 \ = \ -2\,(\mathbf{x}, \mathbf{y}) \end{aligned}$$

Hence the images $\mathbf{e}'_i = \rho(\mathbf{e}_i)$ of the standard basis vectors are also an orthonormal basis.

Now let A be the matrix whose i^{th} column is $\mathbf{e}'_i = \mathrm{col}(0, \ldots, 1, \ldots, 0)$, so $L_A(\mathbf{e}_i) = A \cdot \mathbf{e}_i = \mathbf{e}'_i$. Then A is in $\mathrm{O}(n)$, L_A and $(L_A)^{-1} = L_{A^{-1}}$ are both *linear* orthogonal transformations on $\mathbb{R}^n$, and the product $L_A^{-1} \circ \rho$ as a rigid motion that fixes each $\mathbf{e}_i$ as well as the zero vector. But any such motion must be the identity map. In fact if $\mathbf{x} \in \mathbb{R}^n$ then $(\mathbf{x}, \mathbf{e}_i) = \big(\rho(\mathbf{x}), \rho(\mathbf{e}_i)\big) = \big(\rho(\mathbf{x}), \mathbf{e}'_i\big)$, and since $\mathbf{e}'_i = \mathbf{e}_i$ we get

$$x_i = (\mathbf{x}, \mathbf{e}_i) = \big(\rho(\mathbf{x}), \mathbf{e}'_i\big) = (\rho(\mathbf{x}), \mathbf{e}_i) = x'_i$$

for all i. Hence $\mathbf{x}' = \rho(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x}$, as claimed. $\square$

Every rigid motion on $\mathbb{R}^n$,

$$T(x) = A{\cdot}x + t_b = (t_b \circ L_A) \qquad \text{with } A \in \mathrm{O}(n) \text{ and } b \in \mathbb{R}^n$$

has two components, an orthogonal *linear* map L_A and a translation t_b. Rigid motions are of two types, **orientation preserving** and **orientation reversing**. Translations always preserve orientation of geometric figures, so the nature of a rigid motion T is determined by its linear component L_A, which preserves orientation if $\det(A) > 0$ and reverses it if $\det(A) < 0$. As a simple illustration, consider the matrices (with respect to the standard basis $\mathfrak{X}$ in $\mathbb{R}^2$) of a rotation about the origin R_θ (orientation preserving), and a reflection r_y across the y-axis (orientation reversing).

$$[R_\theta]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad\qquad [r_y]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Rotation: R_θ, $\det[R_\theta] = +1$ Reflection: r_y, $\det[r_y] = -1$

Rotations and reflections can be described in terms of the inner product in $\mathbb{R}^n$.

5.2 Example (Reflections in Inner Product Spaces). *If V is a finite dimensional inner product space over $\mathbb{R}$, a* **hyperplane** *in V is any vector subspace M with $\dim(M) = n-1$ (so M has "codimension 1" in V). This determines a reflection of vectors across M.*

Discussion: Since $V = M \dot{\oplus} M^\perp$ (orthogonal direct sum) every vector v splits uniquely as $v = v_\parallel + v_\perp$ (with "parallel component" $v_\parallel \in M$, and $v_\perp \in M^\perp$). By definition, reflection r_M across M is the (linear) operator that reverses the "perpendicular component" $v_\perp$, so that

$$r_M(v_\parallel + v_\perp) = v_\parallel - v_\perp = v - 2{\cdot}v_\perp \tag{56}$$

as shown in Figure 6.9.

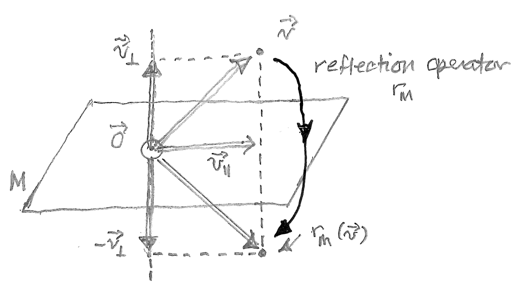

Figure 6.9. Geometric meaning of reflection r_M across an $(n-1)$-dimensional hyperplane in an n-dimensional inner product space over $\mathbb{R}$

Now, let $\{e_1, ..., e_{n-1}\}$ be an orthonormal basis in the subspace M and let e_n be $v_\perp$ renormalized to make $\|e_n\| = 1$, so $M^\perp = \mathbb{R}e_n$. We have seen that

$$v_\parallel = \sum_{k=1}^{n-1} (v, e_k)\, e_k \ ,$$

so $v_\perp = v - v_\parallel = c \cdot e_n$ for some $c \in \mathbb{R}$. But in fact $c = (v, e_n)$ because

$$c = (ce_n, e_n) = (v - v_\parallel, e_n) = (v, e_n) + 0$$

This yields an important formula involving only the inner product.

$$r_M = v_{\|} - v_{\perp} = \big(v_{\|} + v_{\perp}\big) - 2{\cdot}v_{\perp} = v - 2(v, e_n){\cdot}e_n \tag{57}$$

Note: we need $\|e_n\| = 1$ to make this work. $\square$

5.3 Exercise. Show that (57) implies the following properties for any reflection.

(a) $r_M \circ r_M = id_V$, so r_M is its own inverse;

(b) $\det(r_M) = -1$, so all reflections are orientation-reversing.

(c) M is the set of fixed points $\mathrm{Fix}(r_M) = \{x : r_M(x) = x\}$. $\square$

5.4 Exercise. Prove that every reflection r_M on an inner product space preserves distances,

$$\|r_M(x) - r_M(y)\| = \|x - y\|$$

for all $x, y \in V$.

5.5 Exercise. If M is a hyperplane in a finite dimensional real inner product space V and $b \notin M$, the translate $b + M$ (a coset in V/M) is an $n-1$ dimensional hyperplane parallel to M (but is not a vector subspace). Explain why the operation that reflects vectors across $M' = b + M$ must be the rigid motion $T = t_b \circ r_M \circ t_{-b}$.
Hint: Check that $T^2 = T$ and that the set of fixed points $\mathrm{Fix}(T) = \{v \in V : T(v) = v\}$ is precisely M'.

In another direction, we have Euler's famous geometric characterization of orientation preserving orthogonal transformations $L_A : \mathbb{R}^3 \to \mathbb{R}^3$ with $A^{\mathrm{t}}A = I = AA^{\mathrm{t}}$ in $\mathrm{M}(3,\mathbb{R})$ and $\det(A) > 0$. In fact, $\det(A) = +1$ since $A^{\mathrm{t}}A = I$ implies $\big(\det(A)\big)^2 = 1$, so $\det(A) = \pm 1$ for $A \in \mathrm{O}(n)$.

5.6 Theorem (Euler). *Let $A \in \mathrm{SO}(3) = \{A \in \mathrm{M}(3,\mathbb{R}) : A^{\mathrm{t}}A = I$ and $\det(A) = 1\}$. If $A \neq I$ then $\lambda = 1$ is an eigenvalue such that $\dim_{\mathbb{R}}\big(E_{\lambda=1}\big) = 1$. If $v_0 \neq 0$ in $E_{\lambda=1}$ and $\ell = \mathbb{R}\,v_0$ there is some angle $\theta \notin 2\pi\mathbb{Z}$ such that*

$$L_A = R_{\ell,\theta} = (\textit{rotation by } \theta \textit{ radians about the oriented line } \ell \textit{ through the origin}).$$

(*Rotations by a positive angle are determined by the usual "right hand rule," with your thumb pointing in the direction of v_0*).

Proof: The characteristic polynomial $p_T(x)$ for $T = L_A$ has real coefficients. Regarded as a polynomial $p_T \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$, its complex roots are either real or occur in conjugate pairs $z = x + iy$, $\overline{z} = x - iy$ with $y \neq 0$. Since $\mathrm{degree}(p_T) = 3$ there must be at least one real root λ. But because $T = L_A$ is unitary its complex eigenvalues have $|\lambda| = 1$, because if $v \neq 0$ in E_λ,

$$\|v\|^2 = \big(T(v), T(v)\big) = (\lambda v, \lambda v) = |\lambda|^2\,\|v\|^2 \;\Rightarrow\; |\lambda|^2 = 1\ .$$

If λ is real the only possibilities are $\lambda = \pm 1$. The real roots cannot all be -1, for then $\det(T) = (-1)^3 = -1$ and we require $\det(T) = +1$. Thus $\lambda = 1$ is an eigenvalue, and we will see below that $\dim_{\mathbb{R}}\big(E_{\lambda=1}\big) = 1$.

If $v_0 \neq 0$ in $E_{\lambda=1}$, let $M = \mathbb{R}v_0$. Then $M^\perp$ is 2-dimensional and is invariant under both T and $T^* = T^{-1}$. Furthermore (see Exercise 5.7) the restriction $T|_{M^\perp}$ is a unitary (= orthogonal) transformation on the 2-dimensional space $M^\perp$ equipped with the inner product it inherits from $\mathbb{R}^3$. If we fix an orthonormal basis $\{f_1, f_2\}$ in $M^\perp$ and let $f_0 = v_0/\|v_0\|$, we obtain an orthonormal basis for $\mathbb{R}^3$. The matrix A of $T|_{M^\perp}$ with respect to $\mathfrak{X}_0 = \{f_1, f_2\}$ is in

$$\mathrm{SO}(2) = \{A \in \mathrm{M}(2,\mathbb{R}) : A^{\mathrm{t}}A = I \text{ and } \det(A) = 1\}$$

because the matrix of T with respect to the orthonormal basis $\mathfrak{X} = \{f_0, f_1, f_2\}$ is

$$[T]_{\mathfrak{X}} = \left[\begin{array}{cc} 1 & 0 \\ 0 & A \end{array} \right] .$$

Thus $A \in \mathrm{SO}(2)$ because $1 = \det\big([T]_{\mathfrak{X}}\big) = 1 \cdot \det(A)$. As noted below in Exercise VI-5.8, if $A \in \mathrm{SO}(2)$ the rows form an orthonormal basis for $\mathbb{R}^2$ and so do the columns, hence there exist $a, b \in \mathbb{R}$ such that

$$a^2 + b^2 = 1 \qquad \text{and} \qquad A = \left(\begin{array}{cc} a & -b \\ b & a \end{array} \right)$$

It follows easily that there is some $\theta \in \mathbb{R}$ such that

$$A = \left(\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right) .$$

This is the matrix $A = [R_\theta]_{\mathfrak{X}_0}$ of a rotation by θ radians about the origin in $M^\perp$, so $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a rotation $R_{\ell,\theta}$ by θ radians about the axis $\ell = \mathbb{R}v_0$. $\square$

We cannot have $\theta \in 2\pi\mathbb{Z}$ because then $T = \mathrm{id}$ is not really a rotation about any well-defined axis); that's why we required $A \neq I$ in the theorem.

5.7 Exercise. Let $T : V \rightarrow V$ be a linear operator on a finite dimensional inner product space, and M a subspace that is invariant under both T and T^*. Prove that the restriction $(T|M) : M \rightarrow M$ is unitary with respect to the inner product M inherits from V.
Hint: Recall Exercise 4.10.

5.8 Exercise. If $A = [a, b; c, d] \in \mathrm{M}(2, \mathbb{R})$ verify that $A^{\mathrm{t}}A = I \Leftrightarrow$ the rows of A are an orthonormal basis in $\mathbb{R}^2$, so that

$$a^2 + b^2 = 1 \qquad c^2 + d^2 = 1 \qquad ac + bd = 0$$

If, in addition we have

$$\det(A) = ad - bc = +1$$

prove that $c = -b$, $d = a$ and $a^2 + b^2 = 1$, and then explain why there is some $\theta \in \mathbb{R}$ such that $a = \cos(\theta)$ and $b = -\sin(\theta)$.
Note: Thus $L_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a counterclockwise rotation about the origin by θ radians.
Hint: For the last step, think of $a^2+b^2 = 1$ in terms of a right triangle whose hypoteneuse has length $= 1$.

5.9 Exercise. Consider the linear map $L_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ for the matrix

$$A = \left(\begin{array}{cc} -1 & 1 \\ 1 & 1 \end{array} \right) \qquad \text{in } \mathrm{O}(2)$$

What is the geometric action of L_A? If a rotation, find the angle θ; if not, show that the set of fixed points for L_A is a line through the origin L, and $L_A =$ (reflection across L).

5.10 Exercise. If $A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right]$ is in $\mathrm{O}(2)$ and has $\det(A) = -1$,

1. Prove that $L_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is reflection across some line ℓ through the origin.
2. Explain why

$$a^2 + b^2 = 1 \qquad c^2 + d^2 = 1 \qquad ac + bd = 0 \qquad \det(A) = ad - bc = -1$$

and then show there is some θ such that $A = \left[\begin{array}{cc} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{array} \right]$

Note: The preceding matrix is *not* a rotation matrix since $\det(A) = -1$. The angle θ determined here is related to the angle between the line of reflection ℓ and the $+x$-axis. **Hints:** The map $L_A : \mathbb{C}^2 \to \mathbb{C}^2$ is unitary, and in particular is orthogonally diagonalizable. What are the possible patterns of complex eigenvalues (counted according to multiplicity), and how do they relate to the requirement that $\det(A) = -1$?

VI.6. Spectral Theorem for Vector and Inner Product Spaces.

If V is a vector space over a field $\mathbb{K}$ (not necessarily an inner product space), and if $T : V \to V$ is diagonalizable over $\mathbb{K}$, then $V = \bigoplus_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$ (an ordinary direct sum) – see Proposition II-3.9. This decomposition determines projection operators $P_\lambda = P_\lambda^2$ of V onto $E_\lambda(T)$ along the complementary subspaces $\bigoplus_{\mu \neq \lambda} E_\mu(T)$. The projections $P_\lambda = P_\lambda(T)$ have the following easily verified properties:

1. $P_\lambda^2 = P_\lambda$
2. $P_\lambda P_\mu = P_\mu P_\lambda = 0$ if $\lambda \neq \mu$ in $\mathrm{sp}(T)$;
3. $I = \sum_\lambda P_\lambda$;

Condition (1.) simply reflects the fact that P_λ is a projection operator. Each $v \in V$ has a unique decomposition $v = \sum_\lambda v_\lambda$ with $v_\lambda \in E_\lambda(T)$, and (by definition) $P_\lambda(v) = v_\lambda$. Property (3.) follows from this. For (2.) write $v = \sum_\lambda v_\lambda$ and consider distinct $\alpha \neq \beta$ in $\mathrm{sp}(T)$. Then

$$P_\alpha P_\beta(v) = P_\alpha P_\beta\Big(\sum_\lambda v_\lambda\Big) = P_\alpha(v_\beta) = 0 \qquad \text{(since } \alpha \neq \beta\text{)}$$

and similarly for $P_\beta P_\alpha$. The operators $\{P_\lambda : \lambda \in \mathrm{sp}_{\mathbb{K}}(T)\}$ are the **spectral projections** associated with the diagonalizable operator T.

Now let V be an inner product space. If T is orthogonally diagonalizable we have additional information regarding the spectral projections $P_\lambda(T)$:

4. The eigenspaces $E_\lambda(T)$ are orthogonal, $E_\lambda \perp E_\mu$ if $\lambda \neq \mu$, and $V = \dot{\bigoplus}_\lambda E_\lambda(T)$ is an orthogonal direct sum decomposition.
5. The P_λ are *orthogonal* projections, hence they are *self-adjoint* in addition to having the preceeding properties, so that $P_\lambda^2 = P_\lambda = P_\lambda^*$.

In this setting we can prove useful facts relating diagonalizability and eigenspaces of an operator $T : V \to V$ and its adjoint T^*. These follow by recalling that there is a natural isomorphism between any finite dimensional inner product space V and its dual space V^*, as explained in Lemma VI-3.2. Therefore given any basis $\mathfrak{X} = \{e_1, \ldots, e_n\}$ in V there exists *within* V a matching basis $\mathfrak{X}' = \{f_1, \ldots, f_n\}$ that is "dual to" $\mathfrak{X}$ in the sense that

$$(e_i\,,\, f_j) = \delta_{ij} \qquad \text{(Kronecker delta)}$$

These paired bases can be extremely useful in comparing properties of T with those of its adjoint T^* .

6.1 Exercise. Let $\mathfrak{X} = \{e_1, \ldots, e_n\}$ be an arbitrary basis (not necessarily orthonormal) in a finite dimensional inner product space V.

(a) Use induction on n to prove that there exist vectors $\mathfrak{Y} = \{f_1, \ldots, f_n\}$ such that $(e_i, f_j) = \delta_{ij}$.

(b) Explain why the f_j are uniquely determined and a basis for V.

Note: If the initial basis $\mathfrak{X}$ is orthonormal then $f_i = e_i$ and the result trivial; we are interested in *arbitrary* bases in an inner product space.

6.1A Exercise. Let V be an inner product space and T a linear operator that is diagonalizable in the ordinary sense, but not necessarily orthogonally diagonalizable. Prove that

(a) The adjoint operator T^* is diagonalizable. What can you say about its eigenvalues and eigenspaces?

(b) If T is *orthogonally* diagonalizable so is T^*.

Hint: If $\{e_i\}$ diagonalizes T what does the "dual basis" $\{f_j\}$ of Exercise 6.1 do for T^*?

6.1B Exercise. If V is a finite dimensional inner product space and $T : V \to V$ is diagonalizable in the ordinary sense, prove that the spectral projections for T^* are the adjoints of those for T:

$$P_{\overline{\lambda}}(T^*) = \big(P_\lambda(T)\big)^* \quad \text{for all } \lambda \in \mathrm{sp}(T)$$

Hint: Use VI-6.1A and dual diagonalizing bases; we already know $\mathrm{sp}(T^*) = \overline{\mathrm{sp}(T)}$.
Note: $\big(P_\lambda(T)\big)^*$ might differ from $P_\lambda(T)$.

We now procede to prove the spectral theorem and examine its many applications.

6.2 Theorem (The Spectral Theorem). *If a linear operator $T : V \to V$ is diagonalizable on a finite dimensional vector space V over a field $\mathbb{K}$, and if $\{P_\lambda : \lambda \in \mathrm{sp}_{\mathbb{K}}(T)\}$ are the spectral projections, then T has the following description in terms of those projections*

$$T = \sum_{\lambda \in \mathrm{sp}(T)} \lambda \cdot P_\lambda \tag{58}$$

If $f(x) = \sum_{k=0} c_k x^k \in \mathbb{K}[x]$ is any polynomial the operator $f(T) = \sum_{k=0} c_k T^k$ takes the form

$$f(T) = \sum_{\lambda \in \mathrm{sp}(T)} f(\lambda) \cdot P_\lambda \tag{59}$$

In particular, the powers T^k are diagonalizable, with $T^k = \sum_{\lambda \in \mathrm{sp}(T)} \lambda^k \cdot P_\lambda$.

If we define the map $\Phi : \mathbb{K}[x] \to \mathrm{Hom}_{\mathbb{K}}(V, V)$ from polynomials to linear operators on V, letting $\Phi(1) = I$ and

$$\Phi(f) = \sum_{k=0} c_k T^k \qquad \textit{for} \qquad f(x) = \sum_{k=0} c_k x^k \quad ,$$

then Φ is linear and a homomorphism of associative algebras over $\mathbb{K}$, so that

$$\Phi(fg) = \Phi(f) \circ \Phi(g) \qquad \textit{for } f, g \in \mathbb{K}[x] \tag{60}$$

Finally, $\Phi(f) = 0$ (the zero operator on V) if and only if $f(\lambda) = 0$ for each $\lambda \in \mathrm{sp}_{\mathbb{K}}(T)$. Thus $\Phi(f) = \Phi(g)$ if and only if f and g take same values on the spectrum $\mathrm{sp}(T)$, so many polynomials $f \in \mathbb{K}[x]$ can yield the same operator $f(T)$.

Note: This is all remains true for orthogonally diagonalizable operators on an inner product space, but in this case we have the additional property

$$\Phi(\,\overline{f}\,) = \Phi(f)^* \qquad \text{(adjoint operator)} \tag{61}$$

where $\overline{f}(x) = \sum_{k=0} \overline{c_k} x^k$ and $\bar{c}$ is the complex conjugate of c. $\square$

Proof of (6.2): If $v \in V$ decomposes as $v = \sum_\lambda v_\lambda \in \bigoplus_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$, then

$$\begin{aligned} T(v) = T\big(\sum_\lambda v_\lambda\big) &= \sum_\lambda \lambda \cdot v_\lambda = \sum_\lambda \lambda \cdot P_\lambda(v) \\ &= \Big(\sum_{\lambda \in \mathrm{sp}(T)} \lambda \cdot P_\lambda\Big)\, v \end{aligned}$$

for all $v \in V$, proving (58). Then $T^k = \sum_\lambda \lambda^k P_\lambda$ becomes

$$T^k(v) = T^k\big(\sum_\lambda v_\lambda\big) = \sum_\lambda T^k v_\lambda$$

But $T(v_\lambda) = \lambda v_\lambda \;\Rightarrow$

$$T^2(v_\lambda) = T(\lambda \cdot v_\lambda) = \lambda^2 v_\lambda, \quad T^3(v_\lambda) = \lambda^3 v_\lambda, \quad \text{etc}$$

so if $v = \sum_\lambda v_\lambda$ we get

$$T^k(v) = \sum_\lambda \lambda^k v_\lambda = \sum_\lambda \lambda^k P_\lambda(v) = \Big(\sum_\lambda \lambda^k P_\lambda\Big)\, v$$

for all $v \in V$. Noting that the powers T^k and the sum $f(T)$ are linear operators, (59) follows: For any $f(x) = \sum_k c_k x^k$ we have

$$\begin{aligned} f(T)(v) &= f(T)\big(\sum_\lambda v_\lambda\big) = \sum_\lambda f(T)(v_\lambda) \\ &= \sum_\lambda \Big(\sum_k c_k T^k\Big)(v_\lambda) = \sum_\lambda \sum_k c_k T^k(v_\lambda) \\ &= \sum_\lambda \sum_k c_k \lambda^k v_\lambda = \sum_\lambda \Big(\sum_k c_k \lambda^k\Big)\, v_\lambda \\ &= \sum_\lambda f(\lambda)\, v_\lambda = \sum_\lambda f(\lambda) P_\lambda(v) \\ &= \Big(\sum_\lambda f(\lambda) P_\lambda\Big)\, v \qquad \text{for all } v \in V \end{aligned}$$

Thus $f(T) = \sum_\lambda f(\lambda) P_\lambda$ as operators on V.

When $f(x)$ is the constant polynomial $f(x) = 1$ we get

$$\sum_{\lambda \in \mathrm{sp}(T)} f(\lambda) P_\lambda = \sum_\lambda P_\lambda = I$$

as expected. Linearity of Φ is easily checked by applying the operators on either side to a typical vector. As for the multiplicative property, let $f = \sum_{k=0} a_k x^k$ and $g = \sum_{\ell \geq 0} b_\ell x^\ell$, so $fg = \sum_{k,\ell=0} a_k b_\ell\, x^{k+\ell}$. First notice that the multiplicative property holds for monomials $f = x^k$, $g = x^\ell$ because

$$\begin{aligned} \Phi(x^k)\Phi(x^\ell) &= \Big(\sum_{\lambda \in \mathrm{sp}(T)} \lambda^k P_\lambda\Big) \cdot \Big(\sum_{\mu \in \mathrm{sp}(T)} \mu^\ell P_\mu\Big) \\ &= \sum_{\lambda,\mu} \lambda^k \mu^\ell\, P_\lambda P_\mu = \sum_\lambda \lambda^{k+\ell} P_\lambda \\ &= \Phi(x^{k+\ell}) \end{aligned}$$

($P_\lambda P_\mu = 0$ if $\lambda \neq \mu$, and $P_\lambda^2 = P_\lambda$). Then use linearity of Φ to get

$$\begin{aligned}
\Phi(fg) &= \Phi\Big(\sum_{k,\ell=0} a_k b_\ell \, x^{k+\ell}\Big) = \sum_{k,\ell=0} a_k b_\ell \Phi(x^{k+\ell}) \\
&= \sum_{k,\ell} a_k b_\ell \Phi(x^k)\Phi(x^\ell) = \Big(\sum_{k=0} a_k \Phi(x^k)\Big)\cdot\Big(\sum_{\ell=0} b_\ell \Phi(x^\ell)\Big) \\
&= \Phi(f)\circ\Phi(g)
\end{aligned}$$

That completes the proof of Theorem 6.2. $\square$

Although the operator $\Phi(f) = \sum_\lambda f(\lambda) P_\lambda$ was defined for polynomials in $\mathbb{K}[x]$, this sum involves only the values of f on the finite subset $\mathrm{sp}(T) \subseteq \mathbb{K}$, so it makes sense for *all* functions $h : \mathrm{sp}(T) \to \mathbb{K}$ whether or not they are defined off of the spectrum, or related in any way to polynomials. Thus the spectral decomposition of T determines a linear map

$$\Phi : \mathcal{E} \to \mathrm{Hom}_{\mathbb{K}}(V, V) \qquad \Phi(h) = \sum_{\lambda \in \mathrm{sp}(T)} h(\lambda) P_\lambda \tag{62}$$

defined on the larger algebra $\mathcal{E} \supseteq \mathbb{F}[x]$ whose elements are arbitrary functions h from $\mathrm{sp}_{\mathbb{K}}(T) \to \mathbb{K}$. The same argument used for polynomials shows that the extended version of Φ is again a homomorphism between associative algebras, as in (60). Incidentally, the Lagrange Interpolation formula tells us that any $h(x)$ in $\mathcal{E}$ is the restriction of some (nonunique) polynomial $f(x)$, so that

$$\Phi(h) = \Phi\big(f\big|_{\mathrm{sp}(T)}\big) = \Phi(f)$$

All this applies to matrices as well as operators since a matrix is diagonalizable $\Leftrightarrow$ the left multiplication operator $L_A : \mathbb{K}^m \to \mathbb{K}^m$ on coordinate space is diagonalizable.

We can now define "functions $h(T)$ of an operator" for a much broader class of functions than polynomials, as in the next examples.

6.3 Example. If a diagonalizable linear operator $T : V \to V$ over $\mathbb{C}$ has spectral decomposition $T = \sum_\lambda \lambda \cdot E_\lambda$, we can define such operators $h(T)$ as

1. $|T| = \sum_\lambda |\lambda|\, P_\lambda = h(T)$ taking $h(z) = |z|$.
2. $e^T = \sum_\lambda e^\lambda\, P_\lambda = h(T)$ taking $h(z) = e^z = \sum_{n=0}^{\infty} z^n/n!$
3. $\sqrt{T} = \sum_\lambda \lambda^{1/2} P_\lambda$ assigning any (complex) determination of $h(z) = \sqrt{z}$ at each point in the spectrum. Thus there are r^2 possible operator square roots if T has r distinct eigenvalues that are all nonzero. As in Exercise 6.4 below, every such "square root" has the property $h(T)^2 = T$.
4. The **indicator function** of a finite subset $E \subseteq \mathbb{C}$ is

$$1_E(z) = \begin{cases} 1 & \text{if } z \in E \\ 0 & \text{otherwise} \end{cases}$$

Then by (60), $1_E(T)$ is a projection operator with $1_E(T)^2 = 1_E(T)$. In particular, if $E = \{\lambda_1, \ldots, \lambda_s\} \subseteq \mathrm{sp}(T)$ we have

$$1_E(T) = \sum_{\lambda \in E} P_\lambda = \bigoplus_{i=1}^{s} P_{\lambda_i} \ \Big(\text{projection onto } \bigoplus_{i=1}^{s} E_{\lambda_i}(T)\Big) \ .$$

We get $1_E(T) = I$ if $E = \mathrm{sp}(T)$, and if $E = \{\lambda_0\}$ is a single eigenvalue we recover the individual spectral projections: $1_E(T) = P_{\lambda_0}$. $\square$

6.4 Exercise. Let $T : V \to V$ be a diagonalizable linear operator over any ground field $\mathbb{K}$. If T is invertible ($\lambda = 0$ not in the spectrum), explain why

$$h(T) = \sum_{\lambda \in \mathrm{sp}(T)} \frac{1}{\lambda} \cdot P_\lambda \qquad \Big(h(x) = \frac{1}{x} \text{ for } x \neq 0\Big)$$

is the usual inverse T^{-1} of T.
Hint: Show $T \circ h(T) = h(T) \circ T = I$

Similarly we have

$$T^{-k} = \left(T^{-1}\right)^k = \sum_\lambda \frac{1}{\lambda^k} P_\lambda$$

for $k = 0, 1, 2 \ldots$, with $T^0 = I$.

6.5 Exercise. Prove (61) when V is an inner product space over $\mathbb{C}$. (There is nothing to prove when $\mathbb{K} = \mathbb{R}$.)

6.6 Exercise. Prove that a normal operator $T : V \to V$ on a finite dimensional inner product space over $\mathbb{C}$ is self adjoint if and only if its spectrum is real: $\mathrm{sp}_{\mathbb{C}}(T) \subseteq \mathbb{R} + i0$.
Note: We already explained ($\Rightarrow$); you do ($\Leftarrow$).

6.7 Exercise. If T is diagonalizable over $\mathbb{R}$ or $\mathbb{C}$, prove that

$$e^T = \sum_{\lambda \in \mathrm{sp}(T)} e^\lambda P_\lambda$$

is the same as the linear operator given by the exponential series

$$e^T = \sum_{k=0}^{\infty} \frac{1}{k!} T^k$$

. **Note:** If T has spectral decomposition $T = \sum_\lambda \lambda \cdot P_\lambda$ then $T^k = \sum_\lambda \lambda^k P_\lambda$. To discuss convergence of the operator-valued exponential series in Exercise VI-6.7, fix a basis $\mathfrak{X} \subseteq V$. Then a sequence of operators converges, with $T_n \to T$ as $n \to \infty$, if and only if the corresponding matrices converge entry-by-entry, $[T_n]_{\mathfrak{X}\mathfrak{X}} \to [T]_{\mathfrak{X}\mathfrak{X}}$ as $n \to \infty$ in matrix space, as described in Chapter II, Section 5.3. The partial sums of a series converge to a limit

$$S_n = I + T + \frac{1}{2!} T^2 + \ldots + \frac{1}{n!} T^n \;\to\; S_0 \ ,$$

$\Leftrightarrow$ $(S_n)_{ij} \to (S_0)_{ij}$ in $\mathbb{C}$ for all $1 \leq i, j \leq N$. $\square$

6.8 Exercise. Let $S \in \mathrm{M}(2, \mathbb{C})$ be a *symmetric* matrix, so $A^{\mathrm{t}} = A$

(a) Is $L_A : \mathbb{C}^2 \to \mathbb{C}^2$ diagonalizable in the ordinary sense?

(b) Is $L_A : \mathbb{C}^2 \to \mathbb{C}^2$ orthogonally diagonalizable when $\mathbb{C}^2$ is given the usual inner product?

Prove or provide a counterexample.
Note: If we take $\mathbb{R}$ instead of $\mathbb{C}$ the answer is "yes" for both (a) and (b) because $A^* = A^{\mathrm{t}}$ when $\mathbb{K} = \mathbb{R}$. Recall that $\left(L_A\right)^* = L_{A^*}$ for the standard inner product on $\mathbb{C}^2$ – see Exercise VI-3.9. Self-adjoint matrices are diagonalizable over both $\mathbb{R}$ and $\mathbb{C}$, but we are not assuming $A = A^*$ here, only $A = A^{\mathrm{t}}$.

6.9 Exercise. Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be the operator $T = L_A$ for

$$A = \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$$

Explain why T is self-adjoint with respect to the standard inner product $(z, w) = z_1\overline{w_1} + z_2\overline{w_2}$ on $\mathbb{C}^2$. Then determine

(a) The spectrum $\mathrm{sp}_{\mathbb{C}}(T) = \{\lambda_1, \lambda_2\}$;

(b) The eigenspaces $E_\lambda(T)$ and find an orthonormal basis $\{f_1, f_2\}$ in $\mathbb{C}^2$ that diagonalize T. Then

(c) Find a unitary matrix $U^*U = I$ such that

$$UAU^* = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

where $\mathrm{sp}(T) = \{\lambda_1, \lambda_2\}$.

6.10 Exercise (Uniquess of Spectral Decompositions). Suppose $T : V \to V$ is diagonalizable on an arbitrary vector space (not necesarily an inner product space), so $T = \sum_{i=1}^r \lambda_i P_{\lambda_i}$ where $\mathrm{sp}(T) = \{\lambda_1, \ldots, \lambda_r\}$ and P_{λ_i} is the projection onto the λ_i-eigenspace. Now suppose $T = \sum_{j=1}^s \mu_j Q_j$ is some other decomposition such that

$$Q_j^2 = Q_j \neq 0 \qquad Q_jQ_k = Q_kQ_j = 0 \quad \text{if } j \neq k \qquad \sum_{j=1}^s Q_j = I$$

and $\{\mu_1, \ldots, \mu_r\}$ are distinct. Prove that

(a) $r = s$ and if the μ_j are suitably relabeled we have $\mu_i = \lambda_i$ for $1 \leq i \leq r$.

(b) $Q_i = P_{\lambda_i}$ for $1 \leq i \leq r$.

Hint: First show $\{\mu_1, \ldots, \mu_s\} \subseteq \{\lambda_1, \ldots, \lambda_r\} = \mathrm{sp}(T)$; then relabel.

Here is another useful observation about spectra of diagonalizable operators.

6.11 Lemma (Spectral Mapping Theorem). If $T : V \to V$ is a diagonalizable operator on a finite dimensional vector space, and $f(x)$ is any function $f : \mathrm{sp}(T) \to \mathbb{C}$, then $f(T)$ is diagonalizable and

$$\mathrm{sp}(f(T)) = f\big(\mathrm{sp}(T)\big) = \{f(\lambda) : \lambda \in \mathrm{sp}(T)\} \ .$$

Proof: We have shown that $T = \sum_{\lambda \in \mathrm{sp}(T)} \lambda P_\lambda$ where the P_λ are the spectral projections determined by the direct sum decomposition $V = \bigoplus_\lambda E_\lambda(T)$. Then $f(T) = \sum_\lambda f(\lambda)P_\lambda$, from which it is obvious that $f(T)v = f(\lambda)v$ for $v \in E_\lambda(T)$; hence $f(T)$ is diagonalizable. The eigenvalues are the values $f(\lambda)$ for $\lambda \in \mathrm{sp}(T)$, but notice that we might have $f(\lambda) = f(\mu)$ for different eigenvalues of T. To get the eigenspace $E_\alpha(f(T))$ we must add together all these spaces

$$E_\alpha(f(T)) = \bigoplus\nolimits_{\{\lambda : f(\lambda) = \alpha\}} E_\lambda(T) \qquad \text{for every } \alpha \in f(\mathrm{sp}(T)) \ .$$

The identity is now clear. $\square$

As an extreme illustration, if $f(z) \equiv 1$ then $f(T) = I$ and $\mathrm{sp}(T) = \{1\}$.

VI.7. Positive Operators and the Polar Decomposition.

If $T : V \to W$ but $V \neq W$ one cannot speak of "diagonalizing T." (What would "eigenvector" and "eigenvalue" mean in that context?) But we can still seek other decompositions of T as a product of particularly simple, easily understood operators. Even when $V = W$ one might profitably explore such options if T fails to be diagonalizable – diagonalization is not the only useful decomposition of a linear operator.

When $V = W$ and V is an inner product space over $\mathbb{R}$ or $\mathbb{C}$, all self-adjoint (or normal) operators are orthogonally diagonalizable, and among them the *positive operators* are particularly simple.

7.1 Definition. *A linear operator $T : V \to V$ on an inner product space is* **positive** *if*

$$\text{(i)}\quad T^* = T \qquad \textit{and} \qquad \text{(ii)}\quad (Tv, v) \geq 0 \textit{ for all } v \in V.$$

It is **positive definite** *if $(Tv, v) = 0$ only when $v = 0$. We write $T \geq 0$ or $T > 0$, respectively, to indicate these possibilities. A matrix $A \in \mathrm{M}(n, \mathbb{C})$ is said to be positive (or positive definite) if the multiplication operator $L_A : \mathbb{C}^n \to \mathbb{C}^n$ is positive (positive definite) with respect to the usual inner product, so that $(Av, v) \geq 0$ for all v.*

Note that self-adjoint projections $P^2 = P^* = P$ are examples of positive operators, and sums of positive operators are again positive (but not linear combinations unless the coefficients are posiive).

If T is diagonalizable, $\mathrm{sp}(T) = \{\lambda_1, \ldots, \lambda_r\}$, and if $T = \sum_i \lambda_i P_{\lambda_i}$ is the spectral decomposition, a self-adjoint operator is positive $\Leftrightarrow$ $\lambda_i \geq 0$ for all i, so $\mathrm{sp}(T) \subseteq [0, +\infty) + i0$. [In fact if $T \geq 0$ and $v_i \in E_{\lambda_i}$, we have $(Tv_i, v_i) = \lambda_i \|v_i\|^2 \geq 0$. Conversely if all $\lambda_i \geq 0$ and $v = \sum_{i=1}^r v_i$ we get $(Tv, v) = \sum_{i,j}(Tv_i, v_j)$, and since $E_{\lambda_i} \perp E_{\lambda_j}$ for $i \neq j$ this reduces to $\sum_i (Tv_i, v_i) = \sum_i \lambda_i \|v_i\|^2 \geq 0$.]

If T is positive *definite* then $\lambda_i = 0$ cannot occur in $\mathrm{sp}(T)$ and T is invertible, with

$$T^{-1} = \sum_i \frac{1}{\lambda_i} P_{\lambda_i} \qquad \text{(also a positive definite operator)}\ .$$

Positive Square Roots. If $T \geq 0$ there is a positive square root (a positive operator $S \geq 0$ such that $S^2 = T$), namely

$$\text{(63)} \qquad \sqrt{T} = \sum_i \sqrt{\lambda_i}\, P_{\lambda_i} \qquad \left(\sqrt{\lambda_i} = \text{the nonnegative square root of } \lambda_i \geq 0\right)\ ,$$

which is also denoted by $T^{1/2}$. This is a square root because

$$S^2 = \sum_{i,j} \sqrt{\lambda_i}\sqrt{\lambda_j}\, P_{\lambda_i} P_{\lambda_j} = \sum_i \lambda_i P_{\lambda_i} = T$$

where $P_{\lambda_i} P_{\lambda_j} = \delta_{ij} \cdot P_{\lambda_i}$. Notice that the spectral decompositions of T and $\sqrt{T}$ involve the same spectral projections P_{λ_i}; obviously the eigenspaces match up too, because $E_{\lambda_i}(T) = E_{\sqrt{\lambda_i}}(\sqrt{T})$ for all i.

Subject to the requirement that $S \geq 0$, this square root is *unique*, as a consequence of uniqueness of the spectral decomposition on any vector space (see Exercise VI-6.10)

7.2 Exercise. Use uniqueness of spectral decompositions to show that the positive square root operator $\sqrt{T} = \sum_i \sqrt{\lambda_i} P_{\lambda_i}$ defined above is *unique* – i.e. if $A \geq 0$ and $B \geq 0$ and $A^2 = B^2 = T$ for some $T \geq 0$, then $A = B$.

Positivity of $T : V \to V$ has an interesting connection with the exponential map on matrices $\mathrm{Exp} : \mathrm{M}(n, \mathbb{C}) \to \mathrm{M}(n, \mathbb{C})$,

$$\mathrm{Exp}(A) = e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$$

We indicated in Section V.3 that commuting matrices A, B satisfy the Exponent Law $e^{A+B} = e^A \cdot e^B$, with $e^0 = I$. In particular all matrices in the range of *Exp* are invertible, with $\left(e^A\right)^{-1} = e^{-A}$.

7.3 Exercise. Let $\mathcal{P}$ be the set of *positive definite* matrices A in $\mathrm{M}(n, \mathbb{C})$, which are all self-adjoint by definition of $A > 0$. Let $\mathcal{H}$ be the set of all self-adjoint matrices in $\mathrm{M}(n, \mathbb{C})$, which is a vector subspace *over* $\mathbb{R}$ but not over $\mathbb{C}$ since iA is skew-adjoint if A is self-adjoint. Prove that

(a) The exponential matrix e^H is positive and invertible for self-adjoint matrices H.

(b) The exponential map $\mathrm{Exp} : \mathcal{H} \to \mathcal{P}$ is a *bijection*.

Hint: Explain why $\left(e^A\right)^* = e^{A^*}$ and then use the Exponent Law applied to matrices $e^{tA},\ t \in \mathbb{R}$ (you could also invoke the spectral theorem).

It follows that every positive definite matrix $A > 0$ has a unique self-adjoint *logarithm* $\mathrm{Log}(A)$ such that

$$\begin{aligned} \mathrm{Exp}\big(\mathrm{Log}(A)\big) &= A \qquad \text{for } A \in \mathcal{P} \\ \mathrm{Log}\big(e^H\big) &= H \qquad \text{for } H \in \mathcal{H}\ \ , \end{aligned}$$

namely the inverse of the bijection $\mathrm{Exp} : \mathcal{H} \to \mathcal{P}$. In terms of spectral decompositions,

$$\begin{aligned} &\mathrm{Log}(T) \text{ of a positive definite } T \text{ is } \mathrm{Log}(T) = \sum_i \mathrm{Log}(\lambda_i) P_{\lambda_i} \text{ if } T = \sum_i \lambda_i P_{\lambda_i} \\ &\mathrm{Exp}(H) \text{ of a self-adjoint matrix } H \text{ is } e^H = \sum_i e^{\mu_i} Q_{\lambda_i} \text{ if } H = \sum_i \mu_i Q_{\mu_i} \end{aligned}$$

When $V = W$ the *unitary* operators $U : V \to V$ are another well-understood family of (diagonalizable) operators on an inner product space. They are particularly interesting and easy to understand because they correspond to the possible choices of orthonormal bases in V. Every unitary U is obtained by specifying a pair of orthonormal bases $\mathfrak{X} = \{e_i\}$ and $\mathfrak{Y} = \{f_j\}$ and defining U to be the unique linear map such that

$$U\Big(\sum_{i=1}^n c_i e_i\Big) = \sum_{j=1}^n c_j f_j \qquad (\text{arbitrary } c_i \in \mathbb{C})$$

Polar Decompositions. The positive operators $P \geq 0$ and unitary operators U on an inner product space provide a natural *polar decomposition* $T = U \cdot P$ of any linear operator $T : V \to V$. In its simplest form (when T is invertible) it asserts that any *invertible* map T has a *unique* factorization

$$T = U \cdot P \qquad \begin{cases} U : V \to V & \text{unitary (a bijective isometry of } V) \\ P : V \to V & \text{positive definite, invertible} = e^H \text{ with } H \text{ self-adjoint} \end{cases}$$

Both factors are orthogonally diagonalizable (U because it is normal and P because it is self-adjoint), but the original operator T need not itself be diagonalizable over $\mathbb{C}$, let alone orthogonally diagonalizable.

We will develop the polar decomposition first for an invertible operator $T : V \to V$ since that proof is particularly transparent. We then address the general result (often referred to as the *singular value decomposition* when it is stated for matrices). This involves operators that are not necessarily invertible, and may be maps $T : V \to W$ between quite different inner product spaces. The positive component $P : V \to V$ is still unique but the unitary component U may be nonunique (in a harmless sort of way). The "*singular values*" of T are the eigenvalues $\lambda_i \geq 0$ of the positive component P.

7.4 Theorem (Polar Decomposition I). *Let V be a finite dimensional inner product space over $\mathbb{C}$. Every invertible operator $T : V \to V$ has a unique decomposition $T = U{\cdot}P$ where*

$$U \in \mathrm{U}(n) = (\textit{the group of unitary operators } U^* = U^{-1})$$

$$P \in \mathcal{P} = (\textit{invertible positive definite operators } P > 0)$$

By Exercise 7.3 we can also write $T = U \cdot e^H$ for a unique self-adjoint operator $H \in \mathcal{H}$.

This is the linear operator (or matrix) analog of the polar decomposition

$$z = |z|e^{i\theta} = r \cdot e^{i\theta} \qquad \text{with } r > 0 \text{ and } \theta \text{ real (so } |e^{i\theta}| = 1)$$

for nonzero complex numbers. If we think of "positive definite" = "positive," "self-adjoint" as "real," and "unitary" = "absolute value 1," the analogy with the polar decomposition $z = re^{i\theta}$ of a nonzero complex number z is clear.

Some Preliminary Remarks. If $T : V \to W$ is a linear map between two inner product spaces, its *absolute value* $|T|$ is the linear map from $V \to V$ determined in the following way.

The product T^*T maps $V \to V$ and is a positive operator because

$$(T^*T)^* = T^*T^{**} = T^*T \qquad (T^* : W \to V \text{ and } T^{**} = T \text{ on } V)$$

$$\big(T^*Tv, v\big) = (Tv, Tv) = \|Tv\|^2 \geq 0 \quad \text{ for } v \in V$$

Thus T^*T is self-adjoint and has a spectral decomposition $T^*T = \sum_i \lambda_i P_{\lambda_i}$, with eigenvalues $\lambda_i \geq 0$ and self-adjoint projections $P_{\lambda_i} : V \to E_{\lambda_i}(T^*T)$ onto orthogonal subspaces. The **absolute value** $|T| : V \to V$ is then defined as the unique positive square root

$$|T| = \big(T^*T\big)^{1/2} = \sum_i \sqrt{\lambda_i}\, P_{\lambda_i} \ ,$$

whose spectral decomposition involves the same projections that appeared in T^*T. For any linear operator $T : V \to W$ we have $T^*T = |T|^2$ and hence

$$(64) \qquad \|Tv\|_W^2 = (T^*Tv, v)_V = \big(|T|^2 v, v\big)_V = \big\|\, |T|(v)\,\big\|_V^2 \qquad \text{for all } v \in V \ .$$

Thus $|T|(v) \in V$ and $Tv \in W$ have the same norm for every $v \in V$. It follows from (64) that T, T^*T, and $|T|$ have the same kernel because

$$\begin{aligned} Tv = 0 \Rightarrow T^*T(v) = 0 \quad &\Rightarrow \quad (T^*Tv, v) = \big(|T|^2(v), v\big) = \big\|\, |T|(v)\,\big\|^2 = 0 \\ &\Rightarrow \quad |T|(v) = 0 \Rightarrow Tv = 0 \quad \text{(by (64))}\ , \end{aligned}$$

Thus the kernels coincide

$$(65) \qquad K(T) = K(T^*T) = K(|T|)$$

even if the ranges differ, and one of these operators is invertible if and only if they all are. In particular $|T|$ is positive *definite* on V ($|T| > 0$) if and only if $T : V \to W$ is invertible. (Comparisons between T and $|T|$ do not follow from spectral theory because T itself need not be diagonalizable, even if $V = W$.)

Proof of VI-7.4: The proof in the invertible case is simple. For any linear operator $T : V \to V$ we have $T^*T = |T|^2$ and have seen in (64) that $|T|(v)$ and Tv always have the same norm. When T is invertible, so is $|T|$ and we have $R(T) = R(|T|) = V$. The identities (64) determine a bijective isometry $U : V \to V$ that sends $T(v) \mapsto |T|(v)$ for all v, as indicated in Figure 6.10. This map is also linear because $U = T \circ |T|^{-1}$ is a composite of linear operators on V. Thus when T is invertible the desired polar decomposition is

$$U \circ |T| = \big(T \circ |T|^{-1}\big) \circ |T| = T$$

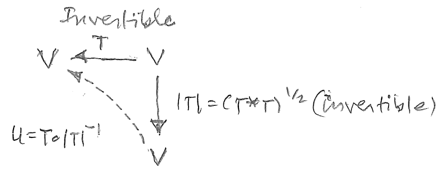

Figure 6.10. The maps involved in defining $|T| : V \to V$ for an invertible map $T : V \to W$ between two inner product spaces. In the discussion we show that the positive operator $|T| = (T^*T)^{1/2}$ is invertible and $R(T^*T) = R\big(|T|\big) = R(T) = V$ when T is invertible. The induced bijection $U = T \circ |T|^{-1} : V \to V$ is a bijective linear isometry (a unitary map of $V \to V$) and the polar decomposition of T is $U \cdot |T|$.

As for uniqueness (valid only in the invertible case), suppose $T = UP = U_0P_0$ with U, U_0 unitary and P, P_0 positive definite. Then $P^* = P, P_0^* = P_0$, and $T^* = P^*U^* = PU^* = P_0^*U_0^* = P_0U_0^*$ since the positive components are self-adjoint; hence

$$P^2 = PU^*UP = PU^*\big(PU^*\big)^* = P_0U_0^*U_0P_0 = P_0^2$$

Now $P^2 = P^*P$ is a positive operator which has a unique positive square root, namely P; likewise for P_0^2. By uniqueness we get $P_0 = P$, from which $U_0 = U$ folllows. $\square$

Computing U for Invertible $T : V \to V$. Determining the positive part $P = |T|$ is straightforward: $P^2 = T^*T$ is self-adjoint and its spectral decomposition can be computed in the usual way. If $\{e_i\}$ is an orthonormal basis of eigenvectors for T^*T, which are also eigenvectors for $P = |T|$, we have

$$T^*T(e_i) = \lambda_i e_i \qquad \text{and} \qquad |T|(e_i) = \sqrt{\lambda_i}\, e_i \tag{66}$$

(with all $\lambda_i > 0$ because $|T|$ is invertible $\Leftrightarrow$ T is invertible $\Leftrightarrow$ all $\lambda_i \neq 0$). From this we get

$$\begin{aligned} |T|^{-1}(e_i) &= |T|^{-1}\left(\frac{1}{\sqrt{\lambda_i}}\,|T|(e_i)\right) = \frac{1}{\sqrt{\lambda_i}}\,e_i \\ &\Downarrow \\ U(e_i) &= T\Big(|T|^{-1}e_i\Big) = \frac{1}{\sqrt{\lambda_i}}\,T(e_i) \end{aligned}$$

By its construction U is unitary on V so the vectors

$$f_i = \frac{1}{\sqrt{\lambda_i}}\,T(e_i)$$

are a new orthonormal basis in V. This completely determines U. $\square$

Note that

$$\|U(e_i)\| = \frac{1}{\sqrt{\lambda_i}}\,\|T(e_i)\| = \frac{1}{\sqrt{\lambda_i}}\,\big\|\,|T|(e_i)\,\big\| = \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_i}}\,\|e_i\| = 1$$

as expected.

The General Polar Decomposition. When $T : V \to V$ is not invertible the polar decomposition is somewhat more complicated. The positive component in $T = U{\cdot}P$ is still the unique positive square root $P = |T| = (T^*T)^{1/2}$. But the unitary part is based on a uniquely determined isometry $U_0 : R(|T|) \to R(T)$ between proper subspaces in V that can have various extensions to a unitary map $U : V \to V$. This ambiguity has no effect on the factorization $T = U \cdot P$; the behavior of U off of $R(|T|)$ is completely irrelevant.

7.5 Theorem (Polar Decomposition II). *Any linear operator $T : V \to V$ on a finite dimensional complex inner product space has a factorization $T = U\,|T|$ where*

1. *$|T|$ is the positive square root of T*
2. *U is a unitary operator on V.*

The unitary factor is uniquely determined only on the range $R(|T|)$, which is all that matters in the decomposition $R = U\,|T|$, but it can be extended in various ways to a unitary map $V \to V$ when T is not invertible.

Proof: First note that

$$R(|T|) = K(|T|)^{\perp} = K(T)^{\perp} = K(T^*T)^{\perp} = R(T^*T)$$
$$R(|T|)^{\perp} = K(|T|) = K(T) = K(T^*T) = R(T^*T)^{\perp}$$

The subspaces in the first row are just the orthocomplements of those in the second. The first and last identities in Row 2 hold because $|T|$ and T^*T are self-adjoint (Proposition VI-4.2); the rest have been proved in (65). We now observe that equation (64)

$$\|Tv\|^2 = (T^*Tv, v) = \big(|T|^2 v, v\big) = \big\|\,|T|(v)\,\big\|^2 \qquad \text{for all } v \in V\ ,$$

implies that there is a norm-preserving bijection U_0 from $R(|T|) \to R(T)$, defined by letting

$$U_0\big(|T|(v)\big) = T(v)\ . \tag{67}$$

This makes sense despite its seeming ambiguity: If an element $y \in R(|T|)$ has realizations $y = |T|(v') = |T|(v)$ we get $|T|(v' - v) = 0$, and then

$$T(v' - v) = T(v') - T(v) = 0$$

because $|T|(v'-v)$ and $T(v'-v)$ have equal norms. Thus $T(v') = T(v)$ and the operator (67) is in fact a well-defined bijective map from $R(|T|)$ into $R(T)$. It is linear because

$$\begin{aligned} U_0\big(|T|v_1 + |T|v_2\big) &= U_0\big(|T|(v_1 + v_2)\big) \;=\; T(v_1 + v_2) \\ &= Tv_1 + Tv_2 \;=\; U_0(|T|v_1) + U_0(|T|v_2) \end{aligned}$$

It is then immediate that $\|U_0(y)\| = \|y\|$ for all $y \in R(|T|)$, and $R(U_0) \subseteq R(T)$, so $\dim R(U_0) \leq \dim R(T)$. But $\dim R(U_0) = \dim R(|T|)$ by definition of U_0, and $K(T) = K(|T|) \Rightarrow \dim R(T) = \dim R(|T|)$. Putting these facts together we get

$$\dim R(|T|) = \dim R(U_0) \leq \dim R(T) = \dim R(|T|)$$

We conclude that $R(U_0) = R(T)$ and $U_0 : R(|T|) \to R(T)$ is a bijective isometry between subspaces of equal dimension. By definition we get

$$T(v) = \big(U_0 \cdot |T|\big)(v) \qquad \text{for all } v \in V\ .$$

We can extend U to a globally defined unitary map $U : V \to V$ because $K(T) = K(|T|) \Rightarrow \dim\, R(T) = \dim\, R(|T|)$ and $\dim\, R(T)^\perp = \dim\, R(|T|)^\perp$; therefore there exist various isometries

$$U_1 : R(|T|)^\perp \to R(T)^\perp \ .$$

corresponding to orthonormal bases in these subspaces. Using the orthogonal decompositions

$$V = R(|T|) \dot{\oplus} R(|T|)^\perp = R(T) \dot{\oplus} R(T)^\perp$$

we obtain a bijective map

$$U(v, v') = \big(U_0(v), U_1(v')\big)$$

such that $U|T| = U_0|T| = T$ on all of V. $\square$

There is a similar decomposition for operators $T : V \to W$ between different inner products spaces; we merely sketch the proof. Once again we define the positive component $|T| = (T^*T)^{1/2}$ as in (63). The identity

$$\big\|\, |T|(v)\, \big\|_V^2 = \|\, T(v)\, \|_W^2 \qquad \text{for all } v \in V$$

holds exactly as in (64), and this induces a linear isometry U_0 from $M = R(|T|) \subseteq V$ to $N = R(T) \subseteq W$ such that

$$T = U_0 \cdot |T| = \Big[T \cdot \big(|T|_M\big)^{-1}\Big] \cdot |T|$$

where $|T|_M =$ (restriction of $|T|$ to M).

The fact that U_0 is only defined on $R(|T|)$ is irrelevant, as it was in Theorem 7.5, but now U_0 cannot be extended unitary map (bijective isometry) from V to W unless $\dim(V) = \dim(W)$. On the other hand since $|T|$ is self-adjoint we have

$$R(|T|) = K(|T|)^\perp = K(T)^\perp$$

and can define $U \equiv 0$ on $K(T)$ to get a globally defined "*partial isometry*" $U : V \to W$

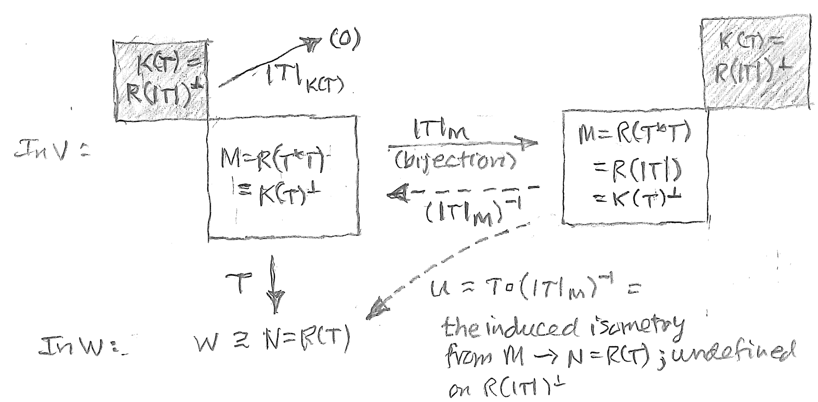

Figure 6.11. The maps involved in defining a polar decomposition $T = U_0 \cdot |T|$ for an arbitrary linear map $T : V \to W$ between different inner product spaces. Here we abbreviate $M = K(T)^\perp \subseteq V$ and $N = R(T) \subseteq W$; $U_0 : M \to N$ is an induced isometry such that $T = U_0 \cdot |T|$.

such that $K(U) = K(T)$, $R(U) = R(U_0) = R(T)$, and

$$U|_{K(T)} = 0 \qquad\qquad U|_{K(T)^\perp} = U|_{R(|T|)} = U_0$$

The players involved are shown in the commutative diagram Figure 6.11.

The **singular value decomposition** is a useful variant of Theorem 7.5.

7.6 Theorem (Singular Value Decomposition). *Let $T : V \to W$ be a linear operator between complex inner product spaces. There exist nonnegative scalars*

$$\lambda_1 \geq \ldots \geq \lambda_r \geq 0 \qquad (r = \mathrm{rank}(T))$$

and orthonormal bases $\{e_1, \ldots, e_r\}$ for $K(T)^\perp \subseteq V$ and $\{f_1, \ldots, f_r\}$ for $R(T) \subseteq W$ such that

$$T(e_i) = \lambda_i f_i \text{ for } 1 \leq i \leq r \qquad \text{and} \qquad T \equiv 0 \text{ on } K(T) = K(T)^{\perp\perp}$$

The λ_i are the eigenvalues of $|T| = (T^*T)^{1/2}$ counted according to their multiplicities.

Chapter V. The Diagonalization Problem.

V.1 The Characteristic Polynomial.

The **characteristic polynomial** $p_A(x)$ of an $n \times n$ matrix is defined to be

$$p_A(x) = \det(A - xI) \qquad (x \text{ an indeterminate})$$

This is a polynomial in $\mathbb{K}[x]$. In fact $\det(A - xI)$ is a polynomial combination of the entries in $(A - xI)$, so it follows that $p_A(x)$ does determine a polynomial in the single unknown x; furthermore $\deg(p_A) = n$. Given a linear operator $T : V \rightarrow V$ on a finite dimensional space V and a basis $\mathfrak{X}$ we have

$$[T - xI]_{\mathfrak{X}\mathfrak{X}} = [T]_{\mathfrak{X}\mathfrak{X}} - x\, I_{n\times n} \qquad (n = \dim(V))$$

so we may define a characteristic polynomial for T in the obvious way.

$$p_T(x) = \det(T - xI) = \det\big([T]_{\mathfrak{X}\mathfrak{X}} - xI_{n\times n}\big) \qquad (x \text{ an indeterminate})$$

The discussions for operators and matrices are so similar that nothing is lost if we focus on matrices for the time being.

Next observe what happens if we write out the characteristic polynomial p_A,

$$p_A(x) = \det(A - xI) = c_0(A) + c_1(A)x + \ldots + c_n(A)x^n \tag{37}$$

In this formula the coefficients $c_i(A)$ are scalar-valued functions from $\mathrm{M}(n, \mathbb{K}) \rightarrow \mathbb{K}$.

1.1. Lemma. *Each coefficient $c_k(A)$ in (37) is a similarity invariant on matrix space*

$$c_k\big(SAS^{-1}\big) = c_k(A) \qquad \textit{for all } A \in \mathrm{M}(n, \mathbb{K}), S \in \mathrm{GL}(n, \mathbb{K})$$

Furthermore, if we identify $\mathrm{M}(n, \mathbb{K})$ with n^2-dimensional coordinate space $\mathbb{K}^{n^2}$ via the correspondence $A \mapsto (a_{11}, \ldots, a_{1n}; \ldots; a_{n_1}, \ldots, a_{nn})$, each coefficient $c_i(A)$ is a polynomial function of the matrix entries: there is a polynomial $F_i \in \mathbb{K}[\mathbf{x}] = \mathbb{K}[x_1, \ldots, x_{n^2}]$ such that $c_i(A) = F_i\big(a_{11}, a_{12}, \ldots, a_{nn}\big)$.

Proof: We have

$$\begin{aligned} \det\big(S(A - xI)S^{-1}\big) &= \det(SAS^{-1} - x\,SS^{-1}) = \det(SAS^{-1} - x\,I) \\ &= c_0\big(SAS^{-1}\big) + c_1\big(SAS^{-1}\big)x + \ldots + c_n\big(SAS^{-1}\big)x^n \ , \end{aligned}$$

while at the same time

$$\begin{aligned} \det\big(S(A - xI)S^{-1}\big) &= \det(S)\cdot\det(A - xI)\cdot\det(S^{-1}) \\ &= \det(A - xI) = c_0(A) + c_1(A)x + \ldots + c_n(A)x^n \end{aligned}$$

for all $x \in \mathbb{K}$. Since these are the same polynomial in $\mathbb{K}[x]$ the coefficients must agree, hence $c_i\big(SAS^{-1}\big) = c_i(A)$.

The polynomial nature of the coefficients as functions of $A \in \mathbb{K}^{n^2}$ follows because $\det(A - xI)$ is a polynomial combination of entries $(A - xI)_{ij}$; the coefficients $c_k(A)$ are then polynomial functions of the a_{ij} when like powers of the unknown "x" are gathered together. $\square$

It is interesting to examine how the coefficients $c_k(A)$ are obtained from entries in A. Starting from the original definition of the determinant in Chapter IV,

$$\det(B) = \sum_{\sigma \in S_n} sgn(\sigma) \cdot \Big(\prod_{i=1}^{n} b_{i,\sigma(i)} \Big) \ ,$$

if we take $B = A - xI$ we have

$$B \ = \ A - xI \ = \ \begin{pmatrix} (a_{11} - x) & a_{12} & . & a_{1n} \\ a_{12} & (a_{22} - x) & . & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & . & . & (a_{nn} - x) \end{pmatrix}$$

It is clear that the only template yielding a product $b_{1,\sigma(1)} ... b_{n\sigma(n)}$ involving x^n is the diagonal template corresponding to the trivial permutation $\sigma = e$; furthermore, in expanding the product $\prod_i (a_{ii} - x)$ we must take the "$-x$" instead of "a_{ii}" from each factor to get the power x^n. Thus $c_n(A) \equiv (-1)^n$ is constant on matrix space (and certainly a similarity invariant).

We claim that

$$\begin{aligned} \det(A - xI) \quad &= \quad (-1)^n x^n + (\text{terms of lower degree}) \\ &= \quad (-1)^n x^n + (-1)^{n-1} \mathrm{Tr}(A)\, x^{n-1} + \ldots + \det(A) \cdot 1 \end{aligned} \tag{38}$$

To get the coefficient of x^{n-1} observe that a product $\prod_i b_{i,\sigma(i)}$ involving x^{n-1} must come from a template having $(n-1)$ marked spots on the diagonal, but then *all* marked spots must lie on the diagonal and we are again dealing with the diagonal template (for $\sigma = e$). In expanding the product $\prod_i (a_{ii} - x)$ we must now select the "$-x$" from $n-1$ factors and the "a_{ii}" from just one. Thus

$$c_{n-1}(A) = (-1)^{n-1} \cdot \sum_{i=1}^{n} a_{ii} = (-1)^{n-1} \mathrm{Tr}(A)$$

as in (38). Determining the other coefficients is tricky business, except for the constant term which is

$$a_0(A) = \det(A)$$

This follows because *every* template yields a product that contributes to this constant term. However if a template marks a spot on the diagonal we must select the "a_{ii}" term rather than the "x" from that diagonal entry $(a_{ii} - x)$. It follows that the constant term in (38) is:

$$\sum_{\sigma \in S_n} sgn(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma(i)} = \det(A)$$

as claimed. We leave discussion of other terms in the expansion (38) for more advanced courses.

Factoring Polynomials. It is well known that if a nonconstant polynomial $f(x)$ in $\mathbb{K}[x]$ has a root $\alpha \in \mathbb{K}$, so $f(\alpha) = \sum_{i=0}^{n} c_i \alpha^i = 0$, then we can factor $f(x) = (x - \alpha) \cdot g(x)$ by long division of polynomials, with $\deg(g) = \deg(f) - 1$. In fact, applying the Euclidean algorithm for division with remainder in $\mathbb{K}[x]$: if $P, Q \in \mathbb{K}[x]$ and $\deg(Q) \geq 1$ we can always write

$$P(x) = A(x)Q(x) + R(x) \qquad (\text{with remainder } R \equiv 0 \text{ or } \deg(R) < \deg(Q))$$

Taking P to be any nonconstant polynomial in $\mathbb{K}[x]$ and $Q = (x - \alpha)$, we get $f(x) = A(x) \cdot (x - \alpha) + R(x)$ where $R(x)$ is either the zero polynomial, or $R(x)$ is nonzero with $\deg(R) < \deg(x - \alpha) = 1$ – i.e. $R(x)$ is then a nonzero *constant* polynomial $R = c\mathbb{1}$. If $\alpha \in \mathbb{K}$ is a root of f, replacing x by α everywhere yields the identity

$$0 = f(\alpha) = A(\alpha)\cdot(\alpha - \alpha) + R(\alpha) = 0 + R(\alpha) = R(\alpha)$$

Since $R = c\mathbb{1}$, this forces $R(x) \equiv 0$ and $f(x) = A(x)(x - \alpha)$ with no remainder – i.e. $(x - \alpha)$ divides $f(x)$ exactly.

If α_1 is a root of f we may split $f(x) = (x - \alpha_1) \cdot g_1(x)$. If we can find a root α_2 of $g_1(x)$ in $\mathbb{K}$ we can continue this process, obtaining $f(x) = (x - \alpha_1)(x - \alpha_2) \cdot g_2(x)$. Pushing this as far as possible we arrive at a factorization

$$f(x) = \prod_{i=1}^{s} (x - \alpha_i) \cdot g(x)$$

in which $g(x)$ has *no* roots in $\mathbb{K}$. We say that f **splits completely over** $\mathbb{K}$ if g reduces to a constant polynomial, so that $f(x) = c\prod_{i=1}^{n}(x - \alpha_i)$. There may be repeated factors, and if we gather together all factors of the same type this becomes

$$f(x) = c\prod_{j=1}^{r}(x - \alpha_i)^{m_j} \qquad (\alpha_i \in \mathbb{K})$$

The roots $\{\alpha_1, \ldots, \alpha_r\}$ are now *distinct* and the exponents $m_i \geq 1$ are their **multiplicities** as roots of $f(x)$; the constant c out front is the coefficient of the leading term $c_n x^n$ in $f(x)$.

1.2. Corollary. *A nonconstant polynomial $f(x) \in \mathbb{K}[x]$ can have at most $n = \deg(f)$ roots in any field of coefficients $\mathbb{K}$. More generally the sum of the multiplicities of the roots in $\mathbb{K}$ is at most n.*

Proof: If $f, g \neq 0$ in $\mathbb{K}[x]$ (so they have well defined degrees) we know that

$$\deg(f(x) + g(x)) = \deg(f(x)) + \deg(g(x))$$

But, $\deg\big(\prod_{i=1}^{r}(x - \alpha_i)^{m_i}\big) = \sum_{i=1}^{r} m_i$, so

$$r = \#(\text{distinct roots}) \leq (m_1 + \ldots + m_r) + \deg(g) = \deg(f) \quad \square$$

1.3. Exercise. If $f(x), h(x)$ are nonzero polynomials over any field, explain why the "degree formula"

$$\deg(f(x)h(x)) = \deg(f(x)) + \deg(h(x))$$

is valid.

1.4. Exercise. Verify that if $f(x) = \prod_{i=1}^{r}(x - \alpha_i) \cdot g(x)$ and $g(x)$ has no roots in $\mathbb{K}$, then the roots of f in $\mathbb{K}$ are $\{\alpha_1, ..., \alpha_r\}$.
Note: Repetitions are allowed; $f(x)$ might even have the form $(x - \alpha)^r \cdot g(x)$.)

1.5. Definition. *The distinct roots $\{\alpha_1, \ldots, \alpha_r\}$ in $\mathbb{K}$ of a nonconstant polynomial and their multiplicities are uniquely determined, and the set of roots is called the* **spectrum** *of the polynomial f and is denoted by $\mathrm{sp}_{\mathbb{K}}(f)$.*

Over the field $K = \mathbb{C}$ of complex numbers we have:

1.6. Theorem (Fundamental Theorem of Algebra). *If f is a nonconstant polynomial in $\mathbb{C}[x]$ then f has a root $\alpha \in \mathbb{C}$, so that $f(\alpha) = 0$.*

1.7. Corollary. Every non constant $f \in \mathbb{C}[x]$ splits completely over $\mathbb{C}$, with

$$f(x) = c \cdot \prod_{i=1}^{r} (x - \alpha_i)^{m_i} \qquad \text{where } m_1 + ... + m_r = n$$

Proof: Since f has a root we may factor $f = (x - \alpha_1) \cdot g_1(x)$. Unless $g_1(x)$ is a constant it also has a root, allowing us to write $f = (x - \alpha_1)(x - \alpha_2)g_2(x)$. Continue recursively. $\square$

Over $\mathbb{K} = \mathbb{R}$ or $\mathbb{Q}$, things get more complicated and $f(x)$ might not have any roots at all in $\mathbb{K}$. For example if $f(x) = x^2 + 1$ over $\mathbb{R}$, or $f(x) = x^2 - 2$ over $\mathbb{Q}$ since $\mathbb{Q}$ does not contain any element α such that $\alpha^2 = 2$ (there is no "square root of 2" in $\mathbb{Q}$). Nevertheless since $\mathbb{R} \subseteq \mathbb{C}$ we may regard any $f \in \mathbb{R}[x]$ as a complex polynomial that happens to have all real coefficients. All real roots α remain roots $\alpha + i0$ in $\mathbb{C}$ (lying on the real axis), but enough new roots appear in the larger field to split f completely as

$$f(x) = c \cdot \prod (x - \alpha_i) \qquad \text{with } \alpha_i \in \mathbb{C}$$

It is important to realize that the new non-real roots enter in "conjugate pairs."

1.8. Lemma. *If $f(x)$ is nonconstant in $\mathbb{R}[x]$ and $z = x + iy$ is a complex root when we identify $\mathbb{R} \subseteq \mathbb{C}$ and $\mathbb{R}[x] \subseteq \mathbb{C}[x]$, then the complex conjugate $\overline{z} = x - iy$ is also a root.*

Proof: There is nothing to prove if z is real ($y = 0$). Otherwise, recall that conjugation

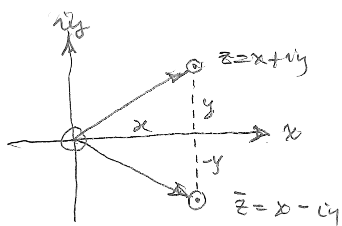

Figure 5.1. Non-real roots of a polynomial with real coefficients come in conjugate pairs $z = x + iy$ and $\overline{z} = x - iy$, mirror images of each other under reflection across the x-axis.

has the following algebraic properties.

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \qquad \text{and} \qquad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}$$

Then

$$\overline{z^n} = \overline{z}^n \qquad \text{for all } n \in \mathbb{Z} \text{ and } z \in \mathbb{C}$$

Hence if $0 = f(z) = \sum_{j=0} c_j z^j$ with c_j real we have

$$\overline{(c_j z^j)} = \overline{c_j}\,\overline{(z^j)} = c_j \overline{z}^j$$

and

$$0 = \overline{0} = \overline{f(z)} = \sum_{j=0} \overline{(c_j z^j)} = \sum_{j=0} c_j \big(\overline{z}\big)^j = f(\overline{z})$$

Hence, $\overline{z}$ is also a root in $\mathbb{C}$. $\square$

The real roots of $f \in \mathbb{R}[x]$ are not subject to any constraints; in fact, all the roots might be real. The number of distinct non-real roots is always *even.*

1.9. Example. If $f \in \mathbb{K}[x]$ is quadratic,

$$f(x) = ax^2 + bx + c \qquad \text{with } a \neq 0,$$

the quadratic formula continues to apply for all fields except those of "*characteristic* 2," in which $2 = 1 + 1$ is equal to 0 (for insatnce $\mathbb{K} = \mathbb{Z}_2$). Except for this, the roots are given by:

$$\text{Quadratic Formula:} \qquad z_{\pm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

If the $\sqrt{\cdots}$ fails to exist in $\mathbb{K}$ the proper conclusion is that $f(x)$ has no roots in $\mathbb{K}$. If $\mathbb{K} = \mathbb{Q}$ or $\mathbb{R}$ this formula gives the correct roots in $\mathbb{C}$ even if there are no roots in $\mathbb{K}$.

Discussion: Complete the square. Adding/subtracting a suitable constant d we may write

$$\begin{aligned} ax^2 + bx + c &= a\Big(x^2 + \frac{b}{a}x + \frac{c}{a}\Big) = a\left[\Big(x^2 + \frac{b}{a}x + d\Big) + \Big(\frac{c}{a} - d\Big)\right] \\ &= a\cdot\Big(x^2 + \frac{b}{a}x + d\Big) + (c - ad) \end{aligned}$$

To make $x^2 + (b/a)x + d$ a "perfect square" of the form $(x + k)^2 = (x^2 + 2kx + k^2)$, we must take $k = b/(2a)$ and $d = k^2 = (b^2/4a^2)$. Then $c - ad = c - (b^2/4a^2)$, so that

$$0 = ax^2 + bx + c = a\Big(x + \frac{b}{2a}\Big)^2 + \Big(\frac{c}{a} - \frac{b^2}{4a^2}\Big) = a\cdot\Big(x + \frac{b}{2a}\Big)^2 + \Big(\frac{4ac - b^2}{4a}\Big)$$

This happens if and only if

$$a\Big(x + \frac{b}{2a}\Big)^2 = \Big(\frac{b^2 - 4ac}{4a}\Big)$$

if and only if

$$\Big(x + \frac{b}{2a}\Big)^2 = \frac{b^2 - 4ac}{4a^2}$$

if and only if

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \qquad \square$$

1.10. Example. Here are some examples of factorization of polynomials.

1. $x^2 - 1 = (x - 1)(x + 1)$ splits over $\mathbb{R}$, with two roots $+1, -1$ each of multiplicity one. On the other hand $x^2 + 1$ has no roots and does not split over $\mathbb{R}$, but it does split over $\mathbb{C}$, with $x^2 + 1 = (x - i)(x + i)$.
2. $x^2 + 2x + 1$ splits over $\mathbb{R}$ as $(x - 1)^2$, but there is just one root, of multiplicity 2;
3. $x^3 - x^2 + x - 1$ has a root $x = 1$ in $\mathbb{R}$. Long division yields a quadratic,

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

   Over $\mathbb{R}$, there is just one root $\lambda_1 = 1$ with multiplicity $m(\lambda_1) = 1$; over $\mathbb{C}$ we get $x^2 + 1 = (x + i)(x - i)$ so there are two more roots roots $\lambda_2 = i$, $\lambda_3 = -i$ in the larger field $\mathbb{C}$.
4. $x^4 - 1 = (x^2 - 1)(x^2 + 1) = (x + 1)(x - 1)(x + i)(x - i)$.

5. $x^3 + x + 1$ has just one real root λ_1 because it is a strictly increasing function of $x \in \mathbb{R}$, and since it goes to $\pm\infty$ as $x \to \pm\infty$ it must cross the x-axis somewhere. But λ_1 is not so easy to write as an explicit algebraic expression involving sums, products, quotients, and cube roots. Such formulas exist, but are algorithms with possible branch points rather than simple expressions like the quadratic formula. A numerical estimate yields the real root $\lambda_1 = -0.6823 + i0$. There is a conjugate pair of complex roots $\lambda_2 = 0.3412 + 1.615\,i$ and $\lambda_3 = 0.3412 - 1.615\,i$, which could be found by (numerically) long dividing $f(x)$ by $(x - \lambda_1)$ and applying the quadratic formula to find the complex roots of the resulting quadratic.

V.2. Finding Eigenvalues.

If V is a finite dimensional vector space we say $\lambda \in \mathbb{K}$ is an **eigenvalue** for a linear operator $T : V \to V$ if there is $v \neq 0$ such that $T(v) = \lambda v$. For any $\lambda \in \mathbb{K}$ the **λ-eigenspace** is $E_\lambda = \{v \in V : (T - \lambda I)v = 0\}$. This vector subspace is nontrivial if and only if λ is an eigenvalue. The set of distinct eigenvalues is called the **spectrum** $\mathrm{sp}_{\mathbb{K}}(T)$ of the operator. When $\lambda = 0$ the eigenspace $E_{\lambda=0}(T)$ is just $\ker(T) = \{v \in V : T(v) = 0\}$ and when $\lambda = 1$ we get the subspace of *fixed vectors* $E_{\lambda=1}(T) = \{v : T(v) = v\}$.

The connection with determinants now emerges: $\lambda \in \mathbb{K}$ is an eigenvalue if and only if

$$\ker(T - \lambda I) \neq (0) \;\Leftrightarrow\; (T - \lambda I) \text{ is singular } \;\Leftrightarrow\; \det(T - \lambda I) = 0$$

Thus the eigenvalues are the roots in $\mathbb{K}$ of the characteristic polynomial $p_T \in \mathbb{K}[x]$.

2.1. Definition. *If $T : V \to V$ is a linear operator on a finite dimensional vector space then $\mathrm{sp}_{\mathbb{K}}(T)$ is the set of distinct roots in $\mathbb{K}$ of the characteristic polynomial $p_T(x) = \det(T - xI)$. We define the* **geometric multiplicity** *of an eigenvalue to be $\dim(E_\lambda)$; its* **algebraic multiplicity** *is the multiplicity of λ as a root of the characteristic polynomial, so that $p_T(x) = (x - \lambda)^m \cdot g(x)$ and $g(x)$ does not have λ as a root.*

2.2. Lemma. *Over any field $\mathbb{K}$,*

$$(\textit{algebraic multiplicity of } \lambda) \geq (\textit{geometric multiplicity})$$

Over $\mathbb{K} = \mathbb{C}$, the sum of the algebraic multiplicities of the (distinct) eigenvalues in $\mathrm{sp}_{\mathbb{C}}(T) = \{\lambda_1, ..., \lambda_r\}$ is $m(\lambda_1) + \ldots + m(\lambda_r) = n = \dim_{\mathbb{C}}(V)$.

Proof: Every eigenspace E_λ is T-invariant because $(T - \lambda I)T(v) = T(T - \lambda I)v = 0$ for $v \in E_\lambda$. This eigenspace has a basis of eigenvectors $\mathfrak{X}_\lambda = \{e_1, \ldots, e_d\}$, with respect to which

$$[T]_{\mathfrak{X}_\lambda} = \begin{pmatrix} \lambda & & & 0 \\ & \lambda & & \\ & & \ddots & \\ 0 & & & \lambda \end{pmatrix}$$

(diagonal). Extending $\mathfrak{X}_\lambda$ to a basis $\mathfrak{X} = \{e_1, ..., e_d, e_{d+1}, ..., e_n\}$ for all of V, we get

$$[T]_{\mathfrak{X}\mathfrak{X}} = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & B \end{array}\right)$$

which implies that

$$[T - x\,I]_{\mathfrak{X}\mathfrak{X}} = \left(\begin{array}{ccc|c} \lambda - x & & & \\ & \ddots & & * \\ 0 & & \lambda - x & \\ \hline & 0 & & B - \lambda I \end{array}\right)$$

2.3. Lemma. *If A is of the form*

$$A = \left(\begin{array}{c|c} B & D \\ \hline 0 & C \end{array}\right)$$

where B and C are two square matrices, then $\det(A) = \det(B) \cdot \det(C)$.

Proof: If B is $m \times m$, a sequence of Type II and III row operations on rows $R_1, \dots, R_m$ puts this block in upper triangular form; similar operations on rows $R_{m+1}, \dots, R_n$ puts block C in upper triangular form without affecting any of the earlier rows. The net result is an echelon form $A' = [B', *; 0, C']$ for which $\det A' = \det(B') \cdot \det(C')$. Each of the determinants $\det(A'), \dots, \det(C')$ differs from its counterpart by a $\pm$ sign; furthermore, the same moves that put B and C in upper triangular form also put A in upper triangular form when applied to the whole $n \times n$ matrix. We leave the reader to check that the sign changes cancel and yield $\det(A) = \det(B) \cdot \det(C)$. $\square$

This can also be seen by noting that if a template contributes to $\det(A)$, every column passing through block B must be marked at a spot in B; otherwise it would marked at a spot below B, whose entry is $= 0$. Likewise for the rows that meet block C, so a template contributes $\Leftrightarrow$ it has the form in Figure 5.2.

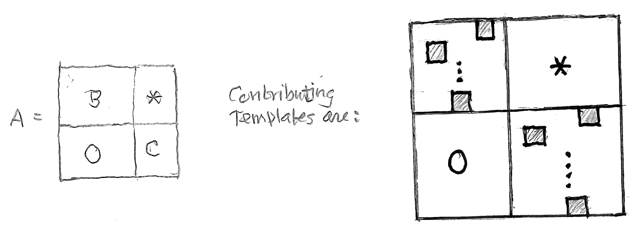

Figure 5.2. If A is a block upper-triangular square matrix, then $\det(A) = \det(B){\cdot}\det(C)$ and the only templates that contribute to $\det(A)$ are those whose marked spots lie entirely within the blocks B and C.

Applying Lemma 2.3 we can complete the proof of Lemma 2.2. We now see that

$$p_T(x) = \det(T - x\,I) = (\lambda - x)^m \cdot Q(x) \qquad \text{where } Q(x) = \det(B - x\,I)$$

Obviously $\deg(Q(x)) = n - m$ and $p_T(x)$ has λ as a root of multiplicity at least m, so (*algebraic multiplicity of* λ) $\geq m = \dim(E_\lambda)$ as claimed. $\square$

It might still be possible for $(x - \lambda)$ to divide $Q(x)$, making the algebraic multiplicity larger than $\dim\big(E_\lambda\big)$. A good example is $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$. The operator $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ has $\dim\big(E_{\lambda=1}\big) = 1$, but $p_T(\lambda) = (\lambda - x)^2$ so the algebraic multiplicity is 2.

The following example illustrates the complete diagonalization process.

2.4. Example. Let $T = L_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ with

$$A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$$

If $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis in $\mathbb{R}^3$ we have $[T]_{\mathfrak{X}\mathfrak{X}} = [L_A]_{\mathfrak{X}\mathfrak{X}} = A$ as in Exercise 4.13 of Chapter II, so

$$\begin{aligned} p_A(x) &= \det(A - x\,I) \;=\; \det \begin{pmatrix} 4-x & 0 & 1 \\ 2 & 3-x & 2 \\ 1 & 0 & 4-x \end{pmatrix} \\ &= \big[(4-x)(3-x)(4-x) + 0 + 0\big] - \big[(3-x) + 0 + 0\big] \\ &= -x^3 + 11x^2 - 39x + 45 \end{aligned}$$

We are looking for roots of a cubic equation. If you can guess a root α, then long divide by $x - \alpha$ to get $p_T(x) = (x-\alpha)\cdot(\mathit{quadratic})$; otherwise you will have to use a numerical root-finding program. Trial and error reveals that $x = 3$ is a root and long division by $(x-3)$ yields

$$\begin{array}{r|rrrr} & -x^2 & +8x & -15 & \\ \hline x-3) & -x^3 & 11x^2 & -39x & +45 \\ & -x^3 & +3x^2 & & \\ \hline & & 8x^2 & -39x & +45 \\ & & 8x^2 & -24x & \\ \hline & & & -15x & +45 \\ & & & -15x & +45 \\ \hline & & & & 0 \end{array}$$

Then

$$\begin{aligned} -x^3 + 11x^2 - 39x + 45 &= (x-3)(-x^2 + 8x - 15) \\ &= -(x-3)(x-5)(x-3) \;=\; -(x-3)^2(x-5) \quad , \end{aligned}$$

so $\mathrm{sp}(A) = \{3, 5\}$ with algebraic multiplicities $m_{\lambda=3} = 2$, $m_{\lambda=5} = 1$. To determine the eigenspaces and geometric multiplicities we must solve systems of equations.

Eigenvalue $\lambda_1 = 3$: We must solve the matrix equation $(A - 3I)X = 0$. Row operations on $[A - 3I\,|\,0]$ yield

$$[A - 3I\,|\,0] = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Solutions: x_2, x_3 are free variables and $x_1 = -x_3$, so

$$X = \begin{pmatrix} -x_3 \\ x_2 \\ x_3 \end{pmatrix} \qquad \text{and} \qquad E_{\lambda=3} = \ker(A - 3I) = \mathbb{R}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \mathbb{R}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

Thus $\lambda = 3$ has geometric multiplicity $\dim\big(E_{\lambda=3}\big) = 2$.

Eigenvalue $\lambda_2 = 5$: Solve matrix equation $(A - 5I)X = 0$. Row operations on $[A - 5I\,|\,0]$ yield

$$[A - 5I\,|\,0] = \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 2 & -2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Solutions: x_3 is a free variable; $x_2 = 2x_3$, $x_1 = x_3$. So

$$X = \begin{pmatrix} x_3 \\ 2x_3 \\ x_3 \end{pmatrix} \quad \text{and} \quad E_{\lambda=5} = \ker(A - 5I) = \mathbb{R}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Thus $\lambda = 5$ has geometric multiplicity $\dim\big(E_{\lambda=5}\big) = 1$.

We showed earlier that the span $M = \sum_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)$ of the eigenspaces of a linear operator is actually a *direct* sum $M = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r}$. In the present situation $M = E_{\lambda=3} \oplus E_{\lambda=5}$ is all of V since the dimension add up to $\dim(V) = 3$. Taking a basis $\mathfrak{Y} = \{\mathbf{f}_1, .., \mathbf{f}_3\}$ that runs first through $E_{\lambda=3} = \mathbb{R}\mathbf{f}_1 \oplus \mathbb{R}\mathbf{f}_2$, and then through $E_{\lambda=5} = \mathbb{R}\mathbf{f}_3$, we obtain a diagonal matrix

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = \left(\begin{array}{cc|c} 3 & 0 & 0 \\ 0 & 3 & 0 \\ \hline 0 & 0 & 5 \end{array}\right)$$

Once we have found the diagonalizing basis

$$\mathfrak{Y} = \{\mathbf{f}_1 = (0,1,0)\,,\, \mathbf{f}_2 = (-1,0,1)\,,\, \mathbf{f}_3 = (1,2,1)\}$$

we determine an invertible matrix Q such that $QAQ^{-1} = [T]_{\mathfrak{Y}\mathfrak{Y}} = \mathrm{diag}(3,3,5)$. To find Q recall that

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} \cdot [T]_{\mathfrak{X}\mathfrak{X}} \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} \cdot A \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = QAQ^{-1}$$

Here $[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = Q^{-1}$ and $[\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}^{-1}$, and by definition $[\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}$ is the transpose of the coefficient array in the system of vector identities

$$\begin{aligned} \mathbf{f}_1 &= [\mathrm{id}]\ \mathbf{f}_1 = 0 + \mathbf{e}_2 + 0 \\ \mathbf{f}_2 &= [\mathrm{id}]\ \mathbf{f}_2 = -\mathbf{e}_1 + 0 + \mathbf{e}_3 \\ \mathbf{f}_3 &= [\mathrm{id}]\ \mathbf{f}_3 = \mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3 \end{aligned}$$

Thus,

$$Q^{-1} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

and $Q = (Q^{-1})^{-1}$ can be found efficiently via row operations.

$$\left(\begin{array}{ccc|ccc} 0 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}\right)$$

Thus

$$Q = \begin{pmatrix} -1 & 1 & -1 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

and

$$QAQ^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

as expected. That completes the "spectral analysis" of A. $\square$

The same sort of calculations determine the eigenspaces in $\mathbb{C}^2$ when $A \in \mathrm{M}(n, \mathbb{R})$ is regarded as a matrix in $\mathrm{M}(n, \mathbb{C})$.

2.5. Example. Diagonalize the operator $L_A : \mathbb{K}^2 \to \mathbb{K}^2$ where

$$A = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}$$

over $\mathbb{C}$ and over $\mathbb{R}$ (insofar as this is possible).

Discussion: The characteristic polynomial of A (or L_A) is

$$p_A(\lambda) = \det \begin{pmatrix} 2-\lambda & 4 \\ -1 & -2-\lambda \end{pmatrix} = -(2-\lambda)(2+\lambda) + 4 = -4 + \lambda^2 + 4 = \lambda^2$$

The only root (real or complex) is $\lambda = 0$ so $\mathrm{sp}_{\mathbb{R}}(A) = \mathrm{sp}_{\mathbb{C}}(A) = \{0\}$. Its algebraic multiplicity is 2, but the geometric multiplicity $\dim_{\mathbb{K}} \left(E_{\lambda=0}\right)$ is equal to 1. The outcome is the same over $\mathbb{C}$ and $\mathbb{R}$.

Eigenvalue $\lambda = 0$. Here $E_{\lambda=0} = \ker(A)$. Row operations on $[A \,|\, 0]$ yield

$$\left(\begin{array}{cc|c} 2-\lambda & 4 & 0 \\ -1 & -2-\lambda & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 2 & 4 & 0 \\ -1 & -2 & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 2 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Solutions: In solving $(A - \lambda I)X = AX = 0$, x_2 is a free variable and $x_1 = -2x_2$ so

$$X = \begin{pmatrix} -2x_2 \\ x_2 \end{pmatrix} \quad \text{and} \quad E_{\lambda=0} = \mathbb{K} \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad .$$

Since there are no other eigenvalues, the best we can do in trying to find a simple matrix description $[T]_{\mathfrak{Y}\mathfrak{Y}}$ is to take a basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ that passes first through $E_{\lambda=0}$: let $\mathbf{f}_1 = (-2, 1)$ and then include one more vector $\mathbf{f}_2 \notin \mathbb{K}\mathbf{f}_1$ to make a basis. We have

$$[T]_{\mathfrak{X}\mathfrak{X}} = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}$$

with respect to the standard basis $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ in $\mathbb{K}^2$ (recall Exercise 4.13 of Chapter II). With respect to the basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ the matrix has block upper-triangular form,

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}$$

But this operator cannot be diagonalized by any choice of basis. $\square$

2.6. Exercise. We have shown that there is a basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ such that

$$A = [T]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}$$

(a) Prove that b must be 0, so
$$A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}$$

(b) Explain how to modify the basis $\mathfrak{Y}$ to get a new basis $\mathfrak{Z}$ such that
$$[T]_{\mathfrak{Z}\mathfrak{Z}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

2.7. Example. The matrix

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \qquad (\theta \text{ real})$$

yields an operator $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ that you will recognize as a rotation counter clockwise about the origin by θ radians. Describe its eigenspaces over $\mathbb{R}$ and over $\mathbb{C}$.

Solution: Over either field we have

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I) = \det \begin{pmatrix} \cos(\theta) - \lambda & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - \lambda \end{pmatrix} \\ &= \big(\cos(\theta) - \lambda\big)^2 + \sin^2(\theta) = \cos^2(\theta) + \sin^2(\theta) - 2\lambda\cos(\theta) + \lambda^2 \\ &= \lambda^2 - 2\lambda\cos(\theta) + 1 \end{aligned}$$

This is zero only when

$$\begin{aligned} \lambda &= \frac{2\cos(\theta) \pm \sqrt{4\cos^2(\theta) - 4}}{2} = \cos(\theta) \pm \sqrt{\cos^2(\theta) - 1} \\ &= \cos(\theta) \pm i\sqrt{1 - \cos^2(\theta)} = \cos(\theta) \pm i\sin(\theta) = e^{\pm i\theta} \end{aligned}$$

The roots are non-real (hence a conjugate pair as shown earlier in Figure 5.1), and they lie on the unit circle in $\mathbb{C}$ because $|e^{\pm i\theta}| = \sin^2(\theta) + \cos^2(\theta) = 1$ for all θ. When $\theta = 0$ or π we have $\lambda = \pm 1 + i0$ (real), and in this case $A = I$ or $-I$. In all other cases A has no real eigenvalues at all, but it can be diagonalized as

$$[L_A]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

for a suitably chosen complex basis $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ in $\mathbb{C}^2$. To find it we need to determine the eigenspaces of L_A in $\mathbb{C}^2$.

Eigenvalue: $\lambda_1 = e^{i\theta} = \cos(\theta) + i\sin(\theta)$.

$$\begin{aligned} [A - \lambda I] &= \begin{pmatrix} \cos(\theta) - e^{i\theta} & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - e^{i\theta} \end{pmatrix} = \begin{pmatrix} -i\sin(\theta) & -\sin(\theta) \\ \sin(\theta) & -i\sin(\theta) \end{pmatrix} \\ &= \sin(\theta) \cdot \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \end{aligned}$$

Now, $(A - \lambda I)X = 0 \Leftrightarrow BX = 0$ where $B = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}$. Row operations yield:

$$\left(\begin{array}{cc|c} -i & -1 & 0 \\ 1 & -i & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -i & 0 \\ 1 & -i & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & -i & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Solutions: Here x_2 is a free variable and $x_1 = ix_2$. So,

$$X = \begin{pmatrix} ix_2 \\ x_2 \end{pmatrix} \quad \text{and} \quad E_{\lambda_1 = e^{it}} = \mathbb{C} \cdot \begin{pmatrix} i \\ 1 \end{pmatrix} \quad .$$

For λ_1, (*algebraic multiplicity*) = (*geometric multiplicity*) = 1.

The discussion for the conjugate eigenvalue $\lambda_2 = e^{-i\theta} = \cos(\theta) - i\sin(\theta)$ is almost the same, with the final result that $E_{\lambda = e^{-i\theta}} = \mathbb{C} \cdot \mathrm{col}(-i, 1)$ Combining these observations we get

$$\mathbb{C}^2 = E_{\lambda = e^{i\theta}} \oplus E_{\lambda = e^{-i\theta}} = \mathbb{C} \cdot \begin{pmatrix} i \\ 1 \end{pmatrix} \oplus \mathbb{C} \cdot \begin{pmatrix} -i \\ 1 \end{pmatrix} = \mathbb{C}\mathbf{f}_1 \oplus \mathbb{C}\mathbf{f}_2$$

Thus, with respect to the basis

$$\mathfrak{Y} = \{\, \mathbf{f}_1 = \text{col}(i,1), \mathbf{f}_2 = \text{col}(-i,1)\,\}$$

we have

$$[L_A]_{\mathfrak{Y}\mathfrak{Y}} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \quad \square$$

As mentioned, the span $M = \sum_{\lambda \in \text{sp}_{\mathbb{K}}(A)} E_\lambda$ is a direct sum $E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r}$ and a suitable chosen basis partially diagonalizes A, with matrix

$$[T]_{\mathfrak{Y}\mathfrak{Y}} = \left(\begin{array}{ccc|c} \boxed{\lambda_1 \cdot I_{d_1 \times d_1}} & & 0 & \\ & \ddots & & * \\ & & \boxed{\lambda_r \cdot I_{d_r \times d_r}} & \\ \hline 0 & & 0 & B \end{array}\right)$$

To proceed further and determine the structure of the lower right-hand block B we would have to develop the theory of nilpotent operators, generalized eigenspaces, and the Jordan decomposition of a linear operator over $\mathbb{C}$. We must leave all that for a subsequent course. However the following observation can be useful.

2.8. Proposition. *If $\dim_{\mathbb{K}}(V) = n$ and $T : V \rightarrow V$ has n distinct eigenvalues in $\mathbb{K}$, then T is diagonalizable and V is the direct sum $\bigoplus_{i=1}^{n} E_{\lambda_i}$. of 1-dimensional eigenspaces.*

Proof: Since $\sum_{\lambda \in \text{sp}(T)} E_{\lambda_i}$ is a direct sum and each λ_i has $\dim\left(E_{\lambda_i}\right) \geq 1$, the dimension of this linear span must equal n, so $V = \bigoplus_{\lambda_i \in \text{sp}(T)} E_{\lambda_i}$. $\square$

In some sense (at least for complex matrices), the "n distinct eigenvalues condition" is *generic*: If entries $a_{ij} \in \mathbb{C}$ are chosen at random, then with "probability 1" the matrix $A = [a_{ij}]$ would have distinct eigenvalues in $\mathbb{C}$, so the characteristic polynomial would split completely into distinct linear factors

$$p_A(x) = c \cdot \prod_{i=1}^{n} (x - \lambda_i)\ .$$

Unfortunately, in many important applications the matrices of interest do not have n distinct eigenvalues, which is why we need the more subtle theory of "generalized eigenvalues" developed in Linear Algebra II, as a backup when diagonalization fails.

2.9. Exercise. What happens to $\text{sp}(T)$ when you replace

$$\text{(a)}\ \ T \rightarrow cT \qquad \text{(b)}\ \ T \rightarrow cT + I \qquad \text{(c)}\ \ T \rightarrow I + cT$$

with $c \neq 0$?

V.3 Diagonalization and Limits of Operators.

We begin by defining limits $\lim_{n\to\infty} A_n = A$ of square matrices over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$; limits $T_n \rightarrow T$ could similarly be defined for linear operators on a finite dimensional vector space V over these fields.

3.1. Definition. *For $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ we may define* **pointwise convergence***, or "***sup norm convergence***" of matrices in* $\text{M}(N, \mathbb{K})$

$$\lim_{n\to\infty} A_n = A \quad \textit{or} \quad A_n \rightarrow A \ \textit{as}\ n \rightarrow \infty$$

to mean that each entry in A_n converges in $\mathbb{C}$ to the corresponding entry in the limit matrix A:

$$\left|a_{ij}^{(n)} - a_{ij}\right| \to 0 \quad \text{in } \mathbb{C} \text{ as } n \to \infty \tag{39}$$

for each $1 \le i, j \le N$, where $A_n = [a_{ij}^{(n)}]$.

Later we will examine other notions of matrix (or operator) convergence. In making the present definition we are, in effect, measurng the "size" of an $N \times N$ matrix by its "**sup-norm**," the size of its largest entry:

$$\|A\|_\infty = \max\{\, |a_{ij}| : 1 \le i, j \le N\}$$

This allows us to define the distance between two matrices in $\mathrm{M}(N, \mathbb{K})$ to be $d(A, B) = \|A - B\|_\infty$, and it should be evident that the limit $A_n \to A$ defined in (39) can be recast in terms of the sup-norm:

$$A_n \to A \text{ as } n \to \infty \quad \Leftrightarrow \quad \|A_n - A\|_\infty \to 0 \text{ as } n \to \infty\ . \tag{40}$$

The sup norm on matrix space has several important properties (easily verified):

3.2. Exercise. If A, $B \in \mathrm{M}(N, \mathbb{K})$ prove that:

(a) $\|\lambda A\|_\infty = |\lambda| \cdot \|A\|_\infty$, for all $\lambda \in \mathbb{K}$,

(b) Triangle Inequality: $\|A + B\|_\infty \le \|A\|_\infty + \|B\|_\infty$;

(c) Multiplicative Property: $\|AB\|_\infty \le n \cdot \|A\|_\infty \cdot \|B\|_\infty$.

Hint: Use the Triangle Inequality in $\mathbb{C}$, which says $|z \pm w| \le |z| + |w|$ for $z, w \in \mathbb{C}$.

A number of theorems regarding sup-norm limits follow from these basic inequalities.

3.3. Exercise. If $A_n \to A$ and $B_n \to B$ in the sup-norm, and $\lambda_n \to \lambda$ in $\mathbb{C}$, prove that:

(a) $A_n + B_n \to A + B$

(b) $A_n B \to AB$ and $AB_n \to AB$;

(c) $A_n B_n \to AB$. Thus matrix multiplication is a "jointly continuous" operation on its two inputs.

(d) If Q is an invertible matrix then $QA_nQ^{-1} \to QAQ^{-1}$. Hence every similarity transformation $A \mapsto QAQ^{-1}$ is a continuous operation on matrix space.

(e) $\lambda_n A_n \to \lambda A$.

Hint: In (c) add and subtract $A_n B$, then apply the triangle inequality.

The triangle inequality has a "converse" that is sometimes indispensable.

3.4. Proposition (Reverse Triangle Inequality). For $A, B \in \mathrm{M}(N, \mathbb{K})$ we have

$$\Big|\, \|A\|_\infty - \|B\|_\infty \,\Big| \le \|A - B\|_\infty$$

Proof: By the Triangle Inequality

$$\|A + B\|_\infty \le \|A\|_\infty + \|B\|_\infty$$

Thus

$$\|A\|_\infty = \|A - B + B\|_\infty \le \|A - B\|_\infty\ , + \|B\|_\infty$$

so that $\|A\|_\infty - \|B\|_\infty \leq \|A-B\|_\infty$. Reversing roles of A, B we also get $\|B\|_\infty - \|A\|_\infty \leq \|A-B\|_\infty$. Since the absolute value of a real number is either $|c| = c$ or $-c$, we conclude that

$$\Big|\, \|A\|_\infty - \|B\|_\infty \,\Big| \leq \|A-B\|_\infty \quad \square$$

As an immediate consequence we have

3.5. Corollary. If $A_n \to A$ in $\mathrm{M}(n,\mathbb{C})$ then $\|A_n\|_\infty \to \|A\|_\infty$ in $\mathbb{R}$. $\square$

3.6. Exercise. If A in $\mathrm{M}(n,\mathbb{C})$ is an *invertible* matrix and $A_n \to A$ in the sup-norm, prove that

(a) $\det(A_n) \to \det(A)$;

(b) $A_n^{-1} \to A^{-1}$ in the sup norm.

Hint: Recall Cramer's Rule for computing A^{-1} for a nonsingular matrix A.

Application #1: Computing the Exponential e^A of a Matrix. We will show that the exponential series

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k \qquad (A \in \mathrm{M}(N,\mathbb{C})\,)$$

converges in the sup-norm, which means that the finite partial sums of the series

$$S_n = I + A + \frac{A^2}{2!} + \ldots + \frac{A^n}{n!} \qquad n \in \mathbb{N}$$

converge to a definite limit e^A in matrix space:

$$\|S_n - e^A\|_\infty \to 0 \quad \text{as} \quad n \to \infty$$

This is not so easy to prove, but if $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ is a *diagonal* matrix

$$D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix}$$

it is quite obvious that the partial sums S_n converge in the sup-norm,

$$S_n = I + D + \ldots + \frac{D^n}{n!} = \begin{pmatrix} 1+\lambda_1+\ldots+\frac{\lambda_1^n}{n!} & & & 0 \\ & 1+\lambda_2+\ldots+\frac{\lambda_2^n}{n!} & & \\ & & \ddots & \\ 0 & & & 1+\lambda_N+\ldots+\frac{\lambda_N^n}{n!} \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} e^{\lambda_1} & & & 0 \\ & e^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_N} \end{pmatrix} \qquad \text{as } n \to \infty$$

because $e^z = \sum_{k=0}^{\infty} z^k/k!$ is absolutely convergent for every complex number $z \in \mathbb{C}$. Therefore $S_n \to e^D$ in the sup-norm and

$$e^D = \sum_{k=0}^{\infty} \frac{D^k}{k!} = \lim_{n\to\infty} S_n = \begin{pmatrix} e^{\lambda_1} & & & 0 \\ & e^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_N} \end{pmatrix}$$

A Digression: The Cauchy Convergence Criterion in Matrix Space. For matrices that are not diagonal we must prove there actually *is* a matrix e^A to which the matrix-exponential series converges in sup-norm,

$$\|S_n - e^A\|_\infty \to 0 \qquad \text{as } n \to \infty\ .$$

This follows because $\mathrm{M}(N, \mathbb{K})$ equipped with the sup-norm $\|\cdot\|_\infty$ has the following **completeness** property, similar to completeness of $\mathbb{R}^n$ and $\mathbb{C}^n$ in the Euclidean norm

$$\|\mathbf{z}\|_2 = \Big(\sum_{k=1}^{N} |z_k|^2\Big)^{1/2} \qquad \text{for } \mathbf{z} = (z_1, \ldots, z_N) \text{ in } \mathbb{C}^N\ ,$$

or completeness of the number fields $\mathbb{K} = \mathbb{R}$ and $\mathbb{C}$.

Theorem (Cauchy Convergence Criterion). *A sequence $\{A_n\}$ in $\mathrm{M}(N, \mathbb{K})$ converges to some limit $A_0 = lim_{n\to\infty} A_n$ in the $\|\cdot\|_\infty$-norm if and only if the sequence has the* **Cauchy property**

$$\|A_m - A_n\|_\infty \to 0 \quad \textit{eventually as} \quad m, n \to \infty$$

(41) *To be precise, this property means: Given any $r > 0$ we can find a cutoff $M > 0$ such that*

$$\|A_m - A_n\|_\infty < r \quad \textit{for all } m, n \geq M$$

Statement (41) is much stronger than saying *successive terms* in the sequence get close, with $\|A_{n+1} - A_n\| \to 0$ as $n \to \infty$; to verify the Cauchy criterion you must show that *all* the terms far along in $\{A_n\}$ are eventually close together as $n \to \infty$.

When you try to prove $A_n \to A_0$ by examining the distances $\|A_n - A_0\|_\infty$ you must actually have the prospective limit A_0 in hand, and that limit might be very hard to guess. The Cauchy criterion gets around this problem. You don't need to identify the value of the limit whose existence is assured in (41), because the Cauchy criterion can be verified by inspecting the terms of the given sequence $\{A_n\}$. Similarly in $\mathbb{R}$, the Integral Test of Calculus shows that the the partial sums

$$S_n = 1 + \frac{1}{2^2} + \ldots + \frac{1}{n^2} \quad \text{of the } \textit{Harmonic Series } \sum_{n=1}^{\infty} 1/n^2$$

have the Cauchy property, and hence by the completeness property (41)

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \lim_{n\to\infty} \{S_n\}$$

exists. It is a lot harder to identify this limit in "closed form," and show it is exactly $\pi^2/6$. We will see one way to do this in Chapter VI.

As for the matrix exponential series $\sum_{n=0}^{\infty} A^k/k!$ we now show that its partial sums $S_n = \sum_{k=0}^{n} A^k/k!$ have the Cauchy property in $||\cdot||_\infty$-norm. Then by completeness of $\mathrm{M}(N, \mathbb{C})$ the partial sums actually have a limit, which we name "e^A"

$$e^A = \sum_{k=1}^{\infty} \frac{A^k}{k!} = \lim_{n\to\infty} S_n$$

Proof: To verify the Cauchy property for $\{S_n\}$ we may assume $m > n$. By the Triangle Inequality and the multiplicative property (c) of Exercise 3.2 we have

$$\|S_m - S_n\|_\infty = \| \sum_{k=n+1}^{m} \frac{A^k}{k!} \|_\infty \leq \sum_{n+1}^{m} \frac{N^k \|A\|_\infty^k}{k!}$$

By the Ratio Test the Taylor series for $f(x) = e^x$ converges (to e^x) for all $x \in \mathbb{R}$:

$$e^x = \sum_{n=0}^{\infty} \frac{D^n f(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

because $D^n\{e^x\} = e^x$ for all x. Taking $x = N \cdot \|A\|_\infty$, we get

$$\sum_{k=0}^{n} \frac{N^k \|A\|_\infty^k}{k!} \to \sum_{k=0}^{\infty} \frac{N^k \|A\|_\infty^k}{k!} = e^{N\|A\|_\infty} < \infty \quad \text{as } n \to \infty \ ,$$

hence for $m \geq n$:

$$0 \leq \|S_m - S_n\|_\infty \leq \sum_{n+1}^{m} \frac{(N \cdot \|A\|_\infty)^k}{k!} = \sum_{k=n+1}^{\infty} \frac{\big(N \cdot \|A\|_\infty\big)^k}{k!} \to 0$$

as $n \to \infty$. Thus, $\{S_n\}$ is Cauchy sequence for the $\| \cdot \|_\infty$-norm and the matrix-valued series $\sum_{k=0}^{\infty} A^k/k!$ converges in $\| \cdot \|_\infty$-norm for every matrix A. $\square$

In general, it is a difficult task to directly compute the sum of a convergent series such as $e^A = \sum_{n=0}^{\infty} A^n/n!$ For instance, consider how one might try to evaluate e^A when

$$A = \begin{pmatrix} 1 & -1 \\ -6 & 2 \end{pmatrix}$$

Computing higher and higher powers A^k is computationally prohibitive, and how many terms would be needed to compute each entry of e^A with an error of at most 1×10^{-6} (6-place accuracy)?

As mentioned earlier, computing e^A is easy if $A = D = \text{diag}(\lambda_1, \ldots, \lambda_N)$ is diagonal. Then,

$$S_n = I + D + \ldots + \frac{D^n}{n!} \to \begin{pmatrix} e^{\lambda_1} & & & 0 \\ 0 & e^{\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_N} \end{pmatrix} = e^D$$

We now show that e^{tA} can be computed in closed form for all $t \in \mathbb{R}$, for any A that is diagonalizable over $\mathbb{R}$ or $\mathbb{C}$.

3.7. Example. Compute e^{tA} ($t \in \mathbb{R}$) for the matrix

$$A = \begin{pmatrix} 1 & -1 \\ -6 & 2 \end{pmatrix}$$

Solution: First observe that A is diagonalizable, with $QAQ^{-1} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} = D$ for suitably chosen Q. The eigenvalues are the roots of the characteristic polynomial

$$\begin{aligned} p_A(x) &= \det \begin{pmatrix} 1-\lambda & -1 \\ -6 & 2-\lambda \end{pmatrix} = (\lambda-2)(\lambda-1) - 6 \\ &= \lambda^2 - 3\lambda + 2 - 6 = \lambda^2 - 3\lambda - 4 = (\lambda-4)(\lambda+1) \ , \end{aligned}$$

so $\text{sp}(A) = \{4, -1\}$. The eigenspaces are computed by row reduction:

- **Eigenvalue** $\lambda = 4$:

$$(A - \lambda I) = \begin{pmatrix} -3 & -1 \\ -6 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{pmatrix}.$$

Solutions of $(A - \lambda I)X = 0$ are

$$X \in \mathbb{K}\cdot \begin{pmatrix} -\frac{1}{3} \\ 1 \end{pmatrix} = \mathbb{K}\cdot \begin{pmatrix} 1 \\ -3 \end{pmatrix} = E_{\lambda=4} .$$

- **Eigenvalue** $\lambda = -1$:

$$(A - \lambda I) = \begin{pmatrix} 2 & -1 \\ -6 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$$

Solutions of $(A - \lambda I)X = 0$ are $X \in \mathbb{K} \cdot \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} = \mathbb{K}\cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = E_{\lambda=-1}$.

Thus $\mathbb{K}^2 = E_{\lambda=4} \oplus E_{\lambda=-1}$ and $\mathfrak{Y} = \{\, \mathbf{f}_1 = (1, -3)\, ,\, \mathbf{f}_2 = (1, 2)\, \}$ is a diagonalizing basis in $\mathbb{K}^2$. On the other hand, from our discussion of "change of basis" in Chapter II we have

$$\begin{aligned} D &= \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} = [L_A]_{\mathfrak{Y}\mathfrak{Y}} = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} \cdot [L_A]_{\mathfrak{X}\mathfrak{X}} \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} \\ &= [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} \cdot A \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} \end{aligned}$$

Since

$$\begin{cases} \mathbf{f}_1 = \mathbf{e}_1 - 3\mathbf{e}_2 \\ \mathbf{f}_2 = \mathbf{e}_1 + 2\mathbf{e}_2 \end{cases} \qquad \big(\text{where } \mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\} = \text{ standard basis in } \mathbb{K}^2\big)$$

we see that $[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix}$. Then $QAQ^{-1} = D$ taking $Q^{-1} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} = \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix}$, and since $\det(Q^{-1}) = 5$ we get $Q = (Q^{-1})^{-1} = \frac{1}{5} \cdot \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}$. Now

$$\begin{cases} D = QAQ^{-1} \\ A = Q^{-1}DQ \end{cases} \quad \Rightarrow \quad A^k = (Q^{-1}DQ)\cdot(Q^{-1}DQ) \cdot \ldots \cdot (Q^{-1}DQ) = Q^{-1}D^kQ$$

for $k = 0, 1, 2 \ldots$, hence by (d) of Exercise 3.3 we get

$$\begin{aligned} e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} &= \sum_{k=0}^{\infty} \frac{(Q^{-1}DQ)^k}{k!} = \sum_{k=0}^{\infty} \frac{Q^{-1}D^kQ}{k!} \\ &= Q^{-1}\Big(\sum_{k=0}^{\infty} \frac{D^k}{k!}\Big) \cdot Q = Q^{-1}e^DQ \ , \end{aligned}$$

We conclude that

$$e^A = Q^{-1} \begin{pmatrix} e^4 & 0 \\ 0 & e^{-1} \end{pmatrix} Q$$

which exhibits e^A as a product of *just three explicit matrices.*

Similarly, for $t \in \mathbb{R}$ we compute e^{tA}

$$\begin{aligned} e^{tA} &= Q^{-1} \begin{pmatrix} e^{4t} & 0 \\ 0 & e^{-t} \end{pmatrix} Q = \begin{pmatrix} 1 & 1 \\ -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} e^{4t} & 0 \\ 0 & e^{-t} \end{pmatrix} \cdot \frac{1}{5} \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \\ &= \frac{1}{5} \begin{pmatrix} 2e^{4t} + 3e^{-t} & -e^{4t} + e^{-t} \\ -6e^{4t} + 6e^{-t} & 3e^{4t} + 2e^{-t} \end{pmatrix} = \frac{1}{5} e^{4t} \cdot \begin{pmatrix} 2 & -1 \\ -6 & 3 \end{pmatrix} + \frac{1}{5} e^{-t} \cdot \begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix} \end{aligned}$$

Setting $t = 0$, we get $e^0 = I$; setting $t = 1$, we get the answer to our original question

$$e^A = \frac{1}{5}e^4 \cdot \begin{pmatrix} 2 & -1 \\ -6 & 3 \end{pmatrix} + \frac{1}{5}e^{-1} \cdot \begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix} \qquad \square$$

Application #2: Solving Linear Systems of Differential Equations.

In the next application we see why one might want to compute the matrix-valued function $\phi(t) = e^{tA}$, $\phi : \mathbb{R} \to \mathrm{M}(N, \mathbb{C})$. First we must sketch some additional properties of the exponential map on matrices (mostly without proofs).

1. *If A and B commute then*

   Exponent Law: $e^{A+B} = e^A \cdot e^B$

   In particular, e^A is always invertible, with $(e^A)^{-1} = e^{-A}$. Futhermore,

   One-Parameter Group Law: $e^{(s+t)A} = e^{sA} \cdot e^{tA}$ *for all $s, t \in \mathbb{R}$*

   and e^{-tA} is the inverse of e^{tA} for $t \in \mathbb{R}$.

Proof (sketch): We give an informal proof involving rearrangement of a matrix-valued double series. But beware: rearrangement and regrouping of series are delicate matters even for scalar-valued series, and a proof that would pass muster with analysts would require considerably more detail – see any text on Mathematical Analysis.

The series $e^A = \sum_{k=0}^{\infty} A^k/k!$ and $e^B = \sum_{\ell=0}^{\infty} B^\ell/\ell!$ are sup-norm convergent. Expanding the product of the two series term-by-term (which in itself requires some justification!) we get

$$\begin{aligned}
e^A \cdot e^B &= \Big(\sum_{k=0}^{\infty} A^k/k!\Big) \cdot \Big(\sum_{\ell=0}^{\infty} B^\ell/\ell!\Big) = \sum_{k,\ell \ge 0} \frac{1}{k!}\frac{1}{\ell!}\, A^k B^\ell \\
&= \sum_{k,\ell \ge 0} \frac{1}{(k+\ell)!} \cdot \frac{(k+\ell)!}{k!\ell!}\, A^k B^\ell \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \cdot \Big(\sum_{k=0}^{n} \tbinom{n}{k}\, A^k B^\ell\Big) \qquad \text{where } \tbinom{n}{k} = \text{(binomial coefficient)} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} (A+B)^n \qquad \text{(Binomial Formula)} \\
&= e^{A+B} \quad \square
\end{aligned}$$

2. Differentiation Law. *The derivative of $\phi(t) = e^{tA}$ exists and is continuous, with*

$$\frac{d}{dt}\big(e^{tA}\big) = A \cdot e^{tA} \qquad \text{for all } t \in \mathbb{R}$$

Proof: Using the Exponent Law we get

$$\begin{aligned}
\frac{d}{dt}\big(e^{tA}\big) &= \lim_{\Delta t \to 0} \frac{e^{(t+\Delta t)A} - e^{tA}}{\Delta t} \\
&= \lim_{\Delta t \to 0} \Big(\frac{e^{(\Delta t)A} - I}{\Delta t}\Big) \cdot e^{tA} \qquad \text{(since } e^{(t+\Delta t)A} = e^{tA} \cdot e^{(\Delta t)A}\text{)} \\
&= \Big(\lim_{\Delta t \to 0} \frac{e^{(\Delta t)A} - I}{\Delta t}\Big) \cdot e^{tA}
\end{aligned}$$

Using the norm properties listed in Exercises 3.2 -3.3 it is not hard to show that

$$\begin{aligned} e^{(\Delta t)A} - I &= \Big(I + (\Delta t)A + \frac{(\Delta t)^2}{2!}A^2 + ...\Big) - I \\ &= \Delta t \cdot \Big(A + \frac{(\Delta t)^2}{2!}A^2 + ...\Big) \;=\; \Delta t\big(A + \mathcal{O}(\Delta t)\big) \end{aligned}$$

where the matrix-valued remainder $\mathcal{O}(\Delta t)$ becomes very small compared to Δt

$$\frac{\|\mathcal{O}(\Delta t)\|_\infty}{|\Delta t|} \to 0 \qquad \text{as } \Delta t \to 0\ .$$

Thus,

$$\frac{e^{(\Delta t)A} - I}{\Delta t} = \frac{\Delta t}{\Delta t}\big(A + \mathcal{O}(\Delta t)\big) \to A$$

in the $\|\cdot\|_\infty$-norm as $\Delta t \to 0$, proving the formula. $\square$

Any system of n first order constant coefficient linear ordinary differential equations in n unknowns can be written in matrix form as

$$\frac{d\mathbf{y}}{dt} = A \cdot \mathbf{y}(y) \qquad \text{with initial condition } \mathbf{y}(0) = \mathbf{c} \tag{42}$$

where $\mathbf{y}(t) = \big(y_1(t), \ldots, y_n(t)\big)$ is a vector-valued function of t, and the $n \times n$ matrix A provides the coefficients of the system. It is well known that once the initial value $\mathbf{c}$ is specified there is a unique infinitely differentiable vector-valued solution $\mathbf{y}(t)$ if we regard $\mathbf{y}(t)$ as an $n \times 1$ column vector. The solution can be computed explicitly as

$$\mathbf{y}(y) = e^{tA} \cdot \mathbf{y}(0) = e^{tA} {\cdot} \mathbf{c} \qquad \text{for } t \in \mathbb{R} \tag{43}$$

In fact,

$$\frac{d\mathbf{y}}{dt} = \frac{d}{dt}\big(e^{tA} \cdot \mathbf{c}\big) = \frac{d}{dt}\big(e^{tA}\big) \cdot \mathbf{c} = Ae^{tA} \cdot \mathbf{c} = A \cdot \mathbf{y}(t)\ ,$$

and when $t = 0$ we get $\mathbf{y}(0) = \mathbf{c}$ because $e^{0 \cdot A} = I_{n \times n}$. We must of course compute e^{tA} to arrive at $\mathbf{y}(t)$ but in the previous example we have seen how that might be done, at least when the coefficient matrix can be diagonalized.

3.8. Example. If $A = \begin{pmatrix} 1 & -1 \\ -6 & 2 \end{pmatrix}$ determine the unique solution of the first order vector-valued differential equation

$$\frac{d\mathbf{y}}{dt} = A \cdot \mathbf{y}(t) \text{ such that } \mathbf{y}_0 = \mathbf{y}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Likewise for the initial value $\mathbf{y}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Then find all solutions of

$$\frac{d\mathbf{y}}{dt} = A \cdot \mathbf{y}(t) \qquad \text{for an arbitrary initial value } \mathbf{y}_0 = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

Solution: Earlier we found that

$$QAQ^{-1} = \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{for} \qquad Q = \frac{1}{5}\begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix}$$

and showed that

$$e^A = e^{QDQ^{-1}} = Q^{-1}\cdot e^D\cdot Q = \frac{1}{5}e^4\cdot\begin{pmatrix} 2 & -1 \\ -6 & 3 \end{pmatrix} + e^{-1}\cdot\begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix}$$

Taking e^{tA} in place of e^A, we got (with little additional effort):

$$e^{tA} = \frac{1}{5}e^{4t}\cdot\begin{pmatrix} 2 & -1 \\ -6 & 3 \end{pmatrix} + \frac{1}{5}e^{-t}\cdot\begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix} \qquad \text{for all } t\in\mathbb{R}$$

Taking $\mathbf{y}_0 = \mathbf{e}_1 = (1,0)$ we get a solution:

$$\mathbf{y}_1(t) = e^{tA}(\mathbf{e}_1) = \frac{1}{5}e^{4t}\cdot\begin{pmatrix} 2 \\ -6 \end{pmatrix} + \frac{1}{5}e^{-t}\cdot\begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

If $\mathbf{y}_0 = \mathbf{e}_2 = (0,1)$ we get another solution:

$$\mathbf{y}_2(t) = e^{tA}(\mathbf{e}_2) = \frac{1}{5}e^{4t}\begin{pmatrix} -1 \\ 3 \end{pmatrix} + \frac{1}{5}e^{-t}\cdot\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

For an arbitrary initial condition $\mathbf{y}(0) = \mathbf{c} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2$, it is obvious that the solution of $d\mathbf{y}/dt = A\cdot\mathbf{y}(t)$ with this initial condition is the same linear combination of the "basic solutions" $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ namely:

$$\begin{aligned}
\mathbf{y}(t) &= c_1\mathbf{y}_1(t) + c_2\mathbf{y}_2(t) \\
&= \frac{1}{5}e^{4t}\cdot\Big[c_1\begin{pmatrix} 2 \\ -6 \end{pmatrix} + c_2\begin{pmatrix} -1 \\ 3 \end{pmatrix}\Big] + \frac{1}{5}e^{-t}\Big[c_1\begin{pmatrix} 3 \\ 6 \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 2 \end{pmatrix}\Big]
\end{aligned}$$

(Check for yourself that $\mathbf{y}(0) = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 = \mathbf{c}$.)

The full set of differentiable maps $f:\mathbb{R}\to\mathbb{C}^2$ such that $df/dt = A\cdot f(t)$ is a 2-dimensional subspace M in the ∞-dimensional space $\mathcal{C}^\infty(\mathbb{R},\mathbb{C}^2)$ of infinitely differentiable vector valued maps:

$$M = \mathbb{C}\text{-span}\{\mathbf{y}_1(t),\mathbf{y}_2(t)\} = \{c_1\mathbf{y}_1 + c_2\mathbf{y}_2 : c_1,c_2\in\mathbb{C}\}$$

and the "basic solutions" $\mathbf{y}_1,\mathbf{y}_2$ are a vector basis for M. One should check that $\mathbf{y}_1$, $\mathbf{y}_2$ are linearly independent vectors in $\mathcal{C}^\infty(\mathbb{R},\mathbb{C}^2)$. But if there were coefficients α_1,α_2 such that $\alpha_1\mathbf{y}_1(t) + \alpha_2\mathbf{y}_2(t)\equiv 0$ in $\mathbb{C}^2$, and we take any convenient base point (say $t=0$), we would then have the following vector identity in $\mathbb{C}^2$:

$$\begin{aligned}
\begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \frac{\alpha_1}{5}\Big[\begin{pmatrix} 2 \\ -6 \end{pmatrix} + \begin{pmatrix} 3 \\ 6 \end{pmatrix}\Big] + \frac{\alpha_2}{5}\Big[\begin{pmatrix} -1 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}\Big] \\
&\Rightarrow \frac{\alpha_1}{5}\begin{pmatrix} 5 \\ 0 \end{pmatrix} + \frac{\alpha_2}{5}\begin{pmatrix} 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
&\Rightarrow \alpha\mathbf{e}_1 + \alpha_2\mathbf{e}_2 = 0 \\
&\Rightarrow \alpha_1 = \alpha_2 = 0
\end{aligned}$$

as required. $\square$

A similar discussion holds for equations $d\mathbf{y}/dt = A\cdot\mathbf{y}(t)$ when A is $n\times n$ (and diagonalizable). If $\{\mathbf{y}_1(t),\dots,\mathbf{y}_n(t)\}\subseteq\mathcal{C}^\infty(\mathbb{R},\mathbb{C}^n)$ are the "basic solutions," whose initial values are $\mathbf{y}_k(0) = \mathbf{e}_k$ (the standard basis vectors in $\mathbb{C}^n$), then a solution with arbitrary initial value $\mathbf{y}(0) = \sum_{k=1}^n c_k\mathbf{e}_k\in\mathbb{C}^n$ is obtained by taking the same linear combination

$$\mathbf{y}(t) = c_1\mathbf{y}_1(t) + \ldots + c_n\mathbf{y}_n(t) \quad .$$

of basic solution $\mathbf{y}_k(t)$. As above, the $\mathbf{y}_k$ are linearly independent vectors in $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C}^n)$: if $\mathbf{0} = \sum c_k \mathbf{y}_k(t)$ in $\mathcal{C}^\infty(\mathbb{R}, \mathbb{C}^n)$ for all t, then (taking $t = 0$) $\sum c_k \mathbf{e}_k = \mathbf{0}$ in $\mathbb{C}^n$; thus, $c_1 = c_2 = \ldots = c_n = 0$ because $\mathbf{y}_k(0) = \mathbf{e}_k$, by definition. We conclude that the $\{\mathbf{y}_k(t)\}$ are a basis for the full set of solutions (with arbitrary initial value) of the equation $d\mathbf{y}/dt = A \cdot \mathbf{y}(t)$.

$$\begin{aligned} M &= \left\{ f \in \mathcal{C}^\infty : \frac{df}{dt} = A \cdot f(t) \text{ for all } t \in \mathbb{R} \right\} \qquad (f : \mathbb{R} \to \mathbb{C}^n\ ) \\ &= \mathbb{C}\text{-span}\{\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t)\} \end{aligned}$$

which has dimension $\dim_{\mathbb{C}}(M) = n$. $\square$

Chapter VII. Nondiagonalizable Operators.

VII-1. Basic Definitions and Examples.

We continue the convention of previous chapters. writing $\dim(V) = |V|$ where appropriate Nilpotent operators present the first serious obstruction to attempts to diagonalize a given linear operator.

1.1. Definition. *A linear operator* $T : V \to V$ *is* **nilpotent** *if* $T^k = 0$ *for some* $k \in \mathbb{N}$*; it is* **unipotent** *if* $T = I + N$ *with* N *nilpotent.*

Obviously T is unipotent $\Leftrightarrow T - I$ is nilpotent.

Nilpotent operators cannot be diagonalized unless T is the zero operator (or $T = I$, if unipotent). Any analysis of normal forms must examine these operators in detail. Nilpotent and unipotent matrices $A \in \mathrm{M}(n, \mathbb{F})$ are defined the same way. As examples, all strictly upper triangular matrices (with zeros on the diagonal) as well as those that are strictly lower triangular, are nilpotent in view of the following observations.

1.2. Exercise. If A has upper triangular form with zeros on and below the diagonal, prove that

$$A^2 = \begin{pmatrix} 0 & 0 & & * \\ & \cdot & \cdot & \\ & & \cdot & 0 \\ 0 & & & 0 \end{pmatrix} \qquad A^3 = \begin{pmatrix} 0 & 0 & 0 & & * \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & 0 \\ & & & \cdot & 0 \\ 0 & & & & 0 \end{pmatrix},$$

etc, so that $A^n = 0$. $\square$

Matrices of the same form, but with 1's on the diagonal all correspond to unipotent operators.

We will see that if $N : V \to V$ is nilpotent there is a basis $\mathfrak{X}$ such that

$$[N]_{\mathfrak{X}} = \begin{pmatrix} 0 & & * \\ & \cdot & \\ & & \cdot \\ 0 & & 0 \end{pmatrix},$$

but this is not true for all bases. Furthermore, a lot more can be said about the terms $(*)$ for suitably chosen bases.

1.3. Exercise. In $\mathrm{M}(n, \mathbb{F})$, show that the sets of upper triangular matrices:

(a) The **strictly upper triangular group** $\mathcal{N} = \left\{ \begin{pmatrix} 1 & & * \\ & \cdot & \\ & & \cdot \\ 0 & & 1 \end{pmatrix} \right\}$ with entries in $\mathbb{F}$.

(b) The **full upper triangular group** in $\mathrm{M}(n, \mathbb{F})$, $\mathcal{P} = \left\{ \begin{pmatrix} a_{1,1} & & * \\ & \cdot & \\ & & \cdot \\ 0 & & a_{n,n} \end{pmatrix} \right\}$ with entries in $\mathbb{F}$ such that $\prod_{i=1}^n a_{i,i} \neq 0$.

are both subgroups in $\mathrm{GL}(n, \mathbb{F})$, with $\det(A) = \prod_{i=1}^{n} a_{i,i} \neq 0$ for elements of either group. Verify that $\mathcal{N}$ and $\mathcal{P}$ are closed under taking products and inverses. $\square$

1.4. Exercise. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ in $\mathrm{M}(2, \mathbb{F})$. This is a nilpotent matrix and in any ground field the only root of its characteristic polynomial

$$p_A(\lambda) = \det(A - \lambda I) = \lambda^2$$

is $\lambda = 0$. There is a nontrivial eigenvector $e_1 = (1, 0)$, corresponding to eigenvalue $\lambda = 0$, because $\ker(A) = \mathbb{F} \cdot e_1$ is nontrivial (as it must be for any nilpotent operator). But you can easily verify that scalar multiples of e_1 are the only eigenvectors, so there is no basis of eigenvectors. A cannot be diagonalized by any similarity transformation, Regardless of the ground field $\mathbb{F}$. $\square$

"Stable Range" and "Stable Kernel" of a Linear Map. If $T : V \to V$ is a linear operator on a finite dimensional vector space (arbitrary ground field), let $K_i = K(T^i) = \ker(T^i)$ and $R_i = R(T^i) = \mathrm{range}(T^i)$ for $i = 0, 1, 2, \cdots$. Obviously these spaces are nested

$$\begin{gathered} (0) \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_i \subseteq K_{i+1} \subseteq \cdots \\ V \supseteq R_1 \supseteq R_2 \supseteq \cdots \supseteq R_i \supseteq R_{i+1} \supseteq \cdots \ , \end{gathered}$$

and if $\dim(V) < \infty$ they must each stabilize at some point, say with $K_r = K_{r+1} = \cdots$ and $R_s = R_{s+1} = \cdots$ for some integers r and s. In fact if r is the first (smallest) index such that $K_r = K_{r+1} = \cdots$ the sequence of ranges must also stabilize at the same point because $|V| = |K_i| + |R_i|$ at each step. With this in mind, we define (for finite dimensional V)

$$\begin{aligned} R_\infty &= \bigcap_{i=1}^{\infty} R_i \ = \ R_r \ = \ R_{r+1} \ = \ \cdots \quad (\textbf{Stable range} \text{ of } T) \\ K_\infty &= \bigcup_{i=1}^{\infty} K_i \ = \ K_r \ = \ K_{r+1} \ = \ \cdots \quad (\textbf{Stable kernel} \text{ of } T) \end{aligned}$$

1.5. Proposition. *$V = R_\infty \oplus K_\infty$ and the spaces R_∞, K_∞ are T-invariant. Furthermore $R_{i+1} \neq R_i$ and $K_{i+1} \neq K_i$ for $i < r$.*

Note: This splitting is sometimes referred to as the "Fitting decomposition" (after a guy named Fitting).

Proof: To see there is a non-trivial jump $R_{i+1} \subsetneqq R_i$ at every step until $i = r$ if suffices to show that $R_{i+1} = R_i$ at some step implies $R_i = R_j$ for all $j \geq i$ (a similar result for kernels then follows automatically). It suffices to show that $R_i = R_{i+1} \Rightarrow R_{i+1} = R_{i+2}$. Obviously, $R_{i+2} \subseteq R_{i+1}$ for all i; to prove the reverse inclusion $R_{i+1} \subseteq R_{i+2}$, let $v \in R_{i+1}$. Then there is some $w_1 \in V$ such that $v = T^{i+1}(w_1) = T\big(T^i(w_1)\big)$. By hypothesis $R_{i+1} = T^{i+1}(V) = R_i = T^i(V)$ so there is some $w_2 \in V$ such that $T^i(w_1) = T^{i+1}(w_2)$. Thus

$$v = T^{i+1}(w_2) = T(T^i(w_1)) = T(T^{i+1}(w_2)) = T^{i+2}(w_2) \in R_{i+2}$$

So, $R_{i+1} \subseteq R_{i+2}$, $R_i = R_{i+1} = R_{i+2}$, and by induction $R_i = R_{i+1} = \cdots = R_\infty$.

For T-invariance of $R_\infty = R_r$ and $K_\infty = K_r$, T maps $R_i \to R_{i+1} \subseteq R_i$ for all i; taking $i = r$, we get $T(R_\infty) = R_\infty$. As for the kernels, if $v \in K_{i+1}$ then $0 = T^{i+1}(v) = T^i(T(v))$. As a consequence, $T(v) \in K_i$ and $T(K_{i+1}) \subseteq K_i$ for all i. For $i \geq r$, we have $K_i = K_{i+1} = K_\infty$, so $T(K_\infty) = K_\infty$ as claimed.

To see $V = K_\infty \oplus R_\infty$ we show (i) $R_\infty + K_\infty = V$ and (ii) $R_\infty \cap K_\infty = \{0\}$. For (ii), if $v \in R_\infty = R_r$ there is some $w \in V$ such that $T^r(w) = v$; but if $v \in K_\infty = K_r$, then $T^r(v) = 0$ and hence $T^r(v) = 0$. Consequently $T^{2r}(w) = T^r(v) = 0$. We now observe that $T : R_i \to R_{i+1}$ is a bijection for $i \geq r$ so $\ker(T|_{R_r}) = \ker(T|_{R_\infty}) = \{0\}$. In fact, if $i \geq r$ then $R_i = R_{i+1}$ and $T : R_i \to R_{i+1}$ is a surjective linear map, and if $T : R_i \to R_{i+1} = R_i$ is surjective it is automatically a bijection. Now in the preceding discussion $v = T^r(w) \in R_r$ and $T^r : R_r \to R_{2r} = R_r$ is a bijection, so

$$0 = T^{2r}(w) = T^r(T^r(w)) = T^r(v)$$

Then $v = 0$, hence $R_\infty \cap K_\infty = \{0\}$

For $(ii) \Rightarrow (i)$, we know

$$\begin{aligned} |R_\infty + K_\infty| &= |R_r + K_r| = |R_r| + |K_r| - |K_r \cap R_r| \\ &= |K_\infty| + |R_\infty| = |K_r| + |R_r| = |V| \end{aligned}$$

(by the Dimension Theorem). We conclude that $R_\infty + K_\infty = V$, proving (i). $\square$

1.6. Lemma. *$T|_{K_\infty}$ is a nilpotent operator on K_∞ and $T|_{R_\infty}$ is a bijective linear map of $R_\infty \to R_\infty$. Hence, every linear operator T on a finite dimensional space V, over any field, has a direct sum decomposition.*

$$T = (T|R_\infty) \oplus (T|K_\infty)$$

such that $T|_{K_\infty}$ is nilpotent and $T|_{R_\infty}$ bijective on R_∞.

Proof: $T^r(K_\infty) = T^r(\ker(T^r)) = \{0\}$ so $(T|_{K_\infty})^r = 0$ and $T|_{K_\infty}$ is nilpotent of degree $\leq r$, the index at which the ranges stabilize at R_∞.

VII-2. Some Observations about Nilpotent Operators.

2.1. Lemma. *If $N : V \to V$ is nilpotent, the unipotent operator $I + N$ is invertible.*

Proof: If $N^k = 0$ the geometric series $I + N + N^2 + \ldots + N^{k-1} + \ldots = \sum_{k=0}^\infty N^k$ is finite and a simple calculation shows that

$$(I - N)(I + N + \cdots + N^{k-1}) = I - N^k = I \ .$$

Hence

$$(I - N)^{-1} = I + N + \cdots + N^{k-1} \qquad \square \tag{1}$$

if $N^k = 0$. $\square$

2.2. Lemma. *If $T : V \to V$ is nilpotent then $p_T(\lambda) = \det(T - \lambda I)$ is equal to $(-1)^n\lambda^n$ ($n = \dim(V)$), and $\lambda = 0$ is the only eigenvalue (over any field $\mathbb{F}$). [It is an eigenvalue since $\ker(T) \neq \{0\}$ and the full subspace of $\lambda = 0$ eigenvectors is precisely $E_{\lambda=0}(T) = \ker(T)$].*

Proof: Take a basis $\mathfrak{X} = \{e_1, \cdots, e_n\}$ that runs first through $K(T) = K_1 = \ker(T)$, then augments to a basis in $K_2 = \ker(T^2)$, etc. With respect to this basis $[T]_{\mathfrak{X}\mathfrak{X}}$ is an upper triangular matrix with zero matrices blocks on the diagonal (see Exercise 2.4 below). Obviously, $T - \lambda I$ has diagonal values $-\lambda$, so $\det(T - \lambda I) = (-1)^n\lambda^n$ as claimed. $\square$

Similarly a unipotent operator T has $\lambda = 1$ as its only eigenvalue (over any field) and its characteristic polynomial is $p_T(x) = \mathbb{1}$ (constant polynomial $\equiv 1$). The sole eigenspace $E_{\lambda=1}(T)$ is the set of *fixed points* $\text{Fix}(T) = \{v : T(v) = v\}$.

2.3. Exercise. Prove that

(a) A nilpotent operator T is diagonalizable (for some basis) if and only if $T = 0$.

(b) T is unipotent if and only if T is the identity operator $I = \text{id}_V$ $\square$

2.4. Exercise. If $T : V \to V$ is a nilpotent linear operator on a finite dimensional space let $\mathfrak{X} = \{e_1, \ldots, e_n\}$ is a basis that passes through successive kernels $K_i = \ker(T^i)$, $1 \le i \le d = \deg(T)$. Prove that $[T]_{\mathfrak{X}}$ is upper triangular with $m_i \times m_i$ zero-blocks on the diagonal, $m_i = \dim(K_i/K_{i-1})$.
Hints: The problem is to devise efficient notation to handle this question. Partition the indices $1, 2, \ldots, n$ into consecutive intervals $J_1, \ldots, J_d$ ($d = \deg(T)$) such that $\{e_j : j \in J_1\}$ is a basis for K_1, $\{e_i : i \in J_1 \cup J_2\}$ is a basis for K_2, etc. Matrix coefficients T_{ij} are determined by the system of vector equations

$$T(e_i) = \sum_{j=1}^{n} T_{ji} e_j \qquad (1 \le i \le n = \dim(V))$$

What do the inclusions $T(K_i) \subseteq K_{i-1}$ tell you about the coefficients T_{ij}? $\square$

Let $T : V \to V$ be nilpotent. The powers T^k eventually "kill" every vector $v \neq 0$, so there is an $m \in \mathbb{N}$ such that $\{v, T(v), \cdots, T^{m-1}(v)\}$ are nonzero and $T^m(v) = 0$. The **nilpotence degree** $\deg(T)$ is the smallest exponent $d = 0, 1, 2, \cdots$ such that $T^d = 0$.

2.5. Proposition. *Let $T : V \to V$ be nilpotent and $v_0 \neq 0$. If $v_0, T(v_0), \cdots, T^{m-1}(v_0)$ are all nonzero and $T^m(v_0) = 0$ define $W(v_0) = \mathbb{F}\text{−span}\{v_0, T(v_0), \cdots, T^{m-1}(v_0)\}$. This subspace is T-invariant and the vectors $\{v_0, T(v_0), \cdots, T^{m-1}(v_0)\}$ are independent, hence a basis for this "cyclic subspace" determined by v_0 and the action of T.*

Proof: The $\{T^k(v_0) : 0 \le k \le n-1\}$ span $W(v_0)$ by definition. They are independent because if $0 = c_0 + c_1T(v_0) + \cdots + c_{m-1}T^{m-1}(v_0)$ for some choice of $c_k \in \mathbb{F}$, then

$$\begin{aligned} 0 = T^{m-1}(0) &= T^{m-1}\big(c_0v_0 + c_1T(v_0) + \cdots + c_{m-1}T^{m-1}(v_0)\big) \\ &= c_0T^{m-1}(v_0) + c_1 \cdot 0 + \cdots + c_{m-1} \cdot 0 \ , \end{aligned}$$

which implies $c_0 = 0$ since $T^{m-1}(v_0) \neq 0$ by minimality of the exponent m. Next, apply T^{m-2} to the original sum, which has now the form $c_1T(v_0) + \cdots + c_{m-1}T^{m-1}(v_0)$; we get

$$T^{m-2}(0) = T^{m-2}\big(c_1T(v_0) + \cdots + c_{m-1}T^{m-1}(v_0)\big) = c_1T^{m-1}(v_0) + 0 + \cdots + 0$$

and then $c_1 = 0$. We can apply the same process repeatedly to get $c_0 = c_1 = c_2 = \cdots = c_{m-1} = 0$. $\square$

Obviously $W(v_0)$ is T-invariant and $T_0 = T|_{W(v_0)}$ is nilpotent (with degree $m = \deg(T_0) \le \deg(T)$) because for each basis vector $T^k(v_0)$ we have $T_0^m(T^k(v_0)) = T^k(T^m(v_0)) = 0$; but in fact $\deg(T_0) = m$ because $T_0^{m-1}(v_0) \neq \{0\}$. Now consider the ordered basis

$$\mathfrak{X} = \{e_1 = T^{m-1}(v_0), e_2 = T^{m-2}(v_0), \cdots, e_m = v_0\} \text{ in } W(v_0) \ .$$

Since $T(e_{k+1}) = e_k$ for each $k \ge 1$ and $T(e_1) = 0$, the matrix $[T]_{\mathfrak{X},\mathfrak{X}}$ has the form

$$[T]_{\mathfrak{X}} = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & & \cdot & \\ 0 & 0 & 0 & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix}$$

The action on these ordered basis vectors is :

$$0 \xleftarrow{T} e_1 \xleftarrow{T} e_2 \xleftarrow{T} \cdots \xleftarrow{T} e_{m-1} \xleftarrow{T} e_m = v_0$$

The "top vector" $e_m = v_0$ is referred to as a **cyclic vector** for the invariant subspace $W(v_0)$. Any matrix having the form

$$\begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & & \cdot & \\ 0 & 0 & 0 & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix}$$

is called an **elementary nilpotent matrix**.

Cyclic Vectors and Cyclic Subspaces for General Linear Operators. To put this in its proper context we leave the world of nilpotent operators for a moment.

2.6. Definition. *If* $\dim(V) < \infty$*,* $T : V \to V$ *is a linear operator, and* $W \subseteq V$ *a nonzero* T*-invariant subspace, we say* W *is a* **cyclic subspace** *if it contains a "cyclic vector"* $v_0 \in W$ *such that* $W = \mathbb{F}\text{-span}\{v_0, T(v_0), T^2(v_0), \cdots\}$.

Only finitely many iterates $T^i(v_0)$ under the action of T can be linearly independent, so there will be a first (smallest) exponent $k = k(v_0)$ such that $\{v_0, T(v_0), \cdots, T^{k-1}(v_0)\}$ are linearly independent and $T^k(v_0)$ is a linear combination of the previous vectors.

2.7. Proposition. *Let* $T : V \to V$ *be an arbitrary linear operator on a finite dimensional vector space. If* $v_0 \in V$ *is non-zero there is a unique exponent* $k = k(v_0) \geq 1$ *such that* $\{v_0, T(v_0), \cdots, T^{k-1}(v_0)\}$ *are linearly independent and* $T^k(v_0)$ *is a linear combination of these vectors. Obviously,*

$$W = \mathbb{F}\text{−span}\{T^j(v_0) : j = 0, 1, 2, \cdots\} = \mathbb{F}\text{−span}\{v_0, T(v_0), \cdots, T^{k-1}(v_0)\}$$

and $\dim(W) = k$*. Furthermore,* $T(W) \subseteq W$ *and* W *is a cyclic subspace in* V*.*

Proof: By definition of $k = k(v_0)$, $T^k(v_0)$ is a linear combination $T^k(v_0) = \sum_{j=0}^{k-1} c_j T^j(v_0)$. Arguing recursively,

$$\begin{aligned} T^{k+1}(v_0) &= T(T^k(v_0)) = \sum_{j=0}^{k-1} c_j T^{j+1}(v_0) \\ &= (c_{k-1}T^k(v_0)) + \big(\text{linear combinations of } v_0, T(v_0), \cdots, T^{k-1}(v_0)\big) \end{aligned}$$

Since we already know $T^k(v_0)$ lies in $\mathbb{F}$-span$\{v_0, T(v_0), \cdots, T^{k-1}(v_0)\}$, so does $T^{k+1}(v_0)$. Continuing this process, we find that all iterates $T^i(v_0)$ $(i \geq k)$ lie in W. By definition $v_0, T(v_0), \cdots, T^{k-1}(v_0)$ are linearly independent and span W, so $\dim(W) = k$. □

When T is nilpotent there is a simpler alternative description of the cyclic subspace W generated by the action of T on $v_0 \neq 0$. Since $T^d = 0$ on all of V when $d = \deg(T)$, there is a smallest exponent l such that $\{v_0, T(v_0), \cdots, T^{\ell-1}(v_0)\}$ are nonzero and $T^\ell(v_0) = T^{\ell+i}(v_0) = 0$ for all $i \geq 0$. These vectors are independent and the next vector $T^\ell(v_0) = 0$ lies in $\mathbb{F}$-span$\{v_0, T(v_0), \cdots, T^{\ell-1}(v_0)\}$, so ℓ is precisely the exponent of the previous lemma and $C = \mathbb{F}$-span$\{v_0, T(v_0), \cdots, T^{\ell-1}(v_0)\}$ is the cyclic subspace generated by v_0.

XII-3. Structure of Nilpotent Operators.

Resuming the discussion of nilpotent operators, we first observe that if $T : V \to V$ is nilpotent and nonzero the chain of kernels $K_i = \ker(T^i)$,

$$\{0\} = K_0 \underset{\neq}{\subseteq} K_1 = \ker(T) \underset{\neq}{\subseteq} K_2 \underset{\neq}{\subseteq} \cdots \underset{\neq}{\subseteq} K_d = V \qquad (d = \deg(T))$$

terminates at V in finitely many steps. The difference sets partition $V \sim (0)$ into disjoint "layers"

$$V \sim (0) = \big(K_d \sim K_{d-1}\big) \cup \cdots \cup \big(K_i \sim K_{i-1}\big) \cup \cdots \cup \big(K_1 \sim K_0\big)$$

where $K_0 = (0)$. The layers $K_i \sim K_{i-1}$ correspond to the quotient spaces K_i/K_{i-1}, and by examining the action of T on these quotients we will be able to determine the structure of the operator T.

3.1. Exercise. If v_0 is in the "top layer" $V \sim K_{d-1}$, prove that $\mathbb{F}\text{-span}\{T^j(v_0) : j \geq 0\}$ has dimension d and *every* such v_0 is a cyclic vector under the iterated action of T on W. $\square$

Since $\dim(K_{d-1}) < \dim(K_d) = \dim(V)$, K_{d-1} is a very thin subset of V and has "measure zero" in V when $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. If you could pick a vector $v_0 \in V$ "at random," you would have $v_0 \in V \sim K_{d-1}$ "with probability 1," and every such choice of v_0 would generate a cyclic subspace of dimension d. "Unsuccessful" choices, which occur with "probability zero," yield cyclic subspaces $W(v_0)$ of dimension $< d$.

We now state the main structure theorem for nilpotent operators .

3.2. Theorem (Cyclic Subspace Decomposition). *Given a nilpotent linear operator $T : V \rightarrow V$ on a finite dimensional vector space V, there is a decomposition $V = V_1 \oplus \cdots \oplus V_r$ into cyclic T-invariant subspaces. Obviously the restrictions $T_i = T|_{V_i}$ are nilpotent, with degrees*

$$m_i = \dim(V_i) = (\textit{smallest exponent } m \textit{ such that } T^m \textit{ kills the cyclic generator } v_i \in V_i)$$

These degrees are unique when listed in descending order $m_1 \geq m_2 \geq \cdots \geq m_r > 0$ (repeats allowed), and $\sum_{i=1}^{r} m_i = \dim(V)$.

While it is nice to know such structure exists, it is equally important to develop a constructive procedure for finding suitable cyclic subspaces $V_1, \cdots, V_r$. This is complicated by the fact that the cyclic subspaces are not necessarily unique, unlike the eigenspaces $E_\lambda(T)$ associated with a diagonalizable operator. Any algorithm for constructing suitable V_i will necessarily involve some arbitrary choices.

The rest of this section provides a proof of Theorem 3.2 that yields on an explicit construction of the desired subspaces. There are some very elegant proofs of Theorem 3.2, but they are existential rather than constructive and so are less informative.

3.3. Corollary. *If $T : V \rightarrow V$ is nilpotent, there is a decomposition into cyclic spaces $V = V_1 \oplus \ldots \oplus V_r$, so there is a basis $\mathfrak{X}$ such that $[T]_{\mathfrak{X}}$ consists of elementary nilpotent diagonal blocks.*

$$[T]_{\mathfrak{X}} = \begin{pmatrix} \boxed{B_1} & 0 & 0 & \cdot & & 0 \\ 0 & \boxed{B_2} & 0 & & \cdot & \\ 0 & 0 & 0 & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \\ 0 & \cdot & \cdot & \cdot & 0 & \boxed{B_r} \end{pmatrix}$$

with

$$B_i = \begin{pmatrix} 0 & 1 & 0 & & & 0 \\ 0 & 0 & 1 & & & \cdot \\ 0 & 0 & 0 & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix}$$

We start with the special case in which T has the largest possible degree of nilpotence.

3.4. Lemma. *If T is nilpotent and $\deg(T) = \dim(V)$, there is a cyclic vector in V and a basis such that $[T]_{\mathfrak{X}}$ has the form B_i of an elementary nilpotent matrix.*

Proof: If $\deg(T) = d$ is equal to $\dim(V)$, the spaces $K_i = \ker(T^i)$ increase with $|K_{i+1}| \geq 1 + |K_i|$ at each step in the chain $\{0\} \subsetneqq K_1 \subseteq \cdots \subseteq K_{d-1} \subseteq K_d = V$. There are $d = \dim(V)$ steps so we must have $|K_{i+1}| = 1 + |K_i|$. Take any vector $v_0 \in V \sim K_{d-1}$. Then $T^d(v_0) = 0$ but by definition of K_{d-1}, $v_0, T(v_0), \cdots, T^{d-1}(v_0)$ are all nonzero, so v_0 is a cyclic vector for the iterated action of T. $\square$

If $T : V \to V$ is nilpotent of degree d, the idea behind proof of Theorem 3.1 is to look at the kernels $K_i = \ker(T^i)$.

$$V = K_d \supsetneqq K_{d-1} \supsetneqq \cdots \supsetneqq K_2 \supsetneqq K_1 = \ker(T) \supsetneqq \{0\}$$

As the kernels get smaller, more of V is "uncovered" (the difference set $V \sim K_s$ and the quotient V/K_s get bigger) and the action in V/K_s reveals more details about the full action of T on V.

It will be important to note that $T(K_i) \subseteq K_{i-1}$ (since $0 = T^i(x) = T^{i-1}(T(x))$ and $T(x) \in K_{i-1}$). Furthermore, $x \notin K_i$ implies that $0 \neq T^i(x) = T^{i-1}(T(x))$ so that $T(x) \notin K_{i-1}$. Thus

(2) $\qquad$ T maps $K_{i+1} \sim K_i$ into $K_i \sim K_{i-1}$ for all i.

But it is not generally true that $T(K_j) = K_{j-1}$.

3.5. Definition. *Let $T : V \to V$ be an arbitrary linear map and W a T-invariant subspace. We say that vectors $e_1, \cdots, e_m$ in V are:*

1. **Independent (mod W)** *if their images $\overline{e}_1, \cdots, \overline{e}_m$ in V/W are linearly independent. Since $\sum_i c_i\overline{e}_i = 0$ in V/W if and only if $\sum_i c_ie_i \in W$ in V, that means:*
$$\sum_{i=1}^{m} c_ie_i \in W \ \Rightarrow\ c_1 = \cdots = c_m = 0 \qquad (c_i \in \mathbb{F})$$
2. **Span V (mod W)** *if $\mathbb{F}$-span$\{\overline{e}_i\} = V/W$, which means: given $v \in V$, there are $c_i \in \mathbb{F}$ such that $\left(v - \sum_i c_ie_i\right) \in W$, or $\overline{v} = \sum_{i=0} c_i\overline{e}_i$ in V/W.*
3. *A* **basis for V (mod W)** *if the images $\{\overline{e}_i\}$ are a basis in V/W, which happens if and only if* 1. *and* 2. *hold.*

3.6. Exercise. Let $W \subseteq \mathbb{R}^5$ be the solution set of system

$$\begin{cases} x_1 + x_3 &= 0 \\ x_1 - x_4 &= 0 \end{cases}$$

and let $\{e_i\}$ be the standard basis in $V = \mathbb{R}^5$.

1. Find vectors v_1, v_2 that are a basis for V (mod W).
2. Is $\mathfrak{X} = \{e_1, e_2, e_3, v_1, v_2\}$ a basis for V where v_1, v_2 are the vectors in (1.)?
3. Find a basis $\{f_1, f_2, f_3\}$ for the subspace W. $\square$

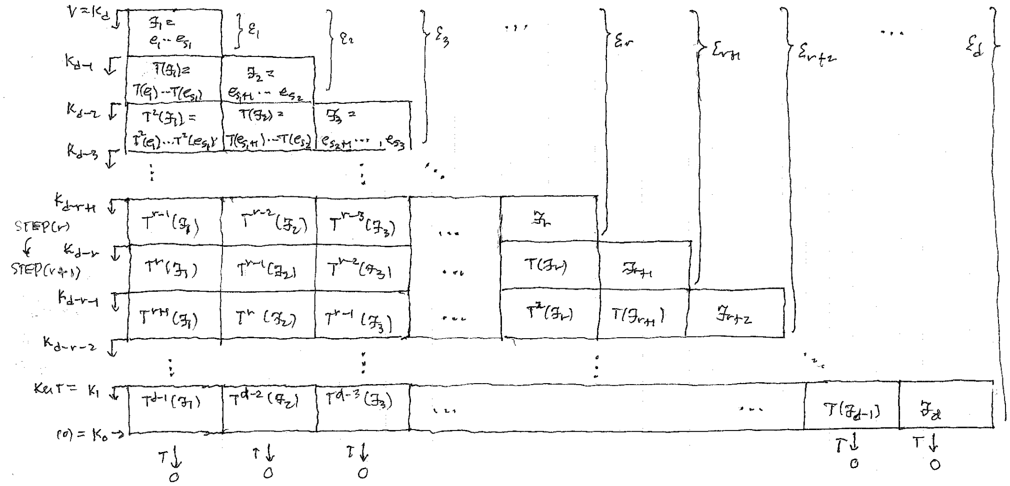

Figure 7.1. Steps in the construction of a basis that decomposes vector space V into cyclic subspaces under the action of a nilpotent linear operator $T : V \to V$. The subspaces K_i are the kernels of the powers T^i for $1 \le i \le d = \deg(T)$, with $K_d = V$ and $K_0 = (0)$.

3.7. Exercise. Let $T : V \to V$ be an arbitrary linear map and W a T-invariant subspace. Independence of vectors $f_1, \cdots, f_r$ mod a T-invariant subspace $W \subseteq V$ implies the independence (mod W') for any smaller T-invariant subspace $W' \subseteq W \subseteq V$. $\square$

Proof of Theorem 3.2. Below we will construct two related sets of vectors $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \cdots$ and $\mathcal{E}_1 = \mathcal{F}_1 \subseteq \mathcal{E}_2 \subseteq \mathcal{E}_3 \subseteq \cdots \subseteq \mathcal{E}_r$ such that $\mathcal{E}_r$ is a basis for V aligned with the kernels $K_d = V \supseteq K_{d-1} \supseteq \cdots \supseteq K_1 = \ker(T) \supseteq \{0\}$. When the construction terminates, the vectors in $\mathcal{E}_r$ will be a basis for all of V that provides the desired decomposition into cyclic subspaces.

(Initial) Step1: Let $\mathcal{F}_1 = \mathcal{E}_1 = \{e_i : i \in \text{ index set } I_1\}$ be any set of vectors in $V \sim K_{d-1}$ that are a basis for V (mod K_{d-1}), so their images $\{\overline{e}_i\}$ are a basis in V/K_{d-1}. Obviously the index set I_1 has cardinality $|I_1| = |V/K_{d-1}| = |V| - |K_{d-1}|$, the dimension of the quotient space.

You might feel more comfortable indicating the index sets $I_1, I_2, \cdots$ being constructed here as consecutive blocks of integers, say $I_1 = \{1, 2, \cdots, s_1\}$, $I_2 = \{s_1 + 1, \cdots, s_2\}$ etc, but this notation becomes really cumbersome after the first two steps. And in fact there is no need to explicitly name the indices in each block. From here on you should refer to the chart shown in Figure 7.1, which lists all the players that will emerge in our discussion.

Step 2: The T-images $T(\mathcal{F}_1)$ lie in the layer $T(V \sim K_{d-1}) \subseteq K_{d-1} \sim K_{d-2}$, as noted in (2). In this step we shall verify two assertions.

Claim (i): The vectors in $T(\mathcal{F}_1) = \{T(e_i) : i \in I_1\} \subseteq K_{d-1} \sim K_{d-2}$ are independent (mod K_{d-2}).

If these vectors are not already representatives of a basis for K_{d-1}/K_{d-2} we can adjoin additional vectors $\mathcal{F}_2 = \{e_i : i \in I_2\} \subseteq K_{d-1} \sim K_{d-2}$ chosen so that $T(\mathcal{F}_1) \cup \mathcal{F}_2$ corresponds to a basis for K_{d-1}/K_{d-2}; otherwise we take $\mathcal{F}_2 = \emptyset$.

Claim (ii): The vectors $\mathcal{E}_2 = \mathcal{F}_2 \cup \big[\mathcal{E}_1 \cup T(\mathcal{F}_1)\big] = \mathcal{E}_1 \cup \big[T(\mathcal{F}_1) \cup \mathcal{F}_2\big]$ are a basis for all of V (mod K_{d-2}).

Remarks: In Linear Algebra I we saw that if $W \subseteq V$ and $\{e_1, \cdots, e_r\}$ is a basis for W, we can adjoin successive "outside vectors" $e_{r+1}, \cdots, e_s$ to get a basis for V. (These can even be found by deleting some of the vectors in a pre-ordained basis in V.) Then the images $\{\overline{e}_{r+1}, \cdots, \overline{e}_s\}$ are a basis for the quotient space V/W. That is how we proved the dimension formula $|V| = |W| + |V/W|$ for finite dimensional V.] $\square$

Proof: Claim (i). If $\sum_{i\in I_1} a_i T(e_i) = T(\sum_{i\in I_1} a_i e_i) \equiv 0 \pmod{K_{d-2}}$ then $\sum_{i\in I_1} a_i e_i$ is in K_{d-2} and also lies in the larger space $K_{d-1} \supseteq K_{d-2}$. But by definition vectors in $\mathcal{F}_1 = \{e_i : i \in I_1\}$ are independent (mod K_{d-1}), so we must have $a_i = 0$ for $i \in I_1$, proving independence (mod K_{d-1}) of the vectors in $T(\mathcal{F}_1)$.

Proof: Claim (ii). Suppose there exist coefficients $a_i^{(1)}, a_i^{(2)}, b_i \in \mathbb{F}$ such that

$$\sum_{i\in I_1} a_i^{(1)} e_i + \sum_{i\in I_2} a_i^{(2)} e_i + \sum_{i\in I_1} b_i T(e_i) \equiv 0 \ \ (mod\ K_{d-2}), \tag{3}$$

This sum lies in K_{d-2}, hence also in the larger subspace K_{d-1}, and the last two terms are already in K_{d-1} because $\mathcal{F}_2 \cup T(\mathcal{F}_1) \subseteq K_{d-1} \sim K_{d-2}$. Thus

$$\sum_{i\in I_1} a_i^{(1)} e_i \equiv 0 \ \ (mod\ K_{d-1}) \ ,$$

and since the e_i, $i \in I_1$, are independent (mod K_{d-1}) we must have $a_i^{(1)} = 0$ for all $i \in I_1$. Now the sum (3) reduces to its last two terms, which all lie in K_{d-1}. But by construction, $\mathcal{F}_2 \cup T(\mathcal{F}_1)$ is a basis for K_{d-1} (mod K_{d-2}), which implies $a_i^{(2)} = 0$ for $i \in I_2$ and $b_i = 0$ for $i \in I_1$. Thus $\mathcal{E}_2 = \mathcal{F}_1 \cup [T(\mathcal{F}_1) \cup \mathcal{F}_2]$ is an independent set of vectors (mod K_{d-2}).

It remains to show $\mathcal{E}_2$ spans V (mod K_{d-2}). If $v \in V$ is not contained in K_{d-1} there is some $v_1 \in \mathbb{F}\text{-span}\{\mathcal{F}_1\}$ such that $v - v_1 \equiv 0 \pmod{K_{d-1}}$, so $v - v_1 \in K_{d-1}$. If this difference is lies outside of K_{d-2} we can find some $v_2 \in T(\mathcal{F}_1) \cup \mathcal{F}_2$ such that $v = (v_1 + v_2) \in K_{d-2}$. Thus $v = v_1 + v_2 \pmod{K_{d-2}}$, and since $v_1 + v_2 \in \mathbb{F}\text{-span}\{\mathcal{F}_1 \cup T(\mathcal{F}_1) \cup \mathcal{F}_2\}$, statement (ii) is proved. $\square$

That completes Step 2. Further inductive steps fill in successive rows in Figure 7.1. They involve no new ideas, but things can get out of hand unless the notation is carefully managed. Below we include a complete discussion of the general inductive step in this process, which could be skipped on first reading. It is followed by a final paragraph proving uniqueness of the multiplicities m_i (which you should read).

The General Inductive Step in Proving Theorem 3.2. This should probably be read with the chart from Figure 7.1 in hand to keep track of the players.

Continuing the recursive construction of basis vectors: at step r we have defined sets of vectors $\mathcal{F}_i \subseteq K_{d-i+1} \sim K_{d-i}$ for $1 \leq i \leq r$ with the properties $\mathcal{E}_1 = \mathcal{F}_1$ and

$$\mathcal{E}_r = \mathcal{E}_{r-1} \cup \left[T^{r-1}(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_{r-1}) \cup \mathcal{F}_r\right]$$

is a basis for V/K_{d-r}. At the next step we take the new vectors

$$T^{r-1}(\mathcal{F}_1) \cup T^{r-2}(\mathcal{F}_2) \cup \cdots \cup \mathcal{F}_r \subseteq K_{d-r+1} \sim K_{d-r}$$

created in the previous step and form their T-images

$$T^r(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_r) \subseteq K_{d-r} \sim K_{d-r-1}$$

To complete the inductive step we show:

1. These vectors are independent (mod K_{d-r-1})

2. We then adjoin additional vectors $\mathcal{F}_{r+1} \subseteq K_{d-r} \sim K_{d-r-1}$ as needed to produce a basis for K_{d-r}/K_{d-r-1}, taking $\mathcal{F}_{r+1} = \emptyset$ if the vectors $T^r(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_r)$ are already representatives for a basis in K_{d-r}/K_{d-r-1}. The vectors

$$\mathcal{E}_{r+1} = \mathcal{E}_r \cup \big[T^r(\mathcal{F}_1) \cup \ldots \cup T(\mathcal{F}_r) \cup \mathcal{F}_{r+1}\big]$$

will then be a basis for V (mod K_{d-r-1}).

Proof details:

1. If the vectors $T^r(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_r)$ are not representatives for an independent set of vectors in K_{d-r}/K_{d-r-1} there are coefficients $\{c_i^{(1)} : i \in I_1\}, \cdots, \{c_i^{(r)} : i \in I_r\}$ such that

$$\sum_{i \in I_1} c_i^{(1)} T^r(e_i) + \ldots + \sum_{i \in I_r} c_i^{(r)} T(e_i) \equiv 0 \quad (mod\ K_{d-r-1})$$

So, this sum is in K_{d-r-1} and in K_{d-r}. But $T^{r-1}\{e_i : i \in I_1\} \cup \cdots \cup \{e_i : i \in I_r\}$ are independent vectors (mod K_{d-r}) by hypothesis, and are a basis for K_{d-r+1}/K_{d-r}. We may rewrite the last congruence as

$$T\Big[\sum_{i \in I_1} c_i^{(1)} T^{r-1}(e_i) + \ldots + \sum_{i \in I_r} c_i^{(r)} e_i\Big] \equiv 0 \quad (\text{mod } K_{d-r-1})$$

So, $T[\cdots] \in K_{d-r-1}$, hence $[\cdots] \in K_{d-r}$ too. By independence of the e_i (mod K_{d-r}), we must have $c_i^{(j)} = 0$ in $\mathbb{F}$ for all i, j. Thus the vectors $T^r(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_r)$ are independent (mod K_{d-r-1}) as claimed.

2. To verify independence of the updated set of vectors

$$\mathcal{E}_{r+1} = \mathcal{E}_r \cup \big[T^r(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_r) \cup \mathcal{F}_{r+1}\big]$$

in V/K_{d-r-1}, suppose some linear combination $S = S' + S''$ is zero (mod K_{d-r-1}) where S' is a sum over vectors in $\mathcal{E}_r$ and S'' a sum over vectors in $T^r(\mathcal{F}_1) \cup \cdots \cup \mathcal{F}_{r+1}$. Then $S \equiv 0$ (mod K_{d-r-1}) implies $S \equiv 0$ (mod K_{d-r}), and then by independence of vectors in $\mathcal{E}_r$ (mod K_{d-r}), all coefficients in S' are zero. The remaining term S'' in the reduced sum lies in $K_{d-r} \sim K_{d-r-1}$, and by independence of $T^r(\mathcal{F}_1) \cup \cdots \cup \mathcal{F}_{r+1}$ in K_{d-r}/K_{d-r-1} all coefficients in S'' are also zero. Thus $\mathcal{E}_{r+1} \subseteq V$ corresponds to an independent set in K_{d-r}/K_{d-r-1}.

Dimension counting reveals that

$$\begin{aligned} |V/K_{d-1}| &= |\mathcal{F}_1| \\ |K_{d-1}/K_{d-2}| &= |T(\mathcal{F}_1)| + |\mathcal{F}_2| \ = |\mathcal{F}_1| + |\mathcal{F}_2| \\ &\ \vdots \\ |K_{d-r}/K_{d-r-1}| &= |\mathcal{F}_1| + \ldots + |\mathcal{F}_{r+1}| \end{aligned} \tag{4}$$

Thus $|V/K_{d-r-1}| = |V/K_{d-1}| + \cdots + |K_{d-r}/K_{d-r-1}|$ is precisely the number $|\mathcal{E}_{r+1}|$ of basis vectors appearing in the first $r+1$ rows from the top of the chart in Figure 7.1). But this is also equal to $\dim(V/V_{d-r-1})$, so $\mathcal{E}_{r+1}$ is a basis for V/V_{d-r-1} and Step(r+1) of the induction is complete.

The Cyclic Subspace Decomposition. A direct sum decomposition of V into cyclic subspaces can now be read out of Figure 7.1, in which basis vectors have been constructed row-by-row. Consider what happens when we partition into columns. For each $e_i \in \mathcal{F}_1$, $(i \in I_1)$, we have $e_i, T(e_i), T^2(e_i), \cdots, T^{d-1}(e_i) \neq 0$ and $T^d(e_i) = 0$, so these vectors span a cyclic subspace $E(e_i)$ such that $T|_{E(e_i)}$ has nilpotent degree d with e_i as its cyclic vector. Since the vectors that span $E(e_i)$ are part of a basis $\mathcal{E}_d$ for all of V, we obtain a direct sum of cyclic T-invariant subspaces $\bigoplus_{i\in I_1} E(e_i) \subseteq V$ ($|I_1| = |\mathcal{F}_1|$ subspaces).

Vectors $e_i \in \mathcal{F}_2$ $(i \in I_2)$ generate cyclic subspaces $E(e_i)$ such that $\dim\big(E(e_i)\big) = \deg(T|_{E(e_i)}) = d-1$; these become part of

$$\bigoplus_{i\in I_1} E(e_i) \oplus \bigoplus_{i_2\in I_2} E(e_2) \ ,$$

etc. At the last step, the vectors $e_i \in \mathcal{F}_d$ $(i \in I_d)$ determine T-invariant one-dimensional cyclic spaces such that $T(\mathbb{F}e_i) = (0)$, with nilpotence degree $= 1$ – i.e. the spaces $E(e_i) = \mathbb{F}e_i$ all lie within $\ker(T)$. The end result is a cyclic subspace decomposition

$$(5) \qquad \Big(\bigoplus_{i_1\in I_1} E(e_{i_1})\Big) \oplus \Big(\bigoplus_{i_2\in I_2} E(e_{i_2})\Big) \oplus \ldots \oplus \Big(\bigoplus_{i_d\in I_d} E(e_{i_d})\Big)$$

of the entire space V, since all basis vectors in $\mathcal{E}_r$ are accounted for. (Various summands in (5) may of course be trivial.)

Uniqueness: A direct sum decomposition $V = \bigoplus_{j=1}^{s} E_j$ into T-invariant cyclic subspaces can be refined by gathering together those E_i of the same dimension, writing

$$V = \bigoplus_{k=1}^{d} \mathcal{H}_k \quad \text{where} \quad \mathcal{H}_k = \bigoplus \{E_i : \dim(E_i) = \deg(T|E_i) = k\}$$

for $1 \leq k \leq d = \deg(T)$.

3.8. Proposition. *In any direct sum decomposition $V = \bigoplus_{j=1}^{s} E_j$ into cyclic T-invariant subspaces, the number of spaces of dimension $\dim(E_i) = k$, $1 \leq k \leq d = \deg(T)$ can be computed in terms of the dimensions of the quotients K_i/K_{i-1}. These numbers are the same for all cyclic decompositions.*

Proof: Let us regard Figure 7.1 as a $d \times d$ array of "cells" with C_{ij} the cell in Row(i) (from the top) and Col(j) (from the left) in the array; the "size" $|C_{ij}|$ of a cell is the number of basis vectors it contains. Note that

- *(i)* $|C_{ij}| = 0$ if the cell lies above the diagonal, with $j > i$, because those cells are empty (others may be empty too).
- *(ii)* $|C_{ij}| = |\mathcal{F}_j|$ for all cells on and below the diagonal in Col(j) of the array. In particular $|C_{j1}| = |\mathcal{F}_1|$ for all nonempty cells in Col(1), $|C_{j2}| = |\mathcal{F}_2|$ for those in Col(2), etc.

By our construction, it is evident that vectors in the nonempty cells in Row(r) of Figure 7.1 correspond to a basis for the quotient space K_{d-r}/K_{d-r-1}. Counting the total number of basis vectors in Row(r) we find that

$$\dim(K_{d-r}/K_{d-r-1}) = |C_{r1}| + \ldots + |C_{r+1,r+1}| = |\mathcal{F}_1| + \ldots + |\mathcal{F}_{r+1}| \ ,$$

We may now recursively compute the values of $|C_{rj}|$ and $|\mathcal{F}_j|$ from the dimensions of the quotent spaces K_i/K_{i-1}. But as noted above, each $e_i \in \mathcal{F}_k$ lies in the diagonal cell C_{kk} and generates a distinct cyclic space in the decomposition. □

That completes the proof of Theorem 3.2.

Remarks. To summarize,

1. We define $K_i = \ker(T^i)$ for $1 \leq d = i$ nilpotence degree of T.
2. The following relations hold.

$$\begin{array}{rcl} \mathcal{E}_1 & = & \mathcal{F}_1 \subseteq V \sim K_{d-1} \text{ determines a basis for } V/K_{d-1}, \\ \mathcal{E}_2 & = & \mathcal{E}_1 \cup \big[\, T(\mathcal{F}_1) \cup \mathcal{F}_2 \,\big] \subseteq V \sim K_{d-2} \text{ determines a basis for } V/K_{d-2}, \\ & \vdots & \\ \mathcal{E}_{r+1} & = & \mathcal{E}_r \cup \big[\, T^r(\mathcal{F}_1) \cup T^{r-1}(\mathcal{F}_2) \cup \cdots \cup \mathcal{F}_{r+1} \,\big] \subseteq V \sim K_{d-r} \text{ determines a basis for } V/K_{d-r-1} \\ & \vdots & \\ \mathcal{E}_d & = & \mathcal{E}_{d-1} \cup \big[\, T^{d-1}(\mathcal{F}_1) \cup \cdots \cup T(\mathcal{F}_{d-1}) \cup \mathcal{F}_d \,\big] \text{ is a basis for all of } V. \ \square \end{array}$$

In working examples it usually helps to start by determining a basis $\mathcal{B}^{(0)} = \mathcal{B}^{(1)} \cup \ldots \cup \mathcal{B}^{(d)}$ for V aligned with the kernels so that $\mathcal{B}^{(1)}$ is a basis for K_1, $\mathcal{B}^{(2)}$ determines a basis for K_2/K_1, etc. This yields a convenient basis in V to start the construction.

3.9. Example. Let $V = \mathbb{F}^5$ and $T : V \to V$ the operator $T = L_A$,

$$T(x_1, \cdots, x_5) = (0, x_3 + x_4, 0, x_3, x_1 + x_4)$$

whose matrix with respect to the standard basis $\mathfrak{X} = \{e_1, \cdots, e_5\}$ in $\mathbb{F}^5$ is

$$A = [T]_{\mathfrak{X}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Show that T is nilpotent, then determine $\deg(T)$ and the kernels

$$V = K_d \supseteq K_{d-1} \supseteq \cdots \supseteq K_1 \supseteq \{0\}$$

Find a basis $\mathfrak{Y}$ such that $[T]_{\mathfrak{Y}}$ has block diagonal form, with each block B_i an elementary nilpotent matrix. This is the *Jordan canonical form* for a nilpotent linear operator.

Discussion: First find bases for the kernels $K_i = \ker(T^i)$. We have

$$\begin{array}{rcl} K_1 = \ker(T) & = & \{\mathbf{x} : x_3 + x_4 = 0,\ x_3 = 0,\ x_1 + x_4 = 0\} \\ & = & \{\mathbf{x} : x_4 = x_3 = 0, x_1 + x_4 = 0\} \ = \ \{x : x_1 = x_3 = x_4 = 0\} \\ & = & \{(0, x_2, 0, 0, x_5) : x_2, x_5 \in \mathbb{F}\} = \mathbb{F}\text{-span}\{e_2, e_5\} \end{array}$$

Iteration of T yields

$$\begin{array}{rcl} T(\mathbf{x}) & = & (0, x_3 + x_4, 0, x_3, x_1 + x_4) \\ T^2(\mathbf{x}) & = & T(T(\mathbf{x})) \ = \ (0, x_3, 0, 0, x_3) \\ T^3(\mathbf{x}) & = & (0, \cdots, 0) \end{array}$$

for $\mathbf{x} \in \mathbb{F}^5$. Clearly T is nilpotent with $\deg(T) = 3$, and

$$\begin{array}{ll} |K_1| = 2 : & K_1 = \mathbb{F}\text{-span}\{e_2, e_5\} = \{\mathbf{x} : x_1 = x_3 = x_4 = 0\} \\ |K_2| = 4 : & K_2 = \ker(T^2) = \{\mathbf{x} : x_3 = 0\} = \mathbb{F}\text{-span}\{e_1, e_2, e_4, e_5\} \\ |K_3| = 5 : & K_3 = \mathbb{F}^5 \end{array}$$

In this example, $\mathfrak{X} = \{e_2, e_5; e_1, e_4; e_2\} = \mathcal{B}^{(1)} \cup \mathcal{B}^{(2)} \cup \mathcal{B}^{(3)}$ is an ordered basis for V aligned with the K_i running through $(0) \subseteq K_1 \subseteq K_2 \subseteq K_3 = V$. From this we can determine the families $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ of Theorem 3.2.

STEP 1: Since $|K_3/K_2| = 1$ any nonzero vector in the layer $K_3 \sim K_2 = \{\mathbf{x} : x_3 \neq 0\}$ yields a basis vector for K_3/K_2. We shall take $\mathcal{F}_1 = \{e_3\}$ chosen from the standard basis $\mathfrak{X}$, and then $\mathcal{E}_1 = \{e_3\}$ too. (Any $\mathbf{x}$ with $x_3 \neq 0$ would also work.)

STEP 2: The image set $T(\mathcal{F}_1) = T(e_3) = e_2 + e_4$ lies in the next layer

$$\begin{aligned} K_2 \sim K_1 &= \{\mathbf{x} : x_3 = 0\} \sim \mathbb{F}\text{-span}\{e_2, e_5\} \\ &= \{\mathbf{x} : x_3 = 0 \text{ and } x_1, x_4 \text{ are not both } = 0\} \end{aligned}$$

Since $|T(\mathcal{F}_1)| = 1$ and $\dim(K_2/K_1) = |K_2| - |K_1| = 4 - 2 = 2$, we must adjoin one suitably chosen new vector $\mathbf{x}$ from layer $K_2 \sim K_1$ to $T(\mathcal{F}_1)$ to get the desired basis for K_2/K_1. Then $\mathcal{F}_2 = \{\mathbf{x}\}$ and

$$\mathcal{E}_2 = (\mathcal{F}_1 \cup T(\mathcal{F}_1)) \cup \mathcal{F}_2 = \{e_3, e_2 + e_4, \mathbf{x}\}$$

$\mathcal{E}_2$ is a basis for V/K_2 as in first inductive step of Theorem 3.2.

A suitable vector $\mathbf{x} = (x_1, \ldots, x_5)$ in $K_2 \sim K_1$, $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$ must have $x_3 = 0$ (so $\mathbf{x} \in K_2$) and x_1, x_3, x_4 not all zero (so $x \notin K_1$). This holds if and only if $(x_3 = 0)$ and (x_1, x_4 are not both 0). But we must also insure that our choice of $\mathbf{x}$ makes $\{e_3, e_2 + e_4, \mathbf{x}\}$ independent (mod K_1). The following lemma is helpful.

3.10. Lemma. *Let $V = \mathbb{F}^n$, W a subspace, $\mathfrak{X} = \{v_1, \cdots, v_r\}$ vectors in V, and let $M = \mathbb{F}$-span$\{v_1, \cdots, v_r\}$ (so $r = |V/W|$). Let $\mathfrak{Y} = \{w_1, \cdots, w_{n-r}\}$ be a basis for W. Then the following assertion are equivalent.*

1. *$\mathfrak{X}$ determines a basis for V/W.*
2. *$\mathfrak{Y} \cup \mathfrak{X} = \{v_1, \cdots, v_r, w_1, \cdots, w_{n-r}\}$ is a basis for V.*
3. *$V = W \oplus M$ (direct sum of subspaces).*

Proof: In Linear Algebra I we showed that the images $\overline{v}_1, \cdots, \overline{v}_r$ are a basis for V/W if and only if $\{v_1, \cdots .v_r\} \cup \mathfrak{Y}$ are a basis for V. It is obvious that $(ii) \Leftrightarrow (iii)$. $\square$

3.11. Corollary. *In the setting of the lemma the "outside vectors" $v_1, \cdots, v_r \in V \sim W$ are a basis for V ((mod W), so the images $\{\overline{v}_1, \cdots, \overline{v}_r\}$ are a basis for V/W, if and only if the $n \times n$ matrix A whose rows are $R_1 = v_1, \cdots, R_r = v_r, R_{r+1} = w_1, \cdots, R_n = w_{n-r}$ has rank equal to n.*

Armed of this observation (and the known basis $\{e_2, e_5\}$ for K_1), we seek a vector $\mathbf{x} = (x_1, \ldots, x_5)$ with x_1, x_4 not all equal to 0, such that

$$A = \begin{pmatrix} e_3 \\ e_2 + e_4 \\ (x_1, x_2, 0, x_4, x_5) \\ e_2 \\ e_5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ x_1 & x_2 & 0 & x_4 & x_5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

has Rowrank(A) = 5. Symbolic row operations put this into the form

$$\begin{pmatrix} x_1 & x_2 & 0 & x_4 & x_5 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

which has $rank = 5$ if and only if $x_1 \neq 0$.

Thus we may take e_1 as the additional vector we seek, and then

$$\mathcal{F}_1 = \{e_3\} \quad T(\mathcal{F}_1) = \{e_2 + e_4\} \quad \mathcal{F}_3 = \{e_1\} \ ,$$

and $\mathcal{E}_2 = \big[\mathcal{F}_1 \cup T(\mathcal{F}_1)\big] \cup \mathcal{F}_2$. That completes Step 2. (Actually any $\mathbf{x}$ with $x_1 \neq 0$, $x_3 = 0$ would work.)

Step 3: In the next layer $K_1 \sim K_0$ we have the vectors

$$T^2(\mathcal{F}_1) = \{T^2(e_3) = T(e_2 + e_4) = e_2 + e_5\} \quad \text{and} \quad T(\mathcal{F}_2) = \{T(e_1)\} = \{e_5\}$$

Since, $|K_1/K_0| = |K_1| = 2$ there is no need to adjoin additional vectors from this layer, so $\mathcal{F}_3 = \emptyset$. The desired basis in V is

$$\mathcal{E}_3 = \mathcal{F}_1 \cup \big[\, T(\mathcal{F}_1) \cup \mathcal{F}_2 \,\big] \cup \big[\, T^2(\mathcal{F}_1) \cup T(\mathcal{F}_2) \,\big] = \{e_3; e_2 + e_4, e_1; e_2 + e_5, e_5\}$$

The iterated action of T sends

$$e_3 \rightarrow T(e_3) = e_2 + e_4 \rightarrow T^2(e_3) = e_2 + e_5 \quad \text{and} \quad e_1 \rightarrow T(e_1) = e_5$$

The cyclic subspaces are

$$\begin{aligned} E_1 &= \mathbb{F}\text{ -span}\{e_3, T(e_3), T^2(e_3)\} \ = \{e_3, e_2 + e_4, e_2 + e_5\} \\ E_2 &= \mathbb{F}\text{-span}\{e_1, T(e_1) = e_5\} \end{aligned}$$

and $V = E_1 \oplus E_2$. With respect to this basis $[T]_{\mathfrak{X}}$ has the block diagonal form

$$[T_{\mathfrak{X}}] = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \ ,$$

each diagonal block being an elementary nilpotent matrix. The number and size of such blocks are uniquely determined but the bases are not unique, nor are the cyclic subspaces in the splitting $V = E_1 \oplus E_2$. $\square$

3.12. Exercise. Let W be the 3-dimensional subspace in $V = \mathbb{F}^5$ determined by the equations

$$\left\{ \begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 3x_1 + 5x_3 &= 0 \end{aligned} \right.$$

which is equivalent to the matrix equation $A\mathbf{x} = 0$ with

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 3 & 0 & 5 & -1 \end{pmatrix}$$

(a) Find vectors $\{v_1, v_2, v_3\}$ that are a basis for W.

(b) Find 2 vectors $\{v_4, v_5\}$ that form a basis for $V \pmod{W}$.

(c) Find two of the standard basis vectors $\{e_1, e_2, e_3, e_4, e_5\}$ in $\mathbb{F}^5$ that are a basis for $V \pmod{W}$.

3.13. Exercise. Do either of the vectors in

$$f_1 = 2e_1 - 3e_2 + e_3 + e_4 \qquad f_2 = -e_1 + 2e_2 + 5e_3 - 2e_4$$

in $\mathbb{F}^5$ lie in the subspace W determined by the system of the previous exercises? Do these vectors form a basis for $\mathbb{F}^5$ (mod W)? $\square$

3.14. Exercise. Which of the following matrices A are nilpotent?

$$\text{(a)} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 2 & -1 \\ -1 & -2 & 1 \\ -1 & -2 & 1 \end{pmatrix} \quad \text{(d)} \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & 4 \end{pmatrix}$$

If A is nilpotent, find a basis for $\mathbb{F}^3$ that puts A into block diagonal form with elementary nilpotent blocks. What is the resulting block diagonal form if the blocks are listed in order of decreasing size? $\square$

3.15. Exercise. If N_1, N_2 are nilpotent is N_1N_2 nilpotent? What if N_1 and N_2 commute? $\square$

3.16. Corollary. If N_1, N_2 are nilpotent operators $N_k : V \to V$ and their commutator $[N_1, N_2] = N_1N_2 - N_2N_1$ is $= 0$.

(a) Prove that linear combination $c_1N_1 + c_2N_2$ are also nilpotent.

(b) If $N_1, \cdots, N_r$ are nilpotent and commute pairwise, so $[N_i, N_j] = 0$ for $i \neq j$, prove that all operators in $\mathbb{F}$-span$\{N_1, \cdots, N_r\}$ are nilpotent. $\square$

3.17. Exercise. Let $V = \mathcal{P}_n(\mathbb{F})$ be the space of polynomials $f = \sum_{i=0}^{n} c_i x^i \in \mathbb{F}[x]$ of degree $\leq n$.

(a) Show that the differentiation operator

$$D : V \to V, Df = df/dx = c_1 + 2c_2\, x + \cdots + n \cdot c_n x^{n-1}$$

is nilpotent with $\deg(D) = n + 1$ (Note: $\dim(V) = n + 1$).

(b) Prove that any constant coefficient differential operator $L : V \to V$ of the form $a_1D + a_2D^2 + \cdots + a_nD^n$ (*no constant term* a_0I) is nilpotent on V.

(c) Does this remain true if a nonzero constant term $c_0\mathbb{1}$ is allowed? $\square$

3.18. Exercise. In the space of polynomials $\mathcal{P}_n(\mathbb{R})$ consider the subspaces

$$\begin{aligned} V_1 &= \{f : f(x) = f(-x), \text{ the even polynomials}\} \\ V_2 &= \{f : f(-x) = -f(x), \text{ the odd polynomials}\} \end{aligned}$$

Prove that these subspaces are invariant under differentiation, and that $\mathcal{P}_n$ is their direct sum $V_1 \oplus V_2$. $\square$

3.19. Exercise. Show $\mathrm{Tr}(A) = 0$, for any nilpotent linear operator $A : V \to V$ of a finite dimensional space. Is the converse true? $\square$

VII.4 A Review of the Diagonalization Problem.

We will give a general structure theorem for linear operators T over a field $\mathbb{F}$ large enough that the characteristic polynomials $p_T = \det(T - xI)$ splits into linear factors $f(x) = c \cdot \prod_{i=1}^{s}(x - a_i)^{m_i}$ in $\mathbb{F}[x]$. This is always true if $\mathbb{F} = \mathbb{C}$, but p_T need not split over other fields; and even if $p_T(x)$ does split, that alone is not enough to guarantee T is diagonalizable. In this section we briefly review diagonalizability of linear operators over a general field $\mathbb{F}$, which means that there is a basis of eigenvectors in V (or equivalently that the eigenspaces $E_\lambda(T)$ span V so $V = \sum_\lambda E_\lambda(T)$). If you already have a good understanding of these matters you may want to skip to Section VII.5 where we discuss the *generalized eigenspaces* that lead to the Jordan Decomposition. However, you should at least read the next theorem and its proof since the techniques used are the basis for the more complicated proof that generalized eigenspaces are independent, part of a direct sum decomposition of V.

Diagonalization.

4.1. Definition. *Let $T : V \to V$ be a linear operator on vector space over $\mathbb{F}$. If $\lambda \in \mathbb{F}$, the* **λ-eigenspace** *is $E_\lambda = \{v \in V : (T - \lambda I)v = 0\}$. Then λ is an* **eigenvalue** *if $E_\lambda(T) \neq \{0\}$ and $\dim_{\mathbb{F}}\big(E_\lambda(T)\big)$ is its* **geometric multiplicity**. *We often refer to $\mathrm{sp}_{\mathbb{F}}(T) = \{\lambda \in \mathbb{F} : E_\lambda \neq \{0\}\}$ as the* **spectrum** *of T over $\mathbb{F}$.*

4.2. Exercise. Show that every eigenspace E_λ is a vector subspace in V that is T-invariant. If $\mathfrak{X} = \{e_1, \cdots, e_r, \cdots, e_n\}$ is a basis for V that first passes through E_λ, show that the matrix of T takes the form

$$[T]_{\mathfrak{X}} = \left(\begin{array}{ccc|cc} \lambda & 0 & 0 & * & * \\ \cdot & \cdot & \cdot & * & * \\ 0 & 0 & \lambda & * & * \\ \hline 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{array}\right)$$

The geometric multiplicity of λ is $\dim(E_\lambda(T))$. We have already seen that when $\mathbb{F} = \mathbb{R}$ the operator $T = (90^\circ$ rotation acting on $\mathbb{R}^2)$ has no eigenvalues in $\mathbb{R}$, so $\mathrm{sp}_{\mathbb{R}}(T) = \emptyset$.

An operator T is diagonalizable if there is a basis $\mathfrak{X} = \{e_1, \cdots, e_n\}$ consisting of eigenvectors e_i, so $T(e_i) = \lambda_i e_i$ with respect to this basis. Then $[T]_{\mathfrak{X}}$ has the diagonal form

$$[T]_{\mathfrak{X}} = \begin{pmatrix} \lambda_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & \lambda_n \end{pmatrix}$$

in which there may be repeats among the λ_i. Conversely, any basis such that $[T]_{\mathfrak{X}}$ takes this form consist entirely of eigenvectors for T. A more sophisticated choice of basis vectors puts $[T]_{\mathfrak{X}}$ into *block diagonal form*. First a simple observation:

4.3. Exercise. If $T : V \to V$ is a linear operator on a finite dimensional space, show that the following statements are equivalent.

(a) There is a basis in V consisting of eigenvectors.

(b) The eigenspaces for T span V, so that

$$V = \sum_{\lambda \in \mathrm{sp}(T)} E_\lambda(T)\ .$$

Note: There actually is something to be proved here: (b) requires more care selecting basis vectors than (a). $\square$

So, if T is diagonalizable and $\{\lambda_1, \ldots, \lambda_r\}$ are its *distinct* eigenvalues in $\mathbb{F}$, may choose a basis of eigenvectors e_i that first runs through E_{λ_1}, then through E_{λ_2}, etc. It is obvious that this choice yields a "block disagonal" matrix

$$[T]_{\mathfrak{X}} = \begin{pmatrix} \lambda_1 \boxed{I_{m_1 \times m_1}} & & & 0 \\ 0 & \lambda_2 \boxed{I_{2 \times m_2}} & & \\ & & \ddots & \\ 0 & & & \lambda_r \boxed{I_{m_r \times m_r}} \end{pmatrix}$$

in which $m_i = \dim\big(E_{\lambda_i}(T)\big)$. $\square$

These observations do not quite yield the definitive characterization of diagonalizability.

(6) DIAGONALIZABILITY CRITERION. *A linear operator T on a finite dimensional space is diagonalizable over $\mathbb{F} \Leftrightarrow V$ is the* DIRECT SUM *of its distinct eigenspaces: $V = \bigoplus_{i=1}^r E_{\lambda_i}(T)$.*

The implication ($\Leftarrow$) is trivial, but in the reverse direction we have so far only shown that (diagonalizable) $\Rightarrow V$ is *spanned* by its eigenspaces, so $V = \sum_{i=1}^r E_{\lambda_i}(T)$ and every v has at least one decomposition $v = \sum_i v_i : wq$ with $v_i \in E_{\lambda_i}(T)$. In a direct sum $\bigoplus_i E_{\lambda_i}(T)$ the decomposition is *unique*, and in particular $0 = \sum_i v_i$ with $v_i \in E_{\lambda_i}(T) \Rightarrow$ each term $v_i = 0$.

4.4. Exercise. Finish the proof of the Diagonalizability Criterion (6). If $V = \sum_i E_{\lambda_i}(T)$ prove that every $v \in V$ has a UNIQUE decomposition $v = \sum_i v_i$ such that $v_i \in E_{\lambda_i}(T)$. $\square$

4.5. Proposition. If $\{\lambda_1, \cdots, \lambda_r\}$ are the distinct eigenvalues in $\mathbb{F}$ for a linear operator $T : V \to V$ on a finite dimensional vector space, and if the eigenspaces E_{λ_i} span V, then V is a direct sum $E_{\lambda_1} \bigoplus \cdots \bigoplus E_{\lambda_r}$. Furthermore,

1. $\dim(V) = \sum_{i=1}^r \dim(E_{\lambda_i}) = \sum_{i=1}^r$ (geometric multiplicity of λ_i)
2. T is diagonalizable over $\mathbb{F}$.

Proof: Since $V = \sum_i E_{\lambda_i}$ every vector in V has a decomposition $v = \sum_{i=1}^r v_i$ with $V_i \in E_{\lambda_i}(T)$, so we need only prove uniqueness of this decomposition, which in turn reduces to proving that the v_i are "independent" in the sense that

$$0 = v_1 + \cdot + v_r \text{ with } v_i \in E_{\lambda_i} \;\Rightarrow\; v_1 = \cdots = v_r = 0$$

Note that for $\mu, \lambda \in \mathbb{F}$, the linear operators $(T - \lambda I)$, $(T - \mu I)$ commute with each other, since I commutes with everybody. Now suppose $\sum_{i=1}^r v_i = 0$ with $T(v_i) = \lambda_i v_i$.

Fix an index i and apply the operator $S = \prod_{j \neq i}(T - \lambda_j I)$ to the sum. We get

$$0 = S(0) = S\big(\sum_{k=1}^r v_k\big) = \sum_{k=1}^r S(v_k) \tag{7}$$

But if $k \neq i$, we can write

$$S(v_k) = \prod_{\ell \neq i}(T - \lambda_\ell I)v_k = \Big[\prod_{\ell \neq k,i}(T - \lambda_k I)\Big] \cdot (T - \lambda_k)v_k = 0$$

Hence the sum (7) reduces to

$$0 = \sum_k S(v_k) = S(v_i) + 0 + \cdots + 0 = \prod_{\ell \neq i}(T - \lambda_\ell I)v_i$$

Observe that we may write $(T - \lambda_\ell) = i) = (T - \lambda_i) + (\lambda_i - \lambda_\ell)I$, for all ℓ, so this becomes

$$0 = \Big[\prod_{\ell \neq i}(T - \lambda_i) + (\lambda_i - \lambda_\ell)I\Big]\, v_i = 0 + \Big[\prod_{\ell \neq i}(\lambda_i - \lambda_\ell)\Big]\, v_i \tag{8}$$

(because $(T - \lambda_i)v_i = 0$). The constant $c = \prod_{\ell \neq i}(\lambda_i - \lambda_\ell)$ must be nonzero because $\lambda_\ell \neq \lambda_i$. Therefore (7) $\Rightarrow v_i = 0$. This works for every $1 \leq i \leq r$ so the v_i are independent, as required. $\square$

4.6. Exercise. Let V be finite dimensional, $\{\lambda_1, \cdots, \lambda_r\}$ the distinct eigenvalues in $\mathbb{F}$ for an $\mathbb{F}$-linear operator $T : V \to V$. Let $E = \sum_{i=1}^r E_{\lambda_i}$ be the span of the eigenspaces. ($E \subseteq V$). Show that

(a) E is T-invariant .

(b) $T|_E$ is diagonalizable. $\Box$

4.7. Exercise. Let $T : V \to V$ be an linear operator on a finite dimensional vector space over $\mathbb{F}$, with $n = \dim_{\mathbb{F}}(V)$. If T has n distinct eigenvalues, prove that

(a) $V = \bigoplus_{i=1}^n E_{\lambda_i}$,

(b) The eigenspace are all one-dimensional, and

(c) T is diagonalizable. $\Box$

4.8. Exercise. If a basis $\mathfrak{X}$ for V passes through the successive eigenspaces $E_{\lambda_1}(T), \cdots, E_{\lambda_r}(T)$, and we then adjoin vectors outside of the subspace $E = \sum_{\lambda_i \in \mathrm{sp}(T)} E_{\lambda_i}(T)$ to get a basis for V, explain why the matrix of T has the form

$$[T]_{\mathfrak{X}} = \left(\begin{array}{ccc|cc} \boxed{\lambda_1 I_{m_1 \times m_1}} & 0 & 0 & * & * \\ \cdot & \cdot & \cdot & * & * \\ 0 & 0 & \boxed{\lambda_r I_{m_r \times m_r}} & * & * \\ \hline 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{array}\right)$$

where $m_i = \dim(E_{\lambda_i}(T))$. $\Box$

4.9. Definition. *If $T : V \to V$ is linear operator on a finite dimensional vector space, every root $\alpha \in \mathbb{F}$ of the characteristic polynomial $p_T(x) = \det(T - x\,I)$ is an eigenvalue for T, so $p_T(x)$ is divisible (without remainder) by $(x - \alpha)$.Repeated division by $(x - \alpha)$ may be possible, and yields a factorization $p_T(x) = (x - \alpha)^{m_\alpha} Q(x)$ where $Q \in \mathbb{F}[x]$ does not have α as a root, and thus is not divisible by $(x - \alpha)$. The exponent m_α is the* **algebraic multiplicity** *of the eigenvalue α.*

Now suppose $\mathbb{F}$ is an algebraically closed field (every nonconstant polynomial $f \in \mathbb{F}[x]$ has a root $\alpha \in \mathbb{F}$), for example $\mathbb{F} = \mathbb{C}$. It follows that every f over such a field splits completely into linear factors $f = c \cdot \prod_{i=1}(x - \alpha_i)$ where $\alpha_1, \cdots, \alpha_n$ are the roots of $f(x)$ in $\mathbb{F}$ (repeats allowed). If $T : V \to V$ is a linear operator on a finite dimensional vector space over such a field, and $\lambda_1, \ldots, \lambda_r$ are its *distinct* eigenvalues in $\mathbb{F}$, the characteristic polynomial splits completely

$$p_T(x) = \det(T - xI) = c \cdot \prod_{j=1}^{r} (x - \lambda_j)^{m_j}$$

where $m_j =$ the algebraic multiplicity of λ_j and $\sum_j m_j = \dim(V)$.

4.10. Corollary. *Let $T : V \to V$ be a linear operator on a finite dimensional space V over $\mathbb{F} = \mathbb{C}$. If the characteristic polynomial*

$$p_T(x) = \det(T - x\,I) = c \cdot \prod_{j=1}^{r} (x - \lambda_j)^{m_j}$$

has DISTINCT *roots (so $m_j = 1$ for all j), then $r = n = \dim_{\mathbb{C}}(V)$ and T is diagonalizable.*

Algebraic vs Geometric Multiplicity.

4.11. Proposition. *If $\lambda \in \mathbb{F}$ is an eigenvalue for linear operator $T : V \to V$, its algebraic multiplicity as a root of $p_T(x) = \det(T - x\,I)$ is $\geq$ (geometric multiplicity of λ) $=$*

$\dim E_\lambda$.

Proof: Fix an eigenvalue λ. Then $E = E_\lambda(T)$ is T-invariant and $T|_E = \lambda \cdot \mathrm{id}_E$. So, if we take a basis $\{e_1, \cdots, e_m\}$ in E_λ and then add vectors $e_{m+1}, \cdots, e_n$ to get a basis $\mathfrak{X}$ for V, we have

$$[T]_{\mathfrak{X}} = \left(\begin{array}{c|c} \lambda I_{m\times m} & * \\ \hline 0 & * \end{array}\right) \qquad \big(m = \dim(E_\lambda(T))\big)$$

Here m is the geometric multplicity of λ and the characteristic polynomial is

$$p_T(x) = \det \left(\begin{array}{ccc|ccc} (\lambda - x) & & 0 & & & \\ & \cdot & & & * & \\ 0 & & (\lambda - x) & & & \\ \hline & & & (a_{m+1,m+1} - x) & & * \\ & 0 & & & \cdot & \\ & & & * & & (a_{n,n} - x) \end{array}\right)$$

This determinant can be written as

$$\det(T - xI) = \sum_{\pi \in S_n} sgn(\pi) \cdot (T - xI)_{1,\pi(1)} \cdot \ldots \cdot (T - xI)_{n,\pi(n)} \tag{9}$$

Each term in this sum involves a product of matrix entries, one selected from each row. If the spots occupied by the selected entries in (9) are marked with a "$\Box$," the marked spots provide a "template" for making the selection, and there is one template for each permutation $\pi \in S_n$: in Row(i), mark the entry in Col(j) with $j = \pi(i)$.

The only $n \times n$ templates that can contribute to the determinant of our block-upper triangular matrix $(T - xI)$ are those in which the first m diagonal spots have been marked (otherwise the corresponding product of terms will include a zero selected from the lower left block). The remaining marked spots must then be selected from the lower right block $(*)$ – i.e. from Row(i) and Col(j) with $m + 1 \le i, j \le n$, as indicated in the following diagram.

$$\left(\begin{array}{ccc|cccc} \Box & & 0 & & & & \\ & \cdot & & & * & & \\ 0 & & \Box & & & & \\ \hline & & & \cdot & \cdot & \Box & \cdot \\ & 0 & & \Box & \cdot & \cdot & \cdot \\ & & & \cdot & \Box & \cdot & \cdot \end{array}\right)$$

Thus $p_T(x) = \det(T - xI)$ has the general form $(x - \lambda)^m \cdot G(x)$, in which the factor $G(x)$ might involve additional copies of λ. We conclude that

$$\big(\text{algebraic multiplicity of } \lambda\big) \ge m = \big(\text{ geometric multiplicity of } \lambda\big) ,$$

as claimed. $\Box$

4.12. Example. Let $T = L_A : \mathbb{R}^3 \to \mathbb{R}^3$, with $A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{pmatrix}$ If $\mathfrak{X} = \{e_1, e_2, e_3\}$ is the standard Euclidean basis then $[T]_{\mathfrak{X}} = A$ and the characteristic polynomial is

$$\begin{aligned} p_T(x) &= \det(A - x\,I) = \det \begin{pmatrix} 4-x & 0 & 1 \\ 2 & 3-x & 2 \\ 1 & 0 & 4-x \end{pmatrix} \\ &= [(4-x)(3-x)(4-x) + 0 + 0\,] - [(3-x) + 0 + 0\,] \\ &= (12 - 7x + x^2)(4-x) - 3 + x \\ &= 48 - 28x + 4x^2 - 12x + 7x^2 - x^3 - 3 + x \\ &= -x^3 + 11x^2 - 39x + 45 \end{aligned}$$

To determine $\mathrm{sp}(T)$ we need to find roots of a cubic; however we can in this case guess a root λ and then long divide by $(x-\lambda)$. After a little trial and error it turns out that $\lambda = 3$ is a root, with $p_T(3) = -27 + 99 - 117 + 45 = 0$ and $p_T(x) = -(x-3)(x^2 - 8x + 15) = -(x-3)^2(x-5)$.

Eigenvalues in $\mathbb{F} = \mathbb{R}$ (or $\mathbb{F} = \mathbb{Q}$) are $\lambda_1 = 3, \lambda_2 = 5$ with algebraic multiplicities $m_1 = 2, m_2 = 1$. For the geometric multiplicities we must compute the eigenspaces $E_{\lambda_k}(T)$.

Case 1: $\lambda = 3$. We solve the system $(A - 3I)\mathbf{x} = 0$ by row reduction.

$$[A - 3I\,] = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Columns in the row reduced system that *do not* meet a "step corner" $\boxed{\ast}$ correspond to free variables in the solution; thus x_2, x_3 can take any value in $\mathbb{F}$ while $x_1 = -x_3$. Thus

$$\begin{aligned} E_{\lambda=3} &= \ker(A - 3I) \;=\; \{(-v_3, v_2, v_3) : v_2, v_3 \in \mathbb{F}, v_1 = -v_3\} \\ &= \mathbb{F}\cdot(-1, 0, 1) \oplus \mathbb{F}\cdot(0, 1, 0) \;=\; \mathbb{F}(-e_1 + e_3) \oplus \mathbb{F}e_2 \end{aligned}$$

These vectors are a basis and $2 = \dim(E_{\lambda=3}) =$ (geometric multiplicity) $=$ (algebraic multplicity).

Case 2: $\lambda = 5$. Solving $(A - 5I)\mathbf{v} = 0$ by row reduction yields

$$[A{-}5I\,] = \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 2 & -2 & 2 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 0 & -2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} \boxed{1} & 0 & -1 & 0 \\ 0 & \boxed{1} & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Now there is only one free variable x_3, with $x_2 = 2x_3$ and $x_1 = x_3$. Thus

$$E_{\lambda=5} = \{(x_3, 2x_3, x_3) : x_3 \in \mathbb{F}\} \;=\; \mathbb{F}\cdot(1, 2, 1)$$

and

$$1 = \dim(E_{\lambda=5}) = (\text{geometric multiplicity of } \lambda = 5) = (\text{algebraic multiplicity}).$$

Diagonalization: A basis $\mathfrak{Y}$ consisting of eigenvectors is given by

$$f_1 = -e_1 + e_3 \qquad f_2 = e_2 \qquad f_3 = e_1 + 2e_2 + e_3\ .$$

and for this basis we have

$$[T]_{\mathfrak{Y}} = \left(\begin{array}{cc|c} 3 & 0 & 0 \\ 0 & 3 & 0 \\ \hline 0 & 0 & 5 \end{array}\right)$$

while $[T]_{\mathfrak{X}} = A$ with respect to the standard Euclidean basis $\{e_i\}$.

It is sometimes important to know the similarity transform $SAS^{-1} = [T]_{\mathfrak{Y}}$ that effects the transition between bases. The matrix S can be found by writing

$$\begin{aligned} [T]_{\mathfrak{Y}\mathfrak{Y}} &= [\mathrm{id} \circ T \circ \mathrm{id}\,]_{\mathfrak{Y}\mathfrak{Y}} \;=\; [\mathrm{id}\,]_{\mathfrak{Y}\mathfrak{X}} \cdot [T]_{\mathfrak{X}\mathfrak{X}} \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} \\ &= [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} \cdot A \cdot [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} \end{aligned}$$

Then $S = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}$, with $S_{\mathfrak{X}\mathfrak{Y}} = S^{-1}$ because

$$[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} \cdot [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{X}} = I_{3\times 3} \qquad (3 \times 3 \text{ identity matrix})$$

(All this is discussed in Chapter II.4 of the *Linear Algebra I Notes*.)

The easiest matrix to determine is usually $S^{-1} = [\mathrm{id}\,]_{\mathfrak{X}\mathfrak{Y}}$ which can be written down immediately if we know how to write basis vectors in $\mathfrak{Y}$ in terms of those in the standard basis $\mathfrak{X}$ in $\mathbb{F}^3$). In the present example we have

$$\left\{ \begin{array}{lcl} id(f_1) &=& -e_1 + 0 \cdot e_2 + e_3 \\ id(f_2) &=& 0 + e_2 + 0 \\ id(f_3) &=& e_1 + 2e_2 + e_3 \end{array} \right. \qquad \Rightarrow \; S^{-1} = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}} \;=\; \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$

It is useful to note that the matrix $[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}$ is just the transpose of the coefficient array in the system of vector identities that express the f_i in terms of the e_j.

We can now find the desired inverse $S = (S^{-1})^{-1}$ by row operations (or by Cramer's rule) to get

$$S = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ -1 & 1 & -1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

and then

$$SAS^{-1} \;=\; \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

as expected. That concludes our analysis of this example. $\square$

4.13. Exercise. Fill in the details needed to compute S.

4.14. Example. Let $A = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix}$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Then

$$A^2 = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad A^3 = 0\ ,$$

so with respect to the standard basis in $\mathbb{F}^2$ the matrix of the map $T = L_A : \mathbb{F}^2 \to \mathbb{F}^2$ is $[T]_{\mathfrak{X}} = A$ and the characteristic polynomial is:

$$\det(A - x\,I) = \det \begin{pmatrix} 2 - x & 4 \\ -1 & -2 - x \end{pmatrix} \;=\; -4 + x^2 + 4 = x^2$$

Thus, $\lambda = 0$ is a root (over $\mathbb{R}$ or $\mathbb{C}$) with (algbraic multiplicity) $= 2$, but the geometric multiplicity is $\dim\big(E_{\lambda=0}\big) = 1$. When we solve the system $(A - \lambda I)\mathbf{x} = A\mathbf{x} = 0$ taking $\lambda = 0$ to determine $E_{\lambda=0} = \ker(L_A)$, row reduction yields

$$\left(\begin{array}{cc|c} 2-\lambda & 4 & 0 \\ -1 & -2-\lambda & 0 \end{array} \right) = \left(\begin{array}{cc|c} 2 & 4 & 0 \\ -1 & -2 & 0 \end{array} \right) \to \left(\begin{array}{cc|c} \boxed{1} & 2 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

For this system x_2 is a free variable and $x_1 = -2x_2$, so

$$E_{\lambda=0} = \mathbb{F}\cdot(-2,1) = \mathbb{F}\cdot(2e_1 - e_2)\ ,$$

and $\dim\big(E_{\lambda=0}\big) = 1 <$ (algebraic multiplicity) $= 2$. There are no other eigenvalues so the best we can do in trying to reduce T is to find a basis such that $[T]_{\mathfrak{Y}}$ has form

$$\left(\begin{array}{c|c} 0 & * \\ \hline 0 & * \end{array}\right)$$

by taking $\mathfrak{Y} = \{f_1, f_2\}$ where $f_1 = (2,1) = 2e_1{+}e_2$ and f_2 is any other vector independent of f_1.

However, T is a nilpotent operator (verify this), so we can do better with a slightly different basis $\mathfrak{Z}$ that puts A into the Jordan canonical form for nilpotent operators (as in Theorem 3.2). In the present example this is

$$[T]_{\mathfrak{Z}} = \begin{pmatrix} \boxed{0} & 1 \\ 0 & \boxed{0} \end{pmatrix} \qquad \text{(an elementary nilpotent matrix)}$$

with two 1×1 blocks of zeros on the diagonal. In fact, in the notation of Theorem 3.2 we have kernels $(0) \subseteq K_1 = \ker(T) \subseteq K_2 = \ker(T^2) = V$ with

$$K_1 = E_{\lambda=0} = \mathbb{F}\cdot f_1 \qquad \text{and} \qquad K_2 = \mathbb{F}^2$$

So, if f_2 is any vector transverse to $\ker(T)$, we have $T(f_2) \in \ker(T) = \mathbb{F}\cdot f_1$. But $T(f_2) \neq 0$ since $f_2 \notin K_1$, and by scaling f_2 appropriately we can make $T(f_2) = f_1$. Then $\mathfrak{Z} = \{f_1, f_2\}$ is a basis that puts $[T]_{\mathfrak{Z}}$ into the form shown above. $\square$

4.15. Exercise. *Repeat the analysis of the previous exercise for the matrix* $A = \begin{pmatrix} 4 & 4 \\ -1 & 0 \end{pmatrix}$. $\square$

That concludes our review of Diagonalization.

VII-5. Generalized Eigenspace Decomposition I.

The Fitting Decomposition (Proposition 1.5) is a first step in trying to decompose a linear operator $T : V \to V$ over an arbitrary field.

5.1. Proposition (Fitting Decomposition). *Given linear $T : V \to V$ on a finite dimensional vector space over any field, then $V = N \oplus S$ for T-invariant subspaces N, S such that $T|_S : S \to S$ is a bijection (invertible linear operator on S), and $T|_N : N \to N$ is nilpotent.*

The relevant subspaces are the "stable kernel" and "stable range" of T,

$$\begin{aligned} K_\infty &= \bigcup_{i=1}^{\infty} K_i, \quad \big(K_i = \ker(T^i) \quad \text{with } \{0\} \subsetneqq K_1 \subsetneqq \cdots \subsetneqq K_r = K_{r+1} = \cdots = K_\infty\big) \\ R_\infty &= \bigcap_{i=1}^{\infty} R_i, \quad \big(R_i = \text{range}(T^i) \quad \text{with } \{0\} \supsetneqq R_1 \supsetneqq \cdots \supsetneqq R_r = R_{r+1} = \cdots = R_\infty\big) \end{aligned}$$

(see Section VII-1). Obviously, $T = (T|_{R_\infty}) \oplus (T|_{K_\infty})$ which splits T into canonically defined nilpotent and invertible parts.

5.2. Exercise. Prove that the Fitting decomposition is unique: If $V = N \oplus S$, both T-invariant, such that $T|_N$ is nilpotent and $T|_S : S \to S$ invertible show that $N = K_\infty(T)$ and $S = R_\infty(T)$. $\square$

Given a linear operator $T : V \to V$ we may apply these remarks to the operators $(T - \lambda I)$ associated with eigenvalues λ in $\mathrm{sp}_{\mathbb{F}}(T)$. The eigenspace $E_\lambda(T) = \ker(T - \lambda I)$ is the first in an ascending chain of subspaces, shown below.

$$\{0\} \subsetneqq \ker(T-\lambda) = E_\lambda(T) \subsetneqq \ker(T-\lambda)^2 \subsetneqq \cdots \subsetneqq \ker(T-\lambda)^r = \cdots = K_\infty(T-\lambda)$$

5.3. Definition. *If $\lambda \in \mathbb{F}$ the "stable kernel" of $(T - \lambda I)$*

$$K_\infty(\lambda) = \bigcup_{m=1}^{\infty} \ker(T - \lambda I)^m$$

is called the **generalized λ-eigenspace**, *which we shall hereafter denote by $M_\lambda(T)$. Thus,*

$$\begin{aligned} M_\lambda(T) &= \{v \in V : (T-\lambda I)^k v = 0 \text{ for some } k \in \mathbb{N}\} \\ &\supseteq E_\lambda(T) = \{v : (T - \lambda I)v = 0\} \end{aligned} \tag{10}$$

We refer to any $\lambda \in \mathbb{F}$ such that $M_\lambda(T) \neq (0)$ as a **generalized eigenvalue** *for T. But note that $M_\lambda(T) \neq \{0\} \Leftrightarrow E_\lambda(T) \neq \{0\} \Leftrightarrow \det(T - \lambda I) = 0$, so these are just the usual eigenvalues of T in $\mathbb{F}$.*

Generalized eigenspaces have the following properties.

5.4. Lemma. *The spaces $M_\lambda(T)$ are T-invariant.*

Proof: T commutes with all the operators $(T - \lambda)^m$, which commute with each other. Thus, $v \in M_\lambda(T) \Rightarrow (T - \lambda I)^k v = 0$ for some $k \in \mathbb{N} \Rightarrow$

$$(T - \lambda I)^k T(v) = T(T - \lambda I)^k v = T(0) = 0$$

Hence $T(v) \in M_\lambda(T)$. $\square$

We now show that $T|_{M_\lambda}$ has a nice upper triangular form with respect to a suitably chosen basis in M_λ.

5.5. Proposition. *Every generalized eigenspace $M_\lambda(T)$, $\lambda \in \mathrm{sp}(T)$, has a basis $\mathfrak{X}$ such that the matrix of $T|_{M_\lambda(T)}$ has upper triangular form*

$$[T|_{M_\lambda}]_{\mathfrak{X}} = \begin{pmatrix} \lambda & & * \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda \end{pmatrix}$$

Proof: We already know that any nilpotent operator N on a finite dimensional vector space can be put into strictly upper triangular form by a suitable choice of basis.

$$[N]_{\mathfrak{X}} = \begin{pmatrix} 0 & & * \\ & \cdot & \\ & & \cdot \\ 0 & & 0 \end{pmatrix}$$

Now write

$$T|_{M_\lambda} = (T - \lambda I)|_{M_\lambda} + \lambda I|_{M_\lambda}$$

in which $V = M_\lambda$, $N = (T - \lambda I)|_{M_\lambda}$ and $I|_{M_\lambda}$ is the identity operator on M_λ. Since $[I|_{M_\lambda}]_{\mathfrak{X}} = I_{m\times m}$ for any basis, a basis that puts $(T - \lambda I)|_{M_\lambda}$ into strict upper triangular form automatically yields

$$[T|_{M_\lambda}]_{\mathfrak{X}} = \big[(T - \lambda I)|_{M_\lambda}\big]_{\mathfrak{X}} + \lambda I = \begin{pmatrix} \lambda & & * \\ & \cdot & \\ 0 & & \lambda \end{pmatrix} \quad \square$$

The most precise result of this sort is obtained using the cyclic subspace decomposition for nilpotent operators (Theorem 3.2) to guide our choice of basis. As a preliminary step we might pick a basis $\mathfrak{X}$ aligned with the kernels

$$(0) \subsetneqq K_1 = \ker(T) \subsetneqq K_2 = \ker(T^2) \subsetneqq \ldots \subsetneqq K_d = V$$

where $d = \deg(T)$. As we indicated earlier in Exercise 2.4, $[T]_{\mathfrak{X}}$ is then upper triangular with zero blocks $Z_i, 1 \leq i \leq d = \deg(T)$, on the diagonal. Applying this to a generalized eigenspace $M_\lambda(T)$, the matrix of the nilpotent operator $T - \lambda I$ becomes upper triangular with zero blocks on the diagonal. Writing $T = (T - \lambda I) + \lambda I$ as above we see that the matrix of T with respect to any basis $\mathfrak{X}$ running through successive kernels $K_i = \ker(T - \lambda I)^i$ must have the form

$$\begin{aligned} \big[T|_{M_\lambda}\big]_{\mathfrak{X}} &= \lambda \cdot I_{n\times n} + [\,T - \lambda I\,]_{\mathfrak{X}} \\ &= \begin{pmatrix} \lambda \cdot \boxed{I_{m_1\times m_1}} & & & * \\ & \lambda \cdot \boxed{I_{m_2\times m_2}} & & \\ & & \ddots & \\ 0 & & & \lambda \cdot \boxed{I_{m_r\times m_r}} \end{pmatrix} \end{aligned} \tag{11}$$

with $m_i = \dim(K_i/K_{i-1}) = \dim(K_i) - \dim(K_{i-1})$ and $n = \dim(V) = \sum_i m_i$. The shape of the "block upper triangular form" (11) is completely determined by the dimensions of the kernels $K_i = K_i(T - \lambda I)$.

Note that (11) can be viewed as saying $T|_{M_\lambda} = \lambda I_\lambda + N_\lambda$ where $I_\lambda = \mathrm{id}_{M_\lambda}$, λI_λ is a scalar operator on M_λ, and $N_\lambda = (T - \lambda I)|_{M_\lambda}$ is a nilpotent operator whose matrix with respect to the basis $\mathfrak{X}$ is similar to the matrix in (11), but with $m_i \times m_i$ zero-blocks on the diagonal. The restriction $T|_{M_\lambda}$ has an "additive decomposition" $T|_{M_\lambda} = $ (*diagonal*) + (*nilpotent*) into commuting scalar and nilpotent parts,

$$T|_{M_\lambda} = \lambda \cdot I + N_\lambda = \begin{pmatrix} \lambda & & * \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda \end{pmatrix} + \begin{pmatrix} 0 & & * \\ & \cdot & \\ & & \cdot \\ 0 & & 0 \end{pmatrix}$$

Furthermore, the nilpotent part N_λ turns out to be a polynomial function of $\big(T|_{M_\lambda}\big)$, so both components of this decomposition also commute with $T|M_\lambda$. There is also a "multiplicative decomposition" $T|_{M_\lambda} = $ (*diagonal*) $\cdot$ (*unipotent*) $= (\lambda I) \cdot U_\lambda$ where U_λ is the unipotent operator $(I + N_\lambda)$; for the corresponding matrices we have

$$\begin{pmatrix} \lambda & & * \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda \end{pmatrix} \cdot \begin{pmatrix} 1 & & * \\ & \cdot & \\ & & \cdot \\ 0 & & 1 \end{pmatrix}$$

Note: The off-diagonal entries $(*)$ in N_λ and U_λ need not be the same in these two decompositions.

As we show below, this description of the way T acts on M_λ can be refined to provide much more information about the off-diagonal terms $(*)$, but we will also see that for many purposes the less explicit block upper triangular form (11) will suffice, and is easy to compute since we only need to determine the kernels K_i.

Now consider what happens if we take a basis $\mathfrak{Y}$ in M_λ corresponding to a cyclic subspace decomposition of the nilpotent operator

$$N_\lambda = (T - \lambda I)|_{M_\lambda} = (T|_{M_\lambda}) - \lambda I_\lambda \qquad (I_\lambda = I|_{M_\lambda})$$

Then $[\lambda I_\lambda]_{\mathfrak{Y}}$ is λ times the identity matrix (as it is for any basis in M_λ) while $[N_\lambda]_{\mathfrak{Y}}$ consists of diagonal blocks, each an elementary nilpotent matrix N_i.

$$[\,N_\lambda]_{\mathfrak{Y}} = [\,(T-\lambda I)|_{M_\lambda}]_{\mathfrak{Y}} = \begin{pmatrix} \boxed{N_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \boxed{N_r} \end{pmatrix}$$

and

$$N_i = \begin{pmatrix} 0 & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \;\; 1 \\ 0 & & & 0 \end{pmatrix}$$

of size $d_i \times d_i$, with N_i a 1×1 zero matrix when $d_i = 1$. This yields the "**Jordan block decomposition**" of $T|_{M_\lambda}$

$$(12) \qquad [\,T|_{M_\lambda}]_{\mathfrak{Y}} \;=\; \lambda[I|_{M_\lambda}]_{\mathfrak{Y}} + [\,N_\lambda]_{\mathfrak{Y}} \;=\; \begin{pmatrix} \boxed{T_1} & & & & 0 \\ & \cdot & & & \\ & & \cdot & & \\ & & & \boxed{T_m} & \\ 0 & & & & \lambda\cdot\boxed{I_{r\times r}} \end{pmatrix}$$

with $T_i = \lambda\cdot I_{d_i\times d_i} + (\textit{elementary nilpotent})$ when $d_i > 1$,

$$T_i \;=\; \begin{pmatrix} \lambda & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \;\; 1 \\ 0 & & & \lambda \end{pmatrix}$$

The last block in (12) is exceptional. The other T_i correspond to the restrictions $T|_{C_i(\lambda)}$ to cyclic subspaces of dimension $d_i > 1$ in a cyclic subspace decomposition

$$M_\lambda(T) = C_1(\lambda) \oplus \,\ldots\, \oplus C_m(\lambda)$$

of the generalized eigenspace. However, some of the cyclic subspaces might be one-dimensional, and any such $C_i(\lambda)$ is contained in the ordinary eigenspace $E_\lambda(T)$. If there are r such degenerate cyclic subspaces we may lump them together into a single subspace

$$E = \bigoplus\{C_i(\lambda) : \dim(C_i(\lambda)) = 1\} \subseteq E_\lambda(T) \subseteq M_\lambda(T)$$

such that $\dim(E) = s$ and $T|_E = \lambda\cdot I_E$. It should also be evident that

$$s + d_1 + \ldots + d_m = \dim\big(M_\lambda(T)\big)$$

This is the Jordan Canonical Form for the restriction $\big(T|_{M_\lambda}\big)$ of T to a single generalized eigenspace. If $M_\lambda(T) \neq 0$ the description (12) is valid for any ground field $\mathbb{F}$, since it is really a result about the nilpotent operator $(T-\lambda)_{M_\lambda}$. Keep in mind that the T-invariant subspaces in a cyclic subspace decomposition of M_λ (or of any nilpotent operator) are not unique, but the number of cyclic subspaces in any decomposition and their dimensions are unique, and we get the same matrix form (12) for a suitably chosen basis.

VII-6. Generalized Eigenspace Decomposition of T.

So far we have only determined the structure of T restricted to a single generalized eigenspace $M_\lambda(T)$. Several obstacles must be surmounted to arrive at a similar structure for T on all of V.

- If the generalized eigenspaces $M_{\lambda_i}(T)$ fail to span V, knowing the behavior of T only on their span
$$M = \sum_{\lambda \in \mathrm{sp}(T)} M_\lambda(T)$$
leaves the gobal behavior of T beyond reach.

- It is equally important to prove (as we did in Proposition 4.5 for ordinary eigenspaces) that the span of the generalized eigenspaces is in fact a DIRECT sum,
$$M = \bigoplus_{\lambda \in \mathrm{sp}(T)} M_\lambda(T)$$
That means the actions of T on different M_λ are *independent* and can be examined separately, yielding a decomposition $T|_M = \bigoplus_{\lambda\in\mathrm{sp}(T)} \left(T|_{M_\lambda}\right)$.

Both issues will be resolved in our favor for operators $T : V \to V$ provided that the characteristic polynomial $p_T(x)$ splits into linear factors in $\mathbb{F}[x]$. This is always true if $\mathbb{F} = \mathbb{C}$; we are not so lucky for linear operators over $\mathbb{F} = \mathbb{R}$ or over a finite field such as $\mathbb{F} = \mathbb{Z}_p$. When this program succeeds the result is the Jordan Canonical Decomposition.

6.1. Theorem (Jordan Canonical Form). *If $T : V \to V$ is a linear operator on a finite dimensional space whose characteristic polynomial $p_T(x) = \det(T - x\,I)$ splits over $\mathbb{F}$, then V is a direct sum of its generalized eigenspaces*
$$V = \bigoplus_{\lambda \in \mathrm{sp}(T)} M_\lambda(T) \ ,$$
and since the $M_\lambda(T)$ are T-invariant we obtain a decomposition of T itself
$$T = \bigoplus_{\lambda \in \mathrm{sp}(T)} T|_{M_\lambda(T)} \tag{13}$$
into operators, each of which can be put into Jordan upper triangular form (12) *by choosing bases compatible with a decomposition of M_λ into T-invariant cyclic subspaces.*

Proof that the generalized eigenspaces are independent components in a direct sum follows the same lines as a similar result for ordinary eigenspaces (Proposition VII-4.5), but with more technical complications. Proof that they span V will require some new ideas based on the Fitting decomposition.

6.2. Proposition (Independence of the M_λ). *The span $M = \sum_{\lambda\in\mathrm{sp}(T)} M_\lambda(T)$ of the generalized eigenspaces (which may be a proper subspace in V) is always a direct sum, $M = \bigoplus_{\lambda\in\mathrm{sp}(T)} M_\lambda(T)$.*

Proof: Let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues in $\mathbb{F}$. By definition of "direct sum" we must show the components M_λ are independent, so that
$$0 = v_1 + \cdots + v_r, \ \text{with } v_i \in M_{\lambda_i} \ \Rightarrow \ \text{ each term } v_i \text{ is zero.} \tag{14}$$

Fix an index k. For for each $1 \le i \le r$, let $m_j = \deg(T - \lambda_j I)|_{M_{\lambda_j}}$. If $v_k = 0$ we're done; and if $v_k \ne 0$ let $m \le m_k$ be the smallest exponent such that $(T - \lambda_k I)^m v_k = 0$ and $(T - \lambda_k I)^{m-1} v_k \ne 0$. Then $w = (T - \lambda_k I)^{m-1} v_k$ is a nonzero eigenvector in E_{λ_k}.

Define

$$A = \prod_{i \ne k} (T - \lambda_i I)^{m_i} \cdot (T - \lambda_k)^{m-1}$$

which is keyed to the particular index λ_k as above. We then have

$$\begin{aligned} 0 &= A(0) = 0 + Av_k \qquad (\text{since } Av_i = 0 \text{ for } i \ne k) \\ &= \prod_{i \ne k} (T - \lambda_i)^{m_i} \big((T - \lambda_k)^{m-1} v_k\big) = \prod_{i \ne k} (T - \lambda_i)^{m_i} w \\ &= \prod_{i \ne k} \big[(T - \lambda_k) + (\lambda_k - \lambda_i)\big]^{m_i} w \qquad (\text{a familiar algebraic trick}) \\ &= \prod_{i \ne k} \sum_{s=0}^{m_i} \tbinom{m_i}{s} (T - \lambda_k)^{m_i - s} (\lambda_k - \lambda_i)^s w \qquad (\text{binomial expansion of } [\cdots]^{m_i}) \end{aligned}$$

All terms in the binomial sum are zero except when $s = m_i$, so we get

$$0 = \big[\prod_{i \ne k} (\lambda_k - \lambda_i)^{m_i}\big] \cdot w$$

The factor $[\cdots]$ is nonzero because the λ_i are the distinct eigenvalues of T in $\mathbb{F}$, so w must be zero. This is a contradiction because $w \ne 0$ by definition. We conclude that *every* term v_k in (14) is zero, so the span M is a direct sum of the M_λ. $\square$

Further Properties of Characteristic Polynomials. Before takling up the proof of Theorem 6.1 we digress to develop a few more facts about characteristic polynomials, in order to work out the relationship between $\mathrm{sp}(T)$ and $\mathrm{sp}\big(T|_{R_\infty}\big)$ where $R_\infty = R_\infty(T - \lambda_1 I)$.

6.3. Lemma. *If $A \in \mathrm{M}(n, \mathbb{F})$ has form* $A = \left(\begin{array}{c|c} B & * \\ \hline 0 & C \end{array}\right)$ *where B is $r \times r$ and C is $(n-r) \times (n-r)$, then* $\det(A) = \det(B) \cdot \det(C)$.

6.4. Corollary. *If $A \in \mathrm{M}(n, \mathbb{F})$ is upper triangular with values $c_1, \ldots, c_n$ on the diagonal, then* $\det(A) = \prod_{i=1}^{n} c_i$.

Proof (Lemma 6.3): Consider an $n \times n$ template corresponding to some $\sigma \in S_n$. If any of the marked spots in columns $C_1, \cdots, C_r$ occur in a row R_i with $r + 1 \le i \le n$, then $a_{ij} = a_{i,\sigma(i)} = 0$ and so is the corresponding term in $\sum_{\sigma \in S_n} (\cdots)$. Thus all columns $C_j, 1 \le j \le r$, must be marked in rows $R_1, \ldots, R_r$ if the template is to yield a nonzero term in $\det(A)$. It follows immediately that all columns C_j with $r + 1 \le j \le n$ must be marked in rows R_i with $r + 1 \le i \le n$ if σ is to contribute to

$$\det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma(i)}$$

Therefore only permutations σ that leave invariant the blocks of indices $[1, r]$, $[r+1, n]$ can contribute. These σ are composites of permutations $\mu = \sigma|_{[1,r]} \in S_r$ and $\tau = \sigma|_{[r+1,n]} \in S_{n-r}$, with

$$\sigma(k) = \mu \times \tau(l) = \begin{cases} \mu(k) & \text{if } 1 \le k \le r \\ \tau(k - r) & \text{if } r + 1 \le k \le n \end{cases}$$

Furthermore, we have $sgn(\sigma) = sgn(\mu\times\tau) = sgn(\mu)\cdot sgn(\tau)$ by definition of sgn, because μ, τ operate on disjoint subsets of indices in $[1, n]$.

In the matrix A we have

$$\begin{array}{rcll} B_{k,\ell} & = & A_{k,\ell} & \text{for } 1 \leq k, \ell \leq r \\ C_{k,\ell} & = & A_{k+r,\ell+r} & \text{for } r+1 \leq k, \ell \leq n \end{array}$$

so we get

$$\begin{aligned} \det(A) &= \sum_{(\mu,\tau)\in S_r\times S_{n-r}} sgn(\mu\times\tau)\cdot\Big(\prod_{i=1}^{r} B_{i,\mu(i)}\Big)\cdot\Big(\prod_{j=1}^{n-r} C_{j,\tau(j)}\Big) \\ &= \Big(\sum_{\mu\in S_r} sgn(\mu)\cdot\prod_{i=1}^{r} B_{i,\mu(i)}\Big)\cdot\Big(\sum_{\tau\in S_{n-r}} sgn(\tau)\cdot\prod_{j=1}^{n-r} C_{j,\tau(j)}\Big) \\ &= \det(B)\cdot\det(C) \quad \square \end{aligned}$$

6.5. Corollary. *If $T : V \rightarrow V$ is a linear operator on a finite dimensional vector space and $M \subseteq V$ is a T-invariant subspace, the characteristic polynomial $p_{T|_M}(x)$ divides $p_T(x) = \det(T - x\,I)$ in $\mathbb{F}[x]$.*

Proof: If $M \subseteq V$ is T-invariant and we take a basis $\mathfrak{X} = \{e_i\}$ that first spans M and then picks up additional vectors to get a basis for V, the matrix $[T]_{\mathfrak{X}}$ has block upper triangular form $\left(\begin{array}{c|c} A & * \\ \hline 0 & B \end{array}\right)$, and then

$$[T - xI]_{\mathfrak{X}} = \left(\begin{array}{c|c} A - xI & * \\ \hline 0 & B - xI \end{array}\right)$$

But it is trivial to check that $A - xI = \big[T - xI)|_M\big]_{\mathfrak{X}'}$ where $\mathfrak{X}' = \{e_1, \cdots, e_r\}$ are the initial vectors that span M. Thus $\det(A - xI)] = \det\big((T - xI)|_M\big) = p_{T|M}(x)$ divides $p_T(x) = \det(A - xI)\cdot\det(B - xI)$. $\square$

6.6. Exercise. Let (V, M, T) be as in the previous corollary. Then T induces a linear map $\tilde{T}$ from $V/M \rightarrow V/M$ such that $\tilde{T}(v + M) = T(v) + M$, for $v \in V$. Prove that the characteristic polynomial $p_{\tilde{T}}(x) = \det_{V/M}(\tilde{T} - xI)$, also divides $p_T(x) = \det(A - xI)\cdot\det(B - xI)$. $\square$

6.7. Lemma. *If f and P are nonconstant polynomials in $\mathbb{F}[x]$ and P divides f, so $f(x) = P(x)Q(x)$ for some other $Q \in \mathbb{F}[x]$, then $P(x)$ must split over $\mathbb{F}$ if $f(x)$ does.*

Proof: If Q is constant there is nothing to prove. Nonconstant polynomials $f \neq 0$ in $\mathbb{F}[x]$ have unique factorization into irreducible polynomials $f = \prod_{i=1}^{r} F_i$, where F_i cannot be written as a product of nonconstant polynomials of lower degree. Each polynomial f, P, Q has such a factorization $P = \prod_{k=1}^{m} P_k$, $Q = \prod_{j=1}^{m} Q_j$ so $f = PQ = \prod_k P_k \cdot \prod_j Q_j$. Since f splits over $\mathbb{F}$ it can also be written as a product of linear factors $f(x) = \prod_{i=1}^{n}(x - \alpha_i)$ where $\{\alpha_i\}$ are the roots of $f(x)$ in $\mathbb{F}$, counted according to their multiplicities. Linear factors $(x - \alpha)$ are obviously irreducible and the two irreducible decomposition of $f(x)$ must agree. Thus P (and Q) are products of linear factors and $P(x)$ splits over $\mathbb{F}$. $\square$

This lemma has a useful Corollary.

6.8. Corollary. *If the characteristic polynomial $p_T(x)$ of a linear operator $T : V \rightarrow V$ splits over $\mathbb{F}$, so does $p_{T|_W}$ for any T-invariant subspace $W \subseteq V$.*

Over $\mathbb{F} = \mathbb{C}$, all non-constant polynomials split.

Proof of Theorem 6.1. T has eigenvalues because p_T splits and its distinct eigenvalues $\{\lambda_1, \cdots, \lambda_r\}$ are the distinct roots of p_T in $\mathbb{F}$. Recall that $E_\lambda \neq \{0\} \Leftrightarrow M_\lambda \neq \{0\}$.

Pick an eigenvalue λ_1 and consider the Fitting decomposition $V = K_\infty \oplus R_\infty$ with respect to the operator $(T - \lambda_1 I)$, so K_∞ is the generalized eigenspace $M_{\lambda_1}(T)$ while R_∞ is the stable range of $(T - \lambda_1 I)$. Both spaces are invariant under $T - \lambda_1 I$, and also under T since $\lambda_1 I$ commutes with T. It will be important to note that

$$\lambda_1 \textit{ cannot be an eigenvalue of } T|_{R_\infty},$$

for if $v \in R_\infty$ is nonzero then $(T - \lambda_1 I)v = 0 \Rightarrow v \in K_\infty \cap R_\infty = \{0\}$. Hence $\mathrm{sp}(T|_{R_\infty}) \subseteq \{\lambda_2, \cdots, \lambda_r\}$.

We now argue by induction on $n = \dim(V)$. There is little to prove if $n = 1$. [There is an eigenvalue, so $E_\lambda = V$ and $T = \lambda I$ on V.] So, assume $n > 1$ and the theorem has been proved for all spaces V' of dimension $< n$ and all operators $T' : V' \to V'$ such that $\det(T' - \lambda I)$ splits over $\mathbb{F}$. The natural move is to apply this inductive hypothesis to $T' = T|_{R_\infty(T - \lambda_1 I)}$ since $\dim(R_\infty) = \dim(V) - \dim(M_{\lambda_1}) < \dim(V) = n$. But to do so we must show $p_{T'}$ splits over $\mathbb{F}$. [If $\mathbb{F} = \mathbb{C}$, every polynomial in $\mathbb{C}[x]$ splits and this issue does not arise.]

By Corollary 6.8 the characteristic polynomial of $T' = T|_{R_\infty}$ splits over $\mathbb{F}$, and by induction on dimension $R_\infty(T')$ is a direct sum of generalized eigenspaces for the restricted operator T'.

$$R_\infty(T') = \bigoplus_{\mu \in \mathrm{sp}(T')} M_\mu(T') \ ,$$

where $\mathrm{sp}(T')$ = the distinct roots of $p|_{T'}$ in $\mathbb{F}$. To compare the roots of p_T and $p_{T'}$, we invoke the earlier observation that $p_{T'}$ divides p_T. Thus the roots $\mathrm{sp}(T') = \mathrm{sp}(T|_{R_\infty})$ are a subset of the roots $\mathrm{sp}(T)$ of $p_T(x)$, and in particular every eigenvalue μ for T' is an eigenvalue for T. Let's label the distinct eigenvalues of T so that

$$\mathrm{sp}(T') = \{\lambda_s, \lambda_{s+1}, \cdots, \lambda_r\} \subseteq \mathrm{sp}(T) = \{\lambda_1, \cdots, \lambda_r\}$$

(with $s > 1$ because $\lambda_1 \notin \mathrm{sp}(T|_{R_\infty})$, as we observed earlier).

Furthermore, for each $\mu \in \mathrm{sp}(T')$ the generalized eigenspace $M_\mu(T')$ is a subspace of $R_\infty \subseteq V$, and must be contained in $M_\mu(T)$ because $(T' - \mu I)^k v = 0 \Rightarrow (T - \mu I)^k v = 0$ for all $v \in R_\infty$. Thus,

$$R_\infty = \bigoplus_{\mu \in \mathrm{sp}(T')} M_\mu(T') \subseteq \sum_{\mu \in \mathrm{sp}(T')} M_\mu(T) \subseteq \sum_{\lambda \in \mathrm{sp}(T)} M_\lambda(T)$$

($R_\infty = \bigoplus_{\mu \in \mathrm{sp}(T')} M_\mu(T')$ by the induction hypothesis). Therefore the generalized eigenspaces M_λ, $\lambda \in \mathrm{sp}(T)$, must span V because

$$\begin{aligned}
V \ &= \ K_\infty \oplus R_\infty \ = \ M_{\lambda_1}(T) \oplus R_\infty \ \subseteq \ M_{\lambda_1}(T) \oplus \big(\bigoplus_{\mu \in \mathrm{sp}(T')} M_\mu(T') \big) \\
&\subseteq \ M_{\lambda_1}(T) + \big(\sum_{\mu \in \mathrm{sp}(T')} M_\mu(T) \big) \qquad \big(\text{because } M_\mu(T') \subseteq M_\mu(T)\big) \qquad (15) \\
&\subseteq \ M_{\lambda_1}(T) + \big(\sum_{\lambda \in \mathrm{sp}(T)} M_\lambda(T) \big) \ \subseteq \ V \qquad \big(\text{because } \mathrm{sp}(T') \subseteq \mathrm{sp}(T)\big)
\end{aligned}$$

Conclusion: the $M_\lambda(T)$, $\lambda \in \mathrm{sp}(T)$, span V so by Proposition 6.2 V is a direct sum of its generalized eigenspaces. That finishes the proof of Theorem 6.1. $\square$

It is worth noting that

$$\mathrm{sp}(T') = \{\lambda_2, \ldots, \lambda_r\} \quad \text{and} \quad M_{\lambda_i}(T') = M_{\lambda_i}(T) \text{ for } 2 \leq i \leq r\ .$$

Since $M_\mu(T') \subseteq M_\mu(T)$ for all $\mu \in \mathrm{sp}(T')$, and $M_{\lambda_1}(T) \cap V' = (0)$, λ_1 cannot appear in $\mathrm{sp}(T')$; on the other hand every $\mu \in \mathrm{sp}(T')$ must lie in $\mathrm{sp}(T)$.

Consequences. Some things can be proved using just the block upper-triangular form for T rather than the more detailed Jordan Canonical form.

6.9. Corollary. *If the characteristic polynomial of $T : V \to V$ splits over $\mathbb{F}$, and in particular if $\mathbb{F} = \mathbb{C}$, there is a basis $\mathfrak{X}$ such that $[T]_{\mathfrak{X}}$ has block upper triangular form*

$$[T]_{\mathfrak{X}} = \begin{pmatrix} \boxed{T_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \boxed{T_r} \end{pmatrix} \tag{16}$$

with blocks on the diagonal

$$T_i = \begin{pmatrix} \lambda_i & & & * \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \\ 0 & & & \lambda_i \end{pmatrix}$$

of size $m_i \times m_i$ such that

1. *$\lambda_1, \ldots, \lambda_r$ are the distinct eigenvalues of T.*
2. *The block sizes are the algebraic multiplicities m_i of the λ_i in the splitting of the characteristic polynomial $p_T(t)$ (see the next corollary for details).*
3. *$p_T(x) = (-1)^n \cdot \prod_{j=1}^r (x - \lambda_i)^{m_j}$ with $n = m_1 + \cdots + m_r$.*

The blocks T_i may or may not have off-diagonal terms. $\square$

6.10. Corollary. *If the characteristic polynomial of an $n \times n$ matrix A splits over $\mathbb{F}$, there is a similarity transform $A \mapsto SAS^{-1}$, $S \in \mathrm{GL}(n, \mathbb{F})$, such that SAS^{-1} has the block upper-triangular form shown above.*

6.11. Corollary. *If the characteristic polynomial of $T : V \to V$ splits over $\mathbb{F}$, and in particular if $\mathbb{F} = \mathbb{C}$, then for every $\lambda \in \mathrm{sp}(T)$ we have*

$$(\textit{algebraic multiplicity of } \lambda) = \dim(M_\lambda) = m_i$$

where m_i is the block size in (16)*).*

Proof: Taking a basis such that $[T]_{\mathfrak{X}}$ has the form (16), $[T - xI]_{\mathfrak{X}}$ will have the same form, but with diagonal entries λ_i replaced by $(\lambda_i - x)$. Then

$$\det[T - xI]_{\mathfrak{X}} = \prod_{j=1}^r (\lambda_j - x)^{\dim(M_{\lambda_j})} = p_T(x)$$

since the block T_j correspond to $T|_{M_{\lambda_j}}$. Obviously, the exponent on $(\lambda_j - x)$ is also the multiplicity of λ_j in the splitting of the characteristic polynomial p_T. $\square$

6.12. Corollary. *If the characteristic polynomial of an $n \times n$ matrix A, with distinct eigenvalues $\mathrm{sp}_{\mathbb{F}}(A) = \{\lambda_1, \ldots, \lambda_r\}$, splits over $\mathbb{F}$ then*

1. $\det(A) = \prod_{i=1}^{r} \lambda_i^{m_i}$, *the product of eigenvalues counted according to their algebraic multiplicities* m_i.

2. $\mathrm{Tr}(A) = \sum_{i=1}^{r} m_i \lambda_i$, *the sum of eigenvalues counted according to their algebraic multiplicities* m_i.

3. *More generally, if* $\mathbb{F} = \mathbb{C}$ *there are explicit formulas for all coefficients of the characteristic polynomial when we write it in the form*

$$p_A(x) = \det(A - x\,I) = \sum_{i=0}^{n} (-1)^i c_i(A)\, x^i$$

If eigenvalues are listed according to their multiplicities $m_i = m(\lambda_i)$, *say as* $\mu_1, \ldots, \mu_n$ *with* $n = \dim(V)$,

$$\mu_1 = \ldots = \mu_{m_1} = \lambda_1 \qquad \mu_{m_1+1} = \ldots = \mu_{m_1+m_2} = \lambda_2 \qquad \textit{etc}$$

then $c_n = 1$ *and*

$$\begin{aligned} c_{n-1} &= \sum_{j=1}^{n} \mu_j \; = \; \mathrm{Tr}(A), \\ &\vdots \\ c_{n-k} &= \sum_{j_1 < \ldots < j_k} \mu_{j_1} \cdot \ldots \cdot \mu_{j_k}, \\ &\vdots \\ c_0 &= \mu_1 \cdot \ldots \cdot \mu_n = \det(A) \end{aligned}$$

These formulas fail to be true if $\mathbb{F} = \mathbb{R}$ *and* $p_T(x)$ *has non-real roots in* $\mathbb{C}$.

6.13. Corollary. *If the characteristic polynomial of an* $n \times n$ *matrix* A *splits over* $\mathbb{F}$, *then* $T : V \to V$ *is diagonalizable if and only if*

$$(\textit{algebraic multiplicity}) = (\textit{geometric multiplicity}) \qquad \textit{for each } \lambda \in \mathrm{sp}(T).$$

Both multiplicities are then equal to $\dim(E_\lambda(T))$.

Proof: $E_\lambda \subseteq M_\lambda$ for every eigenvalue, and by the previous corollary we know that

$$(\textit{geometric multiplicity}) = \dim(E_\lambda) \leq \dim(M_\lambda) = (\textit{algebraic multiplicity})\ .$$

Furthemore, $M_\lambda = \ker(T - \lambda I)^N$ for large $N \in \mathbb{N}$. Writing $V = M_{\lambda_1} \oplus \ldots \oplus M_{\lambda_r}$, the implication ($\Leftarrow$) follows because

$$\begin{aligned} (\textit{algebraic multiplicity}) &= (\textit{geometric multiplicity}) \\ &\Rightarrow \dim(E_{\lambda_i}) = \dim(M_{\lambda_i}) \text{ for all } i \\ &\Rightarrow M_{\lambda_i} = E_{\lambda_i} \text{ since } M_{\lambda_i} \supseteq E_{\lambda_i} \\ &\Rightarrow V = \bigoplus_{i=1}^{r} E_{\lambda_i} \text{ and } T \text{ is diagonalizable.} \end{aligned}$$

For ($\Rightarrow$): if T is diagonalizable we have $V = \bigoplus_{i=1}^{r} E_{\lambda_i}$, but $E_{\lambda_i} \subseteq M_{\lambda_i}$ for each i. Comparing this with the Jordan decomposition $V = \bigoplus_{i=1}^{r} M_{\lambda_i}$ we see that $M_{\lambda_i} = E_{\lambda_i}$. □

6.14. Corollary. *If A is an $n \times n$ matrix whose characteristic polynomial splits over $\mathbb{F}$, let $\mathfrak{X}$ be a basis that puts $[T]_{\mathfrak{X}}$ into Jordan form, so that*

$$(17) \qquad [T]_{\mathfrak{X}} = \begin{pmatrix} \boxed{T_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \boxed{T_r} \end{pmatrix} \qquad \textit{with} \qquad T_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \ \ 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

Then with respect to the same basis the powers $I, T, T^2, \cdots$ take form:

$$(18) \qquad [T^k]_{\mathfrak{X}} = \begin{pmatrix} \boxed{T_1^k} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \boxed{T_r^k} \end{pmatrix} \qquad \textit{with} \qquad T_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \ \ 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

Proof: $[T_i^k]_{\mathfrak{X}} = \big([T_i]_{\mathfrak{X}}\big)^k$ for $k = 0, 1, 2, \cdots$. $\square$

In (18) there may be blocks of various sizes with the same diagonal values λ_i.

These particular powers $[T^k] = [T]^k$ are actually easy to compute. Each block T_i has the form $T_i = \lambda_i I + N_i$ with N_i an elementary nilpotent matrix, so we have

$$T_i^k = \sum_{j=0}^{k} \binom{k}{i} \lambda^{k-j} N_i^j \qquad \text{(binomial expansion)} ,$$

with $N_i^j = 0$ when $j \geq \deg(N_i)$.

6.15. Exercise. If N is an $r \times r$ elementary nilpotent matrix

$$N \ = \begin{pmatrix} 0 & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \ \ 1 \\ 0 & & & 0 \end{pmatrix} \qquad \text{show that} \qquad N^2 \ = \begin{pmatrix} 0 & 0 & 1 & 0 \\ & \cdot & \cdot & \\ & & \cdot & 1 \\ & & \cdot & 0 \\ 0 & & & 0 \end{pmatrix} ,$$

Each new multiple of N moves the diagonal file of $1's$ one step to the upper right, with $N^r = 0$ at the last step. $\square$

6.16. Exercise. If the characteristic polynomial of $T : V \to V$ splits over $\mathbb{F}$, there is a basis that puts $[T]_{\mathfrak{X}}$ in Jordan form, with diagonal blocks

$$A \ = \ \lambda I + N \ = \begin{pmatrix} \lambda & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \ \ 1 \\ 0 & & & \lambda \end{pmatrix}$$

Compute the exponential matrix

$$\mathrm{Exp}(tA) = e^{tA} = \sum_{n=0}^{\infty} \frac{t^n}{n!} A^n$$

for $t \in \mathbb{F}$.
Hint: If A and B commute we have $e^{A+B} = e^A \cdot e^B$; apply the previous exercise.
Note: Since N is nilpotent the exponential series is a finite sum. $\square$

The Spectral Mapping Theorem. Suppose A is an $n \times n$ matrix whose characteristic polynomial splits over $\mathbb{F}$, with

$$p_A(x) = \det(A - x\,I) = c \cdot \prod_{i=1}^{r} (\lambda_i - x\,I)^{m_i} \qquad \text{if } \mathrm{sp}(T) = \{\lambda_1, \ldots, \lambda_r\}\ .$$

Examination of the diagonal entries in the Jordan Canonical form (16), or even the upper triangular form (17), immediately reveals that

$$\mathrm{sp}(T^k) = \text{ the } \textit{distinct} \text{ entries in the list of powers } \lambda_1^k, \ldots, \lambda_r^k$$

Be aware that there might be repeated entries among the λ_i^k, even though the λ_i are distinct. (Devise an example in which $\mathrm{sp}(T^k)$ reduces to a single point even though $\mathrm{sp}(T)$ contains several distinct points.)

Therefore the characteristic polynomial of T^k is the product of the diagonal entries $(\lambda_i^k - x)$ in the Jordan form of $(T^k - x\,I)$,

$$p_{T^k}(x) \ = \ \det(T^k - x\,I) \ = \ \prod_{i=1}^{r} (\lambda_i^k - x)^{m_i}\ , \qquad (\text{where } m_i = \dim(M_{\lambda_i}(T))\ .$$

More can be said under the same hypotheses. If $f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_m t^m$ is any polynomial in $\mathbb{F}[t]$ we can form an operator $f(T)$ that maps $V \to V$ to obtain a natural corresponding $\Phi : \mathbb{F}[t] \to \mathrm{Hom}_{\mathbb{F}}(V)$ such that

$$\Phi(1) = I, \quad \Phi(t) = T, \quad \Phi(t^k) = T^k$$

and

$$\begin{aligned}
\Phi(f_1 + f_2) &= \Phi(f_1) + \Phi(f_2) && \text{(sum of linear operators)}\\
\Phi(f_1 \cdot f_2) &= \Phi(f_1) \circ \Phi(f_2) && \text{(composition of linear operator)}\\
\Phi(cf) &= c \cdot \Phi(f) && \text{for all } c \in \mathbb{F}
\end{aligned}$$

I.e. Φ is a homomorphism between associative algebras. With this in mind we can prove:

6.17. Theorem. (Spectral Mapping Theorem). *If the characteristic polynomial of a linear operator $T : V \to V$ splits over $\mathbb{F}$, and in particular if $\mathbb{F} = \mathbb{C}$, every polynomial $f(t) = \sum_{i=0}^{m} a_i t^i$ in $\mathbb{F}[t]$ determines an operator $\Phi(f)$ in $\mathrm{Hom}_{\mathbb{F}}(V, V)$,*

$$\Phi(f) = \sum_{i=0}^{m} a_i T^i$$

The correspondence $\Phi : \mathbb{F}[t| \to \mathrm{Hom}_{\mathbb{F}}(V)$ is a unital homomorphism of associative algebras and has the following SPECTRAL MAPPING PROPERTY

(19) $$\mathrm{sp}(f(T)) = f\big(\mathrm{sp}(T)\big) = \{f(z) : z \in \mathrm{sp}(T)\}$$

In particular, this applies if T is diagonalizable over $\mathbb{F}$.

Proof: It suffices to choose a basis $\mathfrak{X}$ such that $[T]_{\mathfrak{X}}$ has block upper-triangular form

$$[T]_{\mathfrak{X}} \ = \begin{pmatrix} T_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & T_r \end{pmatrix} \qquad \text{with} \qquad T_i \ = \begin{pmatrix} \lambda_i & & & * \\ & \cdot\ \cdot & & \\ & & \cdot\ \cdot & \\ & & & \cdot \\ 0 & & & \lambda_i \end{pmatrix}$$

of size $m_i \times m_i$ since $[T^k]_{\mathfrak{X}} = [T]_{\mathfrak{X}}^k$ (matrix product) for $k = 0, 1, 2, \cdots$. Hence

$$[f(T)]_{\mathfrak{X}} = f\big([T]_{\mathfrak{X}}\big) = a_0 I + a_1 [T]_{\mathfrak{X}} + \cdots + a_m [T]_{\mathfrak{X}}^m$$

because $f(A) = a_0 I + a_1 A + \cdots + a_m A^m$ for any matrix A.

As in the previous corollary, it follows that $[f(T)]_{\mathfrak{X}}$ is made up of blocks on the diagonal, each of which is upper-triangular with diagonal values $f(\lambda_i)$; then the characteristic polynomial of $f(T)$ is

$$p_{f(T)}(x) = \det(f(T) - x\, I) = \prod_{i=1}^{r} (f(\lambda_i) - x\, I)^{m_i}, \qquad m_i = \dim(M_{\lambda_i}) \ .$$

This is zero if and only if $x = f(\lambda_i)$ for some i, so $\mathrm{sp}(f(T)) = \{f(\lambda_i) : 1 \leq i \leq r\} = f(\mathrm{sp}(T))$. Obviously the characteristic polynomial of $f(T)$ splits over $\mathbb{F}$ if $p_T(t)$ splits.

Here $\lambda_i \in \mathrm{sp}(T) \Rightarrow f(\lambda_i) \in \mathrm{sp}(f(T))$, but the multiplicity of $f(\lambda_i)$ as an eigenvalue of $f(T)$ might be greater than the multiplicity of λ_i as an eigenvalues of T because we might have $f(\lambda_i) = f(\lambda_j)$, and then $\mu = f(\lambda_i)$ will have multiplicity at least $m_i + m_j$ in $\mathrm{sp}(f(T))$.

Another consequence is the Cayley-Hamilton theorem, which can be proved in other ways without developing the Jordan Canonical form. However this normal form suggests the underlying reason why the result is true, and makes its validity almost obvious. On the other hand, alternative proofs can be made to work for arbitrary $\mathbb{F}$ and T, without any assumptions about the characteristic polynomial $p_T(x)$. Since the result is true in this generality, we give both proofs.

6.18. Theorem. (Cayley-Hamilton). *For any linear operator $T : V \to V$ on a finite dimensional vector space, over any $\mathbb{F}$, we have*

$$p_T(T) = \big[\, p_T(t)\big|_{t=T}\big] \ = \ 0 \quad (\textit{the zero operator in } \mathrm{Hom}_{\mathbb{F}}(V, V)) \ ,$$

Thus, applying the characteristic polynomial $p_T(x) = \det(T - x\, I)$ to the linear operator T itself yields the zero operator.

Proof: If $p_T(x)$ splits over $\mathbb{F}$ we have $p_T(x) = \prod_{i=1}^{r} (\lambda_i - x)^{m_i}$, where $m_i = \dim(M_{\lambda_i})$ and $\{\lambda_1, \cdots, \lambda_r\}$ are the (distinct) eigenvalues in $\mathrm{sp}_{\mathbb{F}}(T)$. We want to show that

$$0 = \prod_{i=1}^{r} (T - \lambda_i I)^{m_i} = \big[\, p_T(x)\big|_{x=T}\big]$$

But $V = \bigoplus_{i=1}^{r} M_{\lambda_i}$ and $(T - \lambda_i I)^{m_i}(M_{\lambda_i}) = (0)$ [Given a Jordan basis in M_{λ_i}, $A = [(T - \lambda_i I)|_{M_{\lambda_i}}]_{\mathfrak{X}}$ consists of elementary nilpotent blocks N_j on the diagonal; the size $d_j \times d_j$ of such a block cannot exceed $m_i = \dim(M_{\lambda_i})$, so $N_j^{d_j} = N_j^{m_i} = 0$ for each j.] Hence $\prod_{j=1}^{r} (T - \lambda_i I)^{m_i}\big(M_{\lambda_i}\big) = (0)$, so the operator $p_T(T)$ is zero on each M_{λ_i} and on all of V. $\square$

If p_T does not split over $\mathbb{F}$, a different argument shows that $p_T(T)v = 0$ for all $v \in V$.

Alternative Proof (6.18): The result is obvious if $v = 0$. If $v \neq 0$ there is a largest $m \geq 1$ such that $v, T(v), T^2(v), \cdots, T^{m-1}(v)$ are linearly independent. Then

$$W = \mathbb{F}\text{-span}\{T^k(v) : k \in \mathbb{N}\} = \mathbb{F}\text{-span}\{T^k(v) : 0 \leq k \leq m-1\}$$

and $\{v, T(v), \cdots, T^{m-1}(v)\}$ is a basis for the cyclic subspace W. This space is clearly T-invariant, and as we saw before, $p_{(T|_W)}$ divides p_T, so that $p_T(x) = p_{(T|_W)}(x) \cdot Q(x)$ for some $Q \in \mathbb{F}[x]$. We now compute $p_{(T|_W)}(x)$. For the basis $\mathfrak{X} = \{v, T(v), \cdots, T^{m-1}(v)\}$

we note that $T^m(v) = T(T^{m-1}(v))$ is a unique linear combination of the previous vectors $T^k(v)$, say

$$T^m(v) + a_{m-1}T^{m-1}(v) + \cdots + a_1T(v) + a_0 v = 0 \ .$$

Hence,

$$\big[T|_W\big]_{\mathfrak{X}} = \begin{pmatrix} 0 & & & 0 & -a_0 \\ 1 & \cdot & & & -a_1 \\ & \cdot & \cdot & & \cdot \\ & & \cdot & \cdot & \cdot \\ 0 & & & 1 & -a_{m-1} \end{pmatrix} \quad \text{and} \quad \big[(T-x\,I)|_W\big]_{\mathfrak{X}} = \begin{pmatrix} -x & & & 0 & -a_0 \\ 1 & \cdot & & & -a_1 \\ & \cdot & \cdot & & \cdot \\ & & \cdot & \cdot & \cdot \\ 0 & & & 1 & -x-a_{m-1} \end{pmatrix}$$

6.19. Exercise. Show that

$$p_{(T|_W)}(x) = \det(T|_W - x\,I)_{\mathfrak{X}} = (-1)^m\big(t^m + a_{m-1}t^{m-1} + \cdot + a_1t + a_0\big) \ . \ \square$$

It follows that $p_{(T|_W)}(T)$ satisfies the equation

$$p_{(T|_W)}(T)v = (-1)^m\big(T^n(v) + a_{m-1}T^{m-1}(v) + \cdot + a_1T(v) + a_0v\big) = 0$$

by definition of the coefficients $\{a_j\}$. But then

$$p_T(T)v = Q(T)\cdot[\,p_{(T|_W)}(T)v] = Q(T)(0) = 0 \ .$$

(Recall that $W = \mathbb{F}\text{-span}\{T^k(v)\}$ as in Propositions 2.5 and 2.7.) Since this is true for all $v \in V$, $p_T(T)$ is the zero operator in $\mathrm{Hom}_{\mathbb{F}}(V,V)$. $\square$

Remarks: If $T : V \to V$ is a linear operator on a finite dimensional vector space the polynomial $Q(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_0$ in $\mathbb{F}[x]$ of minimal degree such that $Q(T) = 0$ is called the *minimal polynomial* for T. The polynomial function of T defined above by substituting $x = T$

$$T^m + a_{m-1}T^{m-1} + \cdots + a_1T + a_0I = 0$$

is precisely the minimal polynomial for T. The Jordan form (12) can be used to determine the minimal polynomial, but the block upper-triangular form (11) is too crude for this purpose. (The problem is that the nilpotence degree $\deg(N)$ of a nilpotent matrix will be greater than the degree of the minimal polynomial unless there is a cyclic vector in V.) $\square$

6.20. Example. Let $T : V \to V$ be a linear map on $V = \mathbb{R}^4$ whose matrix with respect to the standard basis $\mathfrak{X} = \{e_1, \cdots, e_4\}$ has the form

$$A = [T]_{\mathfrak{X}} \ = \begin{pmatrix} 7 & 1 & 2 & 2 \\ 1 & 4 & -1 & -1 \\ -2 & 1 & 5 & -1 \\ 1 & 1 & 2 & 8 \end{pmatrix} \quad \text{so that} \quad A - 6I \ = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & -2 & -1 & -1 \\ -2 & 1 & -1 & -1 \\ 1 & 1 & 2 & 2 \end{pmatrix}$$

After some computational work which we omit, we find that

$$p_T(t) = \det(A - xI) = (x-6)^4 = x^4 - 4(6x^3) + 6(6^2x^2) - 4(6^3x) + 6^4 \ ,$$

so $\mathrm{sp}_{\mathbb{R}}(T) = \{6\}$ with algebraic multiplicity $m = 4$. Thus $V = M_{\lambda=6}(T)$ and $(T - 6I)$ is nilpotent. We find $K_1 = \ker(T-6I) = E_{\lambda=6}(T)$ by row reduction of $[T-6I]_{\mathfrak{X}} = [A-6I\,]$,

$$[A-6I] \to \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & -3 & -3 & -3 \\ 0 & 3 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} \boxed{1} & 1 & 2 & 2 \\ 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

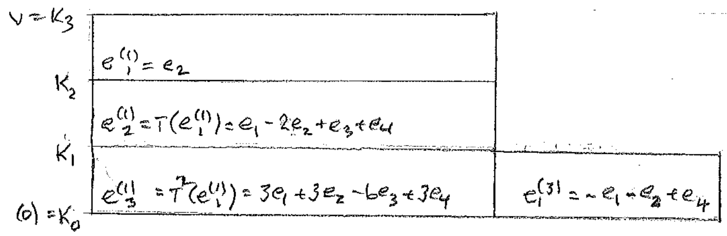

Figure 7.2. The version of Figure 7.1 worked out in Example 6.20. Although there are three columns, Column 2 is empty and has been deleted in the present diagram. All the basis vectors $e_j^{(i)}$ are shown

Thus,

$$\begin{aligned} K_1 &= \{v = (-s-t, -s-t, s, t) : s, t \in \mathbb{R}\} \\ &= \mathbb{R}\text{-span}\{f_1^{(1)} = -e_1 - e_2 + e_3\,,\ f_2^{(1)} = -e_1 - e_2 + e_4\} \\ &= \mathbb{R}\text{-span}\{(-1,-1,1,0)\,,\ (-1,-1,0,1)\} \end{aligned}$$

and $\dim(K_1) = 2$. Next row reduce $\ker(A - 6I\,)^2$ to get

$$[A - 6I\,]^2 \rightarrow \begin{pmatrix} 0 & 3 & 3 & 3 \\ 0 & 3 & 3 & 3 \\ 0 & -6 & -6 & -6 \\ 0 & 3 & 3 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & \boxed{1} & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The first column (which meets no "step corner") corresponds to free variable x_1; the other free variables are x_3, x_4. Thus $K_2 = \ker(A - 6I)^2$ is

$$\begin{aligned} K_2 &= \{v = (a, -b-c, b, c) : a, b, c \in \mathbb{R}\} \\ &= \mathbb{R}\text{-span}\{f_1^{(2)} = e_1\,,\ f_2^{(2)} = e_3 - e_2\,,\ f_3^{(2)} = e_4 - e_2\} \ = \ \mathbb{R}\text{-span}\{e_1, e_3 - e_2, e_4 - e_2\} \end{aligned}$$

and $\dim(K_2) = 3$. Finally $(A - 6I)^3 = 0$, so $\deg(T - 6I\,) = 3$ and $K_3 = V$.

We now apply the procedure for finding cyclic subspaces outlined in Figure 7.1.

Step 1: Find a basis for V mod K_2. Since $\dim(V/K_2) = 1$ this is achieved by taking any $v \in V \sim K_2$. One such choice is $e_1^{(1)} = e_2 = (0,1,0,0)$, which obviously is not in K_2. Then compute its images under powers of $(T - 6I)$,

$$\begin{aligned} e_2^{(1)} &= (T - 6I)e_1^{(1)} \ = \ (1,-2,1,1) \ = \ e_1 - 2e_2 + e_3 + e_4 \in K_2 \sim K_1 \\ e_3^{(1)} &= (T - 6I)^2 e_1^{(1)} \ = \ (3,3,-6,3) \ = \ 3(e_1 + e_2 - 2e_3 + e_4) \in K_1 \sim \{0\} \end{aligned}$$

Step 2: There is no need to augment the vector

$$e_2^{(1)} = (T - 6I)e_1^{(1)} = (1,-2,1,1) \in K_2$$

to get a basis for K_2/K_1, because $\dim(K_2/K_1) = 1$.

Step 3: In $K_1 \sim \{0\}$ we must augment $e_3^{(1)} = (T - 6I)^2 e_1^{(1)}$ to get a basis for $K_1/K_0 \cong K_1$. We need a new vector $e_1^{(3)} \in K_1 \sim \{0\}$ such that $e_1^{(3)}$ and $e_3^{(1)}$ are independent mod $K_0 = (0)$ – i.e. vectors that are actually independent in V. We could try $e_1^{(3)} =$

$(-1,-1,0,1) = -e_1 - e_2 + e_4$ which is in $K_1 \sim K_0$. Independence holds if and only if the matrix M whose rows are $e_3^{(1)}$ and $e_1^{(3)}$ has $rank = 2$. But row operations yield

$$\begin{pmatrix} 1 & 1 & -2 & 1 \\ -1 & -1 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} \boxed{1} & 1 & -2 & 1 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

which has row $rank = 2$, as desired.

Thus $\{T^2(e_1^{(1)}), T(e_1^{(1)}), e_1^{(1)}\,;\, e_1^{(3)}\}$ is a basis for all of V such that

$$\begin{aligned} C_1 &= \mathbb{R}\text{-span}\{T^2(e_1^{(1)}), T(e_1^{(1)}), e_1^{(1)}\}, \\ C_2 &= \mathbb{R}\text{-span}\{e_1^{(3)}\} \end{aligned}$$

are independent cyclic subspaces, generated by the vectors $e_1^{(1)}$ and $e_1^{(3)}$. The ordered basis $\mathfrak{Y} = \{T^2(e_1^{(1)}), T(e_1^{(1)}), e_1^{(1)}\,;\, e_1^{(3)}\}$ puts $[T]_{\mathfrak{Y}}$ in Jordan canonical form

$$[T]_{\mathfrak{Y}} = \left(\begin{array}{ccc|c} 6 & 1 & 0 & 0 \\ 0 & 6 & 1 & 0 \\ 0 & 0 & 6 & 0 \\ \hline 0 & 0 & 0 & 6 \end{array}\right)$$

Basis vectors $T^2(e_1^{(1)})$ and $e_1^{(3)}$ are eigenvectors for the action of T and $E_{\lambda=6}(T) = \mathbb{F}\text{-span}\{e_1^{(3)}, T^2(e_1^{(3)})\}$ is 2-dimensional. $\square$

6.21. Exercise. Find the Jordan canonical form for the linear operators $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ whose matrices with respect to the standard bases $\mathfrak{X} = \{e_1, \cdots, e_n\}$ are

$$\text{(a)} \quad A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix} \qquad \text{(b)} \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

The Minimal Polynomial for $T : V \rightarrow V$. The space of linear operators $\mathrm{Hom}_{\mathbb{F}}(V, V) \cong \mathrm{M}(n, \mathbb{F})$ is finite dimensional, so there is a largest exponent m such that $I, T, T^2, \ldots, T^{m-1}$ are linearly independent. Thus there are coefficients c_j such that $T^m + \sum_{j=0}^{m-1} c_j T^j = 0$ (the zero operator). The monic polynomial

$$x^m + \sum_{j=0}^{m-1} c_j x^j \ \text{ in } \mathbb{F}[x]$$

is the (unique) **minimal polynomial** $m_T(x)$ for this operator. Obviously $d = \deg(m_T)$ cannot exceed $n^2 = \dim\big(\mathrm{M}(n, \mathbb{F})\big)$, but it could be a lot smaller. The minimal polynomial for a matrix $A \in \mathrm{M}(n, \mathbb{F})$ is defined the same way, and it is easy to see that the minimal polynomial $m_T(x)$ for a linear operator is the same as the minimal polynomial of the associated matrix $A = [T]_{\mathfrak{X}}$, and this is so for *every* basis $\mathfrak{X}$ in V. Conversely the minimal polynomial $m_A(x)$ of a matrix coincides with that of the linear operator $L_A : \mathbb{F}^n \rightarrow \mathbb{F}^n$ such that $L_A(v) = A \cdot v$ (matrix product).

Computing $m_T(x)$ could be a chore, but it is easy if we know the Jordan form for T, and this approach also reveals interesting connections between the minimal polynomial $m_T(x)$ and the characteristic polynomial $p_T(x) = \det(T - x\,I)$. We have already seen that the characteristic polynomial is a "similarity invariant" for matrices (or for linear operators), so that

A similarity transformation, $A \mapsto SAS^{-1}$ yields a new matrix with the same characteristic polynomial, so $p_{SAS^{-1}}(x) = p_A(x)$ in $\mathbb{F}[x]$ for all invertible matrices S.

(See Section II.4 of the *Linear Algebra I Class Notes* for details regarding similarity transformations, and Section V.1 for invariance of the characteristic polynomial.) The minimal polynomial is also a similarity invariant, a fact that can easily be proved directly from the definitions.

6.22. Exercise. Explain why the minimal polynomial is the same for:

1. A matrix A and the linear operator $L_A : \mathbb{F}^n \to \mathbb{F}^n$.
2. A linear operator $T : V \to V$ on a finite dimensional vector space and its matrix $A = [T]_{\mathfrak{X}}$ with respect to any basis in V $\square$

6.23. Exercise. Prove that the minimal polynomial $m_A(x)$ of a matrix $A \in \mathrm{M}(n, \mathbb{F})$ is a similarity invariant: $m_{SAS^{-1}}(x) = m_A(x)$ for any invertible $n \times n$ matirx S. $\square$

We will use Example 6.22 to illustrate how the minimal polynomial can be found from the Jordan form, but first let's compute and compare $m_A(x)$ and $p_A(x)$ for a diagonal matrix. If

$$A = \begin{pmatrix} \boxed{\lambda_1 I_{d_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_r I_{d_r}} \end{pmatrix} \qquad \text{where } I_k = k \times k \text{ identity matrix.}$$

The characteristic polynomial obviously depends only on the diagonal values of A, is $p_A(x) = \prod_{j=1}^{r}(\lambda_j - x)^{d_j}$; in contrast, we will show that the minimal polynomial is the product of the distinct factors,

$$m_A(x) = \prod_{j=1}^{r}(\lambda_j - x) \ ,$$

each taken with multiplicity one – i.e. for diagonal matrices, $m_A(x)$ is just $p_A(x)$, ignoring multiplicities.

ADD MORE TEXT RE: MIN POLYN ?

VII-7. The Jordan Form and Differential Equations.

Computing the exponential $e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$ of a matrix turns out to be important in many applications, one of which will be illustrated below. This is not an easy calculation if you try to do it by summing the series. In fact, computing a product of $n \times n$ matrices seems to require n^3 multiplication operations on matrix entries and computing a high power such as A^{200} directly could be a formidable task. The computing effort can be reduced somewhat through clever programming, but it can be done by hand for diagonal matrices, and for elementary nilpotent matrices (see Exercises 6.15 - 6.16), and hence for any matrix that is already in Jordan Canonical form. For a linear operator $T : V \to V$ the Jordan form is obtained by choosing a suitable basis in V; for a matrix A this amounts to finding an invertible matrix S such that the similarity transform $A \mapsto SAS^{-1}$ puts A into Jordan form. Similarity transforms are invertible operations that interact nicely with the matrix-exponential operation, with

$$(20) \qquad Se^A S^{-1} = e^{SAS^{-1}} \qquad \text{for every } A \in \mathrm{M}(n, \mathbb{C}) \ .$$

Thus if we can determine the Jordan form B for A, we can compute e^A in four simple steps,

$$\begin{aligned} A \;&\to\; \text{Jordan form } B = SAS^{-1} \\ &\to\; e^B = e^{SAS^{-1}} \quad \text{(a calculation that can be done by hand)} \\ &\to\; e^A = S^{-1}e^BS = e^{S^{-1}BS} \qquad \text{(because } B = SAS^{-1}\text{)} \end{aligned}$$

Done. Note that e^D will have block upper-triangular form if D has Jordan form.

This idea was illustrated for *diagonalizable* matrices and operators in the *Linear Algebra I Notes*, but in the following example you will see that it is easily adapted to deal with matrices whose Jordan form can be determined. Rather that go through such a calculation just to compute e^A, we will go the extra parsec and show how the ability to compute matrix exponentials is the key to solving systems of constant coefficient differential equations.

Solving Linear Systems of Ordinary Differential Equations.

If $A \in \mathrm{M}(n, \mathbb{F})$ for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))$ is a differentiable function with values in $\mathbb{F}^n$, the vector identity

$$\frac{d\mathbf{x}}{dt} = A \cdot \mathbf{x}(t) \qquad \mathbf{x}(0) = \mathbf{c} = (c_1, \ldots, c_n) \tag{21}$$

is a system of constant coefficient differential equations with initial conditions $x_k(0) = c_k$ for $k = 1, \ldots, n$. There is a unique solution for all $-\infty < t < \infty$, given by $\mathbf{x}(t) = e^{tA} \cdot \mathbf{c}$. This follows because the vector-valued map $\mathbf{x}(t)$ is infinitely differentiable, with

$$\frac{d}{dt}\Big(e^{tA}\Big) = A \cdot e^{tA} \quad \text{for all } t \in \mathbb{R}\ ,$$

from which we get

$$\frac{d\mathbf{x}}{dt} = \frac{d}{dt}\Big(e^{tA}\Big) \cdot \mathbf{c} = Ae^{tA} \cdot \mathbf{c} = A \cdot \mathbf{x}(t)\ .$$

Solving the differential equation (21) therefore reduces to computing e^{tA}, but how do you do that? As noted above, if A has a characteristic polynomial that splits over $\mathbb{F}$ (or if $\mathbb{F} = \mathbb{C}$), we can find a basis that puts A in Jordan canonical form. That means we can, with some effort, find a nonsingular $S \in \mathrm{M}(n, \mathbb{F})$ such that $B = SAS^{-1}$ consists of diagonal blocks

$$B = \begin{pmatrix} \boxed{B_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{B_r} \end{pmatrix}\ ,$$

each having the form

$$B_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \cdot & \cdot & \\ & & \cdot & \cdot \\ & & & \cdot \;\; 1 \\ 0 & & & \lambda_k \end{pmatrix} = \lambda_k I + N_k$$

where $\lambda_k \in \mathrm{sp}_{\mathbb{F}}(A)$, and N_k is either an elementary (cyclic) nilpotent matrix, or a 1×1 block consisting of the scalar λ_k, in which the elementary nilpotent part is degenerate. (Recall the discussion surrounding equation (12)). Obviously,

$$Se^{tA}S^{-1} = S\Big(\sum_{k=0}^{\infty} \frac{t^k}{k!}A^k\Big)S^{-1} = \sum_{k=0}^{\infty} \frac{t^k}{k!}SA^kS^{-1} = e^{t\,SAS^{-1}}\ ,$$

but $SA^kS^{-1} = \left(SAS^{-1}\right)^k$ for $k = 0, 1, 2 \ldots$, so

$$Se^{tA}S^{-1} = \sum_{k=0}^{\infty} \frac{t^k(SAS^{-1})^k}{k!} = \begin{pmatrix} e^{tB_1} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & e^{tB_r} \end{pmatrix} = e^{tB}$$

Here $e^{tB} = e^{t(\lambda I+N)} = e^{t\lambda I} \cdot e^{tN}$ because $e^{A+B} = e^A \cdot e^B$ when A and B commute, and then we have

$$e^{tB} = e^{\lambda t}I \cdot \Big[\sum_{j=0}^{d-1} \frac{t^j}{j!}N^j\Big] = e^{\lambda t}I \cdot \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{d-1}}{(d-1)!} \\ & \cdot & \cdot & & \vdots \\ & & \cdot & \cdot & \frac{t^2}{2!} \\ & & & 1 & t \\ 0 & & & & 1 \end{pmatrix}$$

of size $d \times d$. These matrices can be computed explicitly and so can the scalars $e^{t\lambda}$. Then we can recover e^{tA} by reversing the similarity transform to get

$$e^{tA} = S^{-1}e^{tB}S$$

which requires computing two products of explicit matrices. The solution of the original differential equation

$$\frac{d\mathbf{x}}{dt} = A\mathbf{x}(t) \quad \text{with inital condition} \quad \mathbf{x}(0) = \mathbf{c}$$

is then $\mathbf{x}(t) = e^{tA} \cdot \mathbf{c}$, as above.

7.1. Exercise. Use the Jordan Canonical form to find a list of solutions $A \in \mathrm{M}(2, \mathbb{C})$ to the matrix identity

$$A^2 + I = 0 \qquad \text{(the “square roots of } -I) ,$$

such that every solution is similar to one in your list.
Note: $SA^2S^{-1} = (SAS^{-1})^2$. $\square$

Chapter VIII. Complexification.

VIII-1. Analysis of Linear Operators over $\mathbb{R}$.

How can one analyze a linear operator $T : V \to V$ when the characteristic polynomial $p_T(x)$ does not split over $\mathbb{F}$? One approach is via the "*Rational Canonical form*," which makes no attempt to replace the ground field $\mathbb{F}$ with a larger field of scalars $\mathbb{K}$ over which p_T might split; we will not pursue this topic in these *Notes*. A different approach, which we will illustrate for $\mathbb{F} = \mathbb{R}$, is to enlarge $\mathbb{F}$ by constructing a field of scalars $\mathbb{K} \supseteq \mathbb{F}$; then we may in an obvious way regard $\mathbb{F}[x]$ as a subalgebra within $\mathbb{K}[x]$, and since $\mathbb{K} \supseteq \mathbb{F}$ there is a better chance that $f(x)$ will split into linear factors in $\mathbb{K}[x]$

$$f(x) = c \cdot \prod_{j=1}^{d} (x - \mu_j)^{m_j} \qquad \text{with } c \text{ and } \mu_j \text{ in } \mathbb{K}. \tag{22}$$

It is in fact always possible to embed $\mathbb{F}$ in a field $\mathbb{K}$ that is **algebraically closed**, which means that every polynomial $f(x) \in \mathbb{K}[x]$ splits into linear factors belonging to $\mathbb{K}[x]$, as in (22).

The Fundamental Theorem of Algebra asserts that the complex number field is algebraically closed; but the real number system $\mathbb{R}$ is not – for instance $x^2 + x + 2 \in \mathbb{R}[x]$ cannot split into linear factors in $\mathbb{R}[x]$ because it has no roots in $\mathbb{R}$. However, it does split when regarded as an element of $\mathbb{C}[x]$,

$$x^2 + x + 1 = (x - z_+) \cdot (x - z_-) - \big(x - \tfrac{1}{2}(-1 + i\sqrt{3})\big) \cdot \big(x + \tfrac{1}{2}(-1 - i\sqrt{3})\big)$$

where $i = \sqrt{-1}$. In this simple example one can find the complex roots $z_\pm = \frac{1}{2}(-1 \pm i\sqrt{3})$ using the quadratic formula.

Any real matrix $A \in \mathrm{M}(n, \mathbb{R})$ can be regarded as a matrix in $\mathrm{M}(n, \mathbb{C})$ whose entries happen to be real. Thus the operator

$$L_A(\mathbf{x}) = A \cdot \mathbf{x} \qquad \text{(matrix multiplication)}$$

acting on $n \times 1$ column vectors can be viewed as a linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$, but also as a "complexified" operator $T_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$ on the complexified space $\mathbb{C}^n = \mathbb{R}^n + i\mathbb{R}^n$. Writing vectors $\mathbf{z} = (z_1, \cdots, z_n)$ with complex entries $z_j = x_j + iy_j$ $(x_j, y_j \in \mathbb{R})$, we may regard $\mathbf{z}$ as a combination $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ with complex coefficients of the real vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$. The complexified operator $T_{\mathbb{C}} \in \mathrm{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n)$ can then be expressed in terms of the original $\mathbb{R}$-linear map $T : \mathbb{R}^n \to \mathbb{R}^n$,

$$T_{\mathbb{C}}(x + iy) = T(x) + iT(y),\ T_{\mathbb{C}} \in \mathrm{Hom}_{\mathbb{C}}(\mathbb{C}^n, \mathbb{C}^n) \qquad \text{for } x, y \in V. \tag{23}$$

The result is a $\mathbb{C}$-linear operator $T_{\mathbb{C}}$ whose characteristic polynomial $p_{T_{\mathbb{C}}}(t) \in \mathbb{C}[t]$ turns out to be the same as $p_T(t)$ when we view $p_T \in \mathbb{R}[t]$ as a polynomial in $\mathbb{C}[t]$ that happens to have real coefficients. Since $p_{T_{\mathbb{C}}}(t)$ always splits over $\mathbb{C}$, all the preceding theory applies to $T_{\mathbb{C}}$. Our task is then to translate that theory back to the original real linear operator $T : \mathbb{R}^n \to \mathbb{R}^n$.

1.1. Exercise. If T is a linear operator from $\mathbb{R}^n \to \mathbb{R}^n$ and $T_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$ is its complexification as defined in (23), verify that the characteristic polynomials $p_T(t) \in \mathbb{R}[t]$ and $p_{T_{\mathbb{C}}}(t) \in \mathbb{C}[t]$ are "the same" – i.e. that

$$p_{T_{\mathbb{C}}}(t) = p_T(t)$$

when we identify $\mathbb{R}[t] \subseteq \mathbb{C}[t]$.
Hint: Polynomials in $\mathbb{F}[t]$ are equal $\Leftrightarrow$ they have the same coefficients in $\mathbb{F}$. Here, $p_{T_{\mathbb{C}}}$ has coefficients in $\mathbb{C}$ while p_T has coefficients in $\mathbb{R}$, but we are identifying $\mathbb{R} = \mathbb{R} + i0 \subseteq \mathbb{C}$. $\square$

Relations between $\mathbb{R}[t]$ and $\mathbb{C}[t]$. When we view a real coefficient polynomial $f(t) = \sum_{j=0}^{m} a_j t^j \in \mathbb{R}[t]$ as an element of $\mathbb{C}[t]$ it can have complex roots as well as real roots, but

> *When we view $f(t) \in \mathbb{R}[x]$ as an element of $\mathbb{C}[x]$, any non-real roots must occur in conjugate pairs $z_{\pm} = (u \pm iv)$ with u and v real. Such eigenvalues can have nontrivial multiplicities, resulting in factors $(t - z_+)^m \cdot (t - z_-)^m$ in the irreducible factorization of $f(t)$ in $\mathbb{C}[t]$.*

In fact, if $z = x + iy$ with x, y real and if $f(z) = 0$, the complex conjugate $\overline{z} = x - iy$ is also a root of f because

$$f(\overline{z}) = \sum_{j=0}^{m} a_j \overline{z}^j = \sum_{j=0}^{m} a_j \overline{z^j} = \overline{\sum_j a_j z^j} = \overline{f(z)} = 0$$

(Recall that $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z} \cdot \overline{w}$, and $(\overline{z})^- = z$ for $z, w \in \mathbb{C}$.)

Thus, $\#$(*non-real roots*) is even, if any exist, while the number of real roots is unrestricted, and might be zero. Thus the splitting of f in $\mathbb{C}[t]$ can written as

$$\begin{aligned} f(t) &= c \cdot \prod_{j=1}^{r} (t - \mu_j)^{m_j} (t - \overline{\mu_j})^{m_j} \cdot \prod_{k=r+1}^{s} (t - r_j)^{m_k} \\ &= c \cdot \prod_{j=1}^{r} \big[\, (t - \mu_j)(t - \overline{\mu_j}) \,\big]^{m_j} \cdot \prod_{k=r+1}^{s} (t - r_j)^{m_k} \end{aligned} \tag{24}$$

where the μ_j are complex and nonreal ($\overline{\mu} \neq \mu$), and the r_j are the distinct real roots of f. Obviously $n = \deg(f) = \sum_{j=1}^{r} 2m_j + \sum_{j=r+1}^{s} m_j$. Since f has real coefficients, all complex numbers must disappear when the previous equation is multiplied out. In particular, for each nonreal conjugate pair $\mu, \overline{\mu}$ we have

$$Q_{\mu}(t) = (t - \mu)(t - \overline{\mu}) = t^2 - 2\mathrm{Re}(\mu) + |\mu|^2 \ , \tag{25}$$

a quadratic with real coefficients. Hence,

$$f(t) = c \cdot \prod_{r+1}^{s} (t - r_j)^{m_j} \cdot \prod_{j=1}^{r} (Q_{\mu_j}(t))^{m_j}$$

is a factorization of $f(t)$ into linear and *irreducible* quadratic factors in $\mathbb{R}[t]$, and every $f \in \mathbb{R}[t]$ can be decomposed this way. This is the (unique) decomposition of $f(t)$ into irreducible factors in $\mathbb{R}[t]$: by definition, the $Q_{\mu}(t)$ have no real roots and cannot be a product of polynomials of lower degree in $\mathbb{R}[t]$, while the linear factors are irreducible as they stand.

1.2. Definition. *A nonconstant polynomial $f \in \mathbb{F}[t]$ is* **irreducible** *if it cannot be factored as $f(t) = g(t)h(t)$ with g, h nonconstant and of lower degree than f. A polynomial is* **monic** *if its leading coefficient is* 1*. It is well known that every monic $f \in \mathbb{F}[t]$ factors uniquely as $\prod_{j=1}^{r} h_j(t)^{m_j}$ where each h_j monic and irreducible in $\mathbb{F}[t]$. The exponent $m_j \geq 1$ is its multiplicity, and this factorization is unique.*

The simplest irreducibles (over any $\mathbb{F}$) are the linear polynomials $at + b$ (with $a \neq 0$ since "irreducibility" only applies to nonconstant polynomials). This follows from the *degree formula*

Degree Formula: $\quad \deg(gh) = \deg(g) + \deg(h) \quad$ for all nonzero $g, h \in \mathbb{F}[t]$.

If we could factor $at+b = g(t)h(t)$, either $g(t)$ or $h(t)$ would have degree 0, and the other would have degree $1 = \deg(at+b)$. Thus there are no nontrivial factorization of $at+b$. When $\mathbb{F} = \mathbb{C}$, *all* irreducible polynomials have *degree* $= 1$, but if $\mathbb{F} = \mathbb{R}$ they can have *degree* $= 1$ or 2.

1.3. Lemma. *The irreducible monic polynomials in* $\mathbb{R}[t]$ *have the form*

1. $(t - r)$ *with* $r \in \mathbb{R}$*, or*
2. $t^2 + bt + c$ *with* $b^2 - 4c < 0$*. These are precisely the polynomials* $(t-\mu)(t-\overline{\mu})$ *with* μ *a non-real element in* $\mathbb{C}$*.*

Proof: Linear polynomials $at + b\,(a \neq 0)$ in $\mathbb{R}[t]$ are obviously irreducible. If f has the form of (2.), the quadratic formula applied on be applied to $f(t) = t^2 + bt + c$ in $\mathbb{C}[x]$ to find its roots

$$\mu, \overline{\mu} = \frac{-b \pm \sqrt{b^2 - 4c}}{2} = \frac{-b \pm i\sqrt{4c - b^2}}{2}$$

There are three possible outcomes:

1. $f(x)$ has a single real root with multiplicity $m = 2$ when $b^2 - 4c = 0$ and then we have $f(t) = (t - \frac{1}{2}b)^2$;
2. There are two distinct real roots $r_\pm = \frac{1}{2}\big(-b \pm \sqrt{b^2 - 4c}\big)$ when $b^2 - 4c > 0$, and then $f(t) = (t - r_+)(t - r_-)$;
3. When the discriminant $b^2 - 4c$ is negative there are two distinct conjugate nonreal roots in $\mathbb{C}$,

   $$\mu = \frac{-b + \sqrt{b^2 - 4c}}{2} \quad \text{and} \quad \overline{\mu} = \frac{-b - i\sqrt{4c - b^2}}{2} \qquad (i = \sqrt{-1}),$$

   in which case $f(t) = (t - \mu)(t - \overline{\mu})$ has the form (25) in $\mathbb{R}[t]$.

The quadratic $f(t)$ is irreducible in $\mathbb{R}[t]$ when $f(t)$ has two nonreal roots; otherwise it would have a factorization $(t - r_1)(t - r_2)$ in $\mathbb{R}[t]$ and also in $\mathbb{C}[t]$. That would disagree with $(x - \mu)(x - \overline{\mu})$, contrary to unique factorization in $\mathbb{C}[t]$, and cannot occur. $\square$

Complexification of Arbitrary Linear Operators over $\mathbb{R}$. We now discuss complexification of *arbitrary* vector spaces over $\mathbb{R}$ and complexifications $T_\mathbb{C}$ of the $\mathbb{R}$-linear operators $T : V \to V$ that act on them.

1.4. Definition. (Complexification). *Given an arbitrary vector space* V *over* $\mathbb{R}$ *its* **complexification** $V_\mathbb{C}$ *is the set of symbols* $\{\mathbf{z} = x + iy : x, y \in V\}$ *equipped with operations*

$$\begin{aligned} \mathbf{z} + \mathbf{w} \;=\; (x + iy) + (u + iv) &\;=\; (x + u) + i(y + v) \\ (a + ib)\cdot\mathbf{z} \;=\; (a + ib)(x + iy) &\;=\; (ax - by) + i(bx + ay), \text{ for } a + ib \in \mathbb{C} \end{aligned}$$

Two symbols $\mathbf{z} = (x + iy)$ *and* $\mathbf{z}' = (x' + iy')$ *designate the same element of* $V_\mathbb{C} \Leftrightarrow x' = x$ *and* $y' = y$*.*

1. *The* **real points** *in* $V_\mathbb{C}$ *are those of the form* $V + i0$*. This set is a vector space over* $\mathbb{R}$ *(but* NOT *over* $\mathbb{C}$*), because*

   $$\begin{aligned} (c + i0)\cdot(x + i0) &= (cx) + i0 && \text{for } c \in \mathbb{R}, x \in V \\ (x + i0) + (x' + i0) &= (x + x') + i0 && \text{for } x, x' \in V. \end{aligned}$$

   Clearly the operations $(+)$ *and (scale by real scalar c+i0) match the usual operations in* V *when restricted to* $V + i0$*.*

2. *If $T : V \to V$ is an $\mathbb{R}$-linear operator its* **complexification** $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ *is defined to be the map*

$$T_{\mathbb{C}}(x+iy) = T(x) + iT(y), \;\; x, y \in V \;\; , \tag{26}$$

which turns out to be a $\mathbb{C}$-linear operator on $V_{\mathbb{C}}$.

We indicate this by writing "$T_{\mathbb{C}} = T + iT$."

1.5. Exercise. If $M_1, \dots, M_r$ are vector spaces over $\mathbb{R}$ prove that the complexification of $V = M_1 \oplus \dots \oplus M_r$ is $V_{\mathbb{C}} = (M_1)_{\mathbb{C}} \oplus \dots \oplus (M_r)_{\mathbb{C}}$. $\square$

1.6. Exercise. Prove that $V_{\mathbb{C}}$ is actually a vector space over $\mathbb{C}$. (Check each of the vector space Axioms.)
Hint: In particular, you must check that $(z_1 z_2) \cdot w = z_1 \cdot (z_2 \cdot w)$, for $z_1, z_2 \in \mathbb{C}$ and $w \in V_{\mathbb{C}}$, and that $(c + i0) scdot (x + i0) = (c \cdot x) + i0$ for $c \in \mathbb{R}$, so $V + i0$ is a subspace over $\mathbb{R}$ isomorphic to V. $\square$

1.7. Example. We verify that

1. $T_{\mathbb{C}}$ is in fact a $\mathbb{C}$-linear operator on the complex vector space $V_{\mathbb{C}}$.
2. When we identify V with the subset $V + i0$ in $V_{\mathbb{C}}$ via the map $j : v \mapsto v + i0$, the restriction $T_{\mathbb{C}}|_{V+i0}$ gets identified with the original operator T in the sense that

$$T_{\mathbb{C}}(v + i0) = T(v) + i \cdot 0 \quad \text{ for all } v \in V \; .$$

Thus the folowing diagram is commutative, with $T_{\mathbb{C}} \circ j = j \circ T$

$$\begin{array}{ccc} V & \xrightarrow{j} & V + i0 \subseteq V_{\mathbb{C}} \\ T \downarrow & & \downarrow T_{\mathbb{C}} \\ V & \xrightarrow{j} & V + i0 \subseteq V_{\mathbb{C}} \end{array} \qquad \text{with } T_{\mathbb{C}} \circ j = j \circ T \; .$$

Discussion: Commutativity of the diagram is immediate from the definitions of $V_{\mathbb{C}}$ and $T_{\mathbb{C}}$. The messy part of proving (1.) is showing that $T_{\mathbb{C}}(z \cdot \mathbf{w}) = z \cdot T_{\mathbb{C}}(\mathbf{w})$ for $z \in \mathbb{C}, \mathbf{w} \in V_{\mathbb{C}}$, so we will only do that. If $z = a + ib \in \mathbb{C}$ and $\mathbf{w} = u + iv$ in $V_{\mathbb{C}}$ we get

$$\begin{aligned} T_{\mathbb{C}}\big((a+ib)(u+iv)\big), &= T_{\mathbb{C}}\big((au - bv) + i(bu + av)\big) \\ &= T(au - bv) + iT(bu + av) \;=\; aT(u) - bT(v) + ibT(u) + iaT(v) \\ &= (a + ib) \cdot (T(u) + iT(v)) \;=\; (a + ib) \cdot T_{\mathbb{C}}(u + iv) \quad \square \end{aligned}$$

1.8. Example. If $\{e_j\}$ is an $\mathbb{R}$-basis in V, then $\{\tilde{e_j} = e_j + i0\}$ is a $\mathbb{C}$-basis in $V_{\mathbb{C}}$. In particular, $\dim_{\mathbb{R}}(V) = \dim_{\mathbb{C}}(V_{\mathbb{C}})$.

Discussion: If $w = v + iw$ $(v, w \in V)$ there are real coefficients $\{c_j\}, \{d_j\}$ such that

$$w = \Big(\sum_j c_j e_j\Big) + i\Big(\sum_j d_j e_j\Big) = \sum_j (c_j + id_j)(e_j + i0) \; ,$$

so the $\{\tilde{e_j}\}$ span $V_{\mathbb{C}}$. As for independence, if we have

$$0 + i0 = \sum z_j \tilde{e_j} = \sum (c_j + id_j) \cdot (e_j + i0) = \Big(\sum_j c_j e_j\Big) + i\Big(\sum_j d_j e_j\Big)$$

in $V_{\mathbb{C}}$ for coefficients $z_j = c_j + id_j$ in $\mathbb{C}$, then $\sum_j c_j e_j = 0 = \sum_j d_j e_j$, which implies $c_j = 0$ and $d_j = 0$ because $\{e_j\}$ is a basis in V. Thus $z_j = 0$ for all j. $\square$

1.9. Example. If $V = \mathbb{R}^n$ then $V_{\mathbb{C}} = \mathbb{R}^n + i\mathbb{R}^n$ is, in a obvious sense, the same as $\mathbb{C}^n$.

If $A \in \mathrm{M}(n, \mathbb{R})$, we get a $\mathbb{R}$-linear operator $T = L_A$ that maps $v \to A \cdot v$ (matrix product of $n \times n$ times $n \times 1$), whose matrix with respect to the standard basis $\mathfrak{X} = \{e_j\}$ in $\mathbb{R}^n$ is $[T]_{\mathfrak{X}} = A$. If $\{\tilde{e_j} = e_j + i0\} = \mathfrak{Y}$ is the corresponding basis in $V_{\mathbb{C}}$, it is easy to check that we again have $[T_{\mathbb{C}}]_{\mathfrak{Y}} = A$ – i.e. $T_{\mathbb{C}}$ is obtained by letting the matrix A with real entries act on complex column vectors by matrix multiply (regarding A as a complex matrix that happens to have real entries). $\square$

1.10. Definition (Conjugation). *The* **conjugation operator** $J : V_{\mathbb{C}} \to V_{\mathbb{C}}$ *maps* $x + iy \to x - iy$. *It is an* $\mathbb{R}$*-linear operator on* $V_{\mathbb{C}}$*, with*

$$J(c \cdot w) = c \cdot J(w) \qquad \text{if } c = c + i0 \in \mathbb{R} \ ,$$

but is **conjugate linear** *over* $\mathbb{C}$*, with*

$$\begin{aligned} J(z \cdot w) &= \overline{z} \cdot J(w) \quad \text{for } z \in \mathbb{C}, w \in V_{\mathbb{C}} \\ J(w_1 + w_2) &= J(w_1) + J(w_2) \end{aligned}$$

Further properties of conjugation are easily verified from this definition:

(27)
1. $J^2 = J \circ J = \mathrm{id}$, so $J^{-1} = J$.
2. $w \in V_C$ is a real point if and only if $J(w) = w$.
3. $\dfrac{w + J(w)}{2} = x + i0$ and $\dfrac{w - J(w)}{2i} = y + i0$, if $w = x + iy$ in $V_{\mathbb{C}}$.

The operator J can be used to identity the $\mathbb{C}$-linear maps $S : V_{\mathbb{C}} \to V_{\mathbb{C}}$, of **real type**, those such that $S = T_{\mathbb{C}} = T + iT$ for some $\mathbb{R}$-linear $T : V \to V$.

1.11. Exercise. Whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, a matrix in $\mathrm{M}(n, \mathbb{F})$ determines a linear operator $L_A : \mathbb{F}^n \to \mathbb{F}^n$. Verify the following relationships between operators on $V = \mathbb{R}^n$ and $V_{\mathbb{C}} = \mathbb{C}^n = \mathbb{R}^n + i\mathbb{R}^n$.

1. If $\mathbb{F} = \mathbb{R}$, $\big(L_A\big)_{\mathbb{C}} = L_A + iL_A : \mathbb{C}^n \to \mathbb{C}^n$ is the same as the operator $L_A : \mathbb{C}^n \to \mathbb{C}^n$ we get by regarding A as a complex matrix all of whose entries are real.
2. Consider $A \in \mathrm{M}(n, \mathbb{C})$ and regard $\mathbb{C}^n$ as the complexification of $\mathbb{R}^n$. Verify that $L_A : \mathbb{C}^n \to \mathbb{C}^n$ is of *real type* $\Leftrightarrow$ all entries in A are real, so $A \in \mathrm{M}(n, \mathbb{R})$.
3. If S and T are $\mathbb{R}$-linear operators on a real vector space V, is the map
$$(S + iT) : (x + iy) \to S(x) + iT(y)$$
on $V_{\mathbb{C}}$ a $\mathbb{C}$-linear operator? If so, when is it of real-type? $\square$

1.12. Exercise. If $T : V \to V$ is an $\mathbb{R}$-linear operator on a real vector space, prove that

1. $\big(T_{\mathbb{C}}\big)^k = \big(T^k\big)_{\mathbb{C}}$ for all $k \in \mathbb{N}$
2. $e^{(T_{\mathbb{C}})} = \big(e^T\big)_{\mathbb{C}}$ on $V_{\mathbb{C}}$ $\square$

1.13. Lemma. *If* $T : V \to V$ *is* $\mathbb{R}$*-linear and* $T_{\mathbb{C}}$ *is its complexification, then* $T_{\mathbb{C}}$ *commutes with* J*,*

$$JT_{\mathbb{C}} = T_{\mathbb{C}}J \quad (\text{or equivalently } JT_{\mathbb{C}}J = T_{\mathbb{C}}).$$

Conversely, if $S : V_{\mathbb{C}} \to V_{\mathbb{C}}$ *is any* $\mathbb{C}$*-linear operator the following statements are equivalent*

1. $S = T_{\mathbb{C}} = T + iT$ *for some real-linear* $T : V \to V$.

2. $SJ = JS$.

3. *S leaves invariant the $\mathbb{R}$-subspace of real vectors $V + i0$ in $V_{\mathbb{C}}$.*

Proof: For (1.) $\Rightarrow$ (2.) is trivial:

$$\begin{array}{rcl} JT_{\mathbb{C}}J(x+iy) &=& JT_{\mathbb{C}}(x-iy) \;=\; J(T(x)+iT(-y)) \;=\; J(T(x)-iT(y)) \\ &=& T(x)+iT(y) \;=\; T_{\mathbb{C}}(x+iy) \end{array}$$

For (2.) $\Rightarrow$ (3.), suppose $SJ = JS$. We have $w \in V + i0$ if and only if $w = \frac{1}{2}\big(w + J(w)\big)$, and for these w we have

$$S(w) = \tfrac{1}{2}\big(S(w) + S(J(w)\big) = \tfrac{1}{2}\big(J(S(w)) + S(w)\big)\ .$$

By the properties (27), $S(w)$ is a vector in $V + i0$.

For (3.) $\Rightarrow$ (1.), if S leaves $V + i0$ invariant then $S(x + i0) = T(x) + i0$ for some uniquely determined vector $T(x) \in V$. We claim that $T : x \mapsto T(x)$ is an $\mathbb{R}$-linear map. In fact, if $c_1, c_2 \in \mathbb{R}$ and $v_1, v_2 \in V$, we have

$$S\big((c_1x_1 + c_2x_2) + i0\big) = T(c_1x_1 + c_2x_2) + i0\ ,$$

while S (being $\mathbb{C}$-linear) must also satisfy the identities

$$\begin{array}{rcl} S((c_1x_1+c_2x_2)+i0) &=& S((c_1x_1+i0)+(c_2x_2+i0)) \;=\; S(c_1x_1+i0)+S(c_2x_2+i0) \\ &=& S((c_1+i0)\cdot(x_1+i0)+(c_2+i0)\cdot(x_2+i0)) \\ &=& (c_1+i0)\cdot(Tx+i0)+(c_2+i0)\cdot(T(x_2)+i0) \\ &=& \big(c_1T(x_1)+c_2T(x_2)\big)+i0 \end{array}$$

Thus T is $\mathbb{R}$-linear on V. Furthermore, $S = T_{\mathbb{C}}$ because

$$\begin{array}{rcll} T_{\mathbb{C}}(x+iy) &=& T(x)+iT(y) = (T(x)+i0)+i(T(y)+i0) & \\ &=& S(x+i0)+iS(y+i0) & \text{(by $\mathbb{C}$-linearity of S and definition of T)} \\ &=& S\big((x+i0)+i(y+i0)\big) \;=\; S(x+iy) & \end{array}$$

Thus $S : V_{\mathbb{C}} \to \mathbb{C}$ is of real type if and only if $JS = SJ$, and then $S = \big(S|_{V+i0)}\big)_{\mathbb{C}}$. $\square$

An Application. The complexified operator $T_{\mathbb{C}}$ acts on a complex vector space $V_{\mathbb{C}}$ and therefore can be put into Jordan form (or perhaps diagonalized) by methods worked out previously. We now examine the special case when $T_{\mathbb{C}}$ is diagonalizable, before taking up the general problem: If $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$ is diagonalizable, what can be said about the structure of the original $\mathbb{R}$-linear operator $T : V \to V$?

We start with an observation that holds for any $\mathbb{R}$-linear operator $T : V \to V$, whether or not $T_{\mathbb{C}}$ is diagonalizable.

1.14. Lemma. *If $p_T(t) = \sum_{j=0}^m a_j t^j$ ($a_j \in \mathbb{R}$), then $p_{T_{\mathbb{C}}} = p_T$ in the sense that $p_{T_{\mathbb{C}}}(t) = \sum_{j=0}^m (a_j + i0)t^j$ in $\mathbb{C}[t] \supseteq \mathbb{R}[t]$.*

Proof: Earlier we proved that if $\mathfrak{X} = \{e_j\}$ is an $\mathbb{R}$-basis in V then $\mathfrak{Y} = \{\tilde{e_j} = e_j + i0\}$ is a $\mathbb{C}$-basis in $V_{\mathbb{C}}$, and that $[T]_{\mathfrak{X}} = [T_{\mathbb{C}}]_{\mathfrak{Y}}$ because

$$\begin{array}{rcl} T_{\mathbb{C}}(\tilde{e_j}) &=& T_{\mathbb{C}}(e_j+i0) \;=\; T(e_j)+iT(0) \;=\; \big(\sum_k t_{kj}\cdot e_k\big)+i0 \\ &=& \sum_k (t_{kj}+i0)(e_k+i0) \;=\; \sum t_{kj}\tilde{e_j} \end{array}$$

Thus $[T_{\mathbb{C}}]_{ij} = t_{ij} = [T]_{ij}$. Subtracting tI and taking the determinant, the outcome is the same whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. $\Box$

Hereafter we write p_T for $p_{T_{\mathbb{C}}}$ leaving the context to determine whether p_T is to be regarded as an element of $\mathbb{R}[x]$ or $\mathbb{C}[x]$. As noted earlier, p_T always splits in $\mathbb{C}[x]$, but might not split as an element of $\mathbb{R}[x]$. Furthermore, the nonreal roots in the factorization (24) come in conjugate pairs, so we may list the eigenvalues of $T_{\mathbb{C}}$ as follows, selecting one representative μ from each conjugate pair $\mu, \overline{\mu}$

$$\mu_1, \overline{\mu_1}, \cdots, \mu_r, \overline{\mu_r}; \lambda_{r+1}, \cdots, \lambda_s, \qquad \text{with } \lambda_i \text{ real and } \mu_j \neq \overline{\mu_j} \ . \tag{28}$$

and repeating eigenvalues/pairs according to their multiplicity in $p_{T_{\mathbb{C}}}(t)$.

Now assume $T_{\mathbb{C}}$ is diagonalizable, with (complex) eigenspaces $E_{\lambda_i}, E_{\mu_j}, E_{\overline{\mu_j}}$ in $V_{\mathbb{C}}$ that yield a direct sum decomposition of $V_{\mathbb{C}}$. Now observe that if $\mu \neq \overline{\mu}$ then $w \in V_{\mathbb{C}}$ is an eigenvector for μ if and only if $J(w)$ is an eigenvector for $\overline{\mu}$ because

$$T_{\mathbb{C}}(J(w)) = J\big(T_{\mathbb{C}}(w)\big) = J(\mu w) = \overline{\mu} J(w) \tag{29}$$

Hence, $J(E_\mu(T_{\mathbb{C}})) = E_{\overline{\mu}}(T_{\mathbb{C}})$ and J is an $\mathbb{R}$-linear bijection between $E_\mu(T_{\mathbb{C}})$ and $E_{\overline{\mu}}(T_{\mathbb{C}})$. Observe that $J(E_\mu \oplus E_{\overline{\mu}}) = E_\mu \oplus E_{\overline{\mu}}$ even though neither summand need be J-invariant (although we do have $J(E_\lambda) = E_\lambda$ when λ is a real eigenvalue for $T_{\mathbb{C}}$). The complex subspaces $E_\mu \oplus E_{\overline{\mu}}$ are of a special "real type "in $V_{\mathbb{C}}$ owing to their conjugation-invariance.

1.15. Definition. *If W is a $\mathbb{C}$-subspace of the complexification $V_{\mathbb{C}} = V + iV$, its* **real points** *are those in $W_{\mathbb{R}} = W \cap (V + i0)$. This is a vector space over $\mathbb{R}$ that determines a complex subspace $(W_{\mathbb{R}})_{\mathbb{C}} \subseteq W$ by taking $\mathbb{C}$-linear combinations.*

$$(W_{\mathbb{R}})_{\mathbb{C}} = W_{\mathbb{R}} + iW_{\mathbb{R}} \subseteq W$$

In general, $W_{\mathbb{R}} + iW_{\mathbb{R}}$ can be a lot smaller than the original complex subspace W. We say that a complex subspace $W \subseteq V_{\mathbb{C}}$ is of **real-type** *if*

$$W = W_{\mathbb{R}} + iW_{\mathbb{R}}$$

where $W_{\mathbb{R}} = W \cap (V + i0)$ is the set of real points in W.

Thus a complex subspace of real type W is the complexification of its subspace of real points $W_{\mathbb{R}}$.

Subspaces of real type are easily identified by their conjugation-invariance.

1.16. Lemma. *A complex subspace W in a complexification $V_{\mathbb{C}} = V + iV$ is of real type if and only if $J(W) = W$.*

Proof: $W = W_{\mathbb{R}} + iW_{\mathbb{R}}$ so $J(W) = W_{\mathbb{R}} - iW_{\mathbb{R}} = W_{\mathbb{R}} + iW_{\mathbb{R}}$ since $W_{\mathbb{R}} = -W_{\mathbb{R}}$, proving $(\Rightarrow)$. Conversely, for $(\Leftarrow)$: if $J(W) = W$ and we write $w \in W$ as $w = x + iy$ $(x, y \in V)$, both

$$\tfrac{1}{2}(w + J(w)) = x + i0, \quad \text{and} \quad \frac{1}{2i}(w - J(w)) = y - i0$$

are in $V + i0$, and both are in $W_{\mathbb{R}} = W \cap (V + i0)$. Since $w = (x + i0) + i(y + i0) = x + iy$, we conclude that $w \in W_{\mathbb{R}} + iW_{\mathbb{R}}$, so W is of real type. $\Box$

The spaces $W = E_\mu \oplus E_{\overline{\mu}}$ (and $W = E_\lambda$ for real λ) are obviously of real type since $J(E_\mu) = E_{\overline{\mu}}$. We now compare what is happening in a complex subspace W with what goes on in the real subspace $W_{\mathbb{R}}$. Note that $T_{\mathbb{C}}(W_{\mathbb{R}}) \subseteq W_{\mathbb{R}}$ because

$$\begin{aligned} T_{\mathbb{C}}(W_{\mathbb{R}}) &= T_{\mathbb{C}}(W \cap (V + i0)) \;=\; T_{\mathbb{C}}(W) \cap T_{\mathbb{C}}(V + i0) \\ &\subseteq W \cap (T(V) + i0) \;\subseteq\; W \cap (V + i0) \;=\; W_{\mathbb{R}} \end{aligned}$$

Case 1: $\lambda \in \mathrm{sp}(T_{\mathbb{C}})$ is real. Proposition 8.17 below shows that $E_\lambda(T_{\mathbb{C}}) = \big(E_\lambda(T)\big)_{\mathbb{C}}$ if $\lambda \in \mathbb{R}$. We have previously seen that an arbitrary $\mathbb{R}$-basis $\{f_j\}$ for $E_\lambda(T)$ corresponds to a $\mathbb{C}$-basis $\tilde{f}_j = f_j + i0$ in $\big(E_\lambda(T)\big)_{\mathbb{C}} = E_\lambda(T_{\mathbb{C}})$. But

$$T_{\mathbb{C}}(\tilde{f}_j) = \lambda \cdot \tilde{f}_j = \lambda \cdot (f_j + i0) = \lambda f_j + i0 \ ,$$

since $\tilde{f}_j \in E_\lambda(T_{\mathbb{C}})$, while

$$T_{\mathbb{C}}(\tilde{f}_j) = (T + iT)(f_j + i0) = T(f_j) + i0 \ .$$

Hence $T(f_j) = \lambda f_j$ and $T = T_{\mathbb{C}}|_{(E_\lambda(T)+i0)}$ is diagonalizable over $\mathbb{R}$.

1.17. Proposition. *If $T : V \to V$ is a linear operator on a real vector space and λ is a real eigenvalue for $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$, then $\lambda \in \mathrm{sp}_{\mathbb{R}}(T)$ and*

$$E_\lambda(T_{\mathbb{C}}) = E_\lambda(T) + iE_\lambda(T) = (E_\lambda(T))_{\mathbb{C}}$$

In particular, $\dim_{\mathbb{C}}(E_\lambda(T)) = \dim_{\mathbb{R}}(E_\lambda(T))$ *for real eigenvalues of T.*

Proof: $\lambda + i0 \in \mathrm{sp}_{\mathbb{C}}(T_{\mathbb{C}}) \cap (\mathbb{R} + i0)$ if and only if there is a vector $u + iv \neq 0$ in $V_{\mathbb{C}}$ such that $T_{\mathbb{C}}(u+iv) = (\lambda+i0)(u+iv) = \lambda u + i\lambda v$. But because $T_{\mathbb{C}}(u+iv) = T(u)+iT(v)$ this happens if and only if $T(v) = \lambda v$ and $T(u) = \lambda u$, and since at least one of the vectors $u, v \in V$ is nonzero we get $\lambda + i0 \in \mathrm{sp}_{\mathbb{C}}(T_{\mathbb{C}}) \cap (\mathbb{R} + i0) \subseteq \mathrm{sp}_{\mathbb{R}}(T)$.

Conversely, suppose $x + iy \in E_\lambda(T_{\mathbb{C}})$ for real λ. Then

$$T_{\mathbb{C}}(x + iy) = (\lambda + i0)(x + iy) \ = \ \lambda x + i\lambda y \ ,$$

but we also have

$$T_{\mathbb{C}}(x + iy) = T_{\mathbb{C}}(x + i0) + iT_{\mathbb{C}}(y + i0) \ = \ T(x) + iT(y)$$

because $T_{\mathbb{C}}$ is $\mathbb{C}$-linear. This holds if and only if $T(x) = \lambda x$ and $T(y) = \lambda y$, so that

$$x + iy \in E_\lambda(T) + iE_\lambda(T) = (E_\lambda(T))_{\mathbb{C}} \ . \quad \Box$$

1.18. Corollary. *We have* $\mathrm{sp}_{\mathbb{C}}(T_{\mathbb{C}}) \cap (\mathbb{R} + i0) = \mathrm{sp}_{\mathbb{R}}(T)$ *for any $\mathbb{R}$-linear operator $T : V \to V$ on a finite dimensional real vector space.*

1.19. Exercise. Let $V_{\mathbb{C}} = V + iV$ be the complexification of a real vector space V and let $S : V_{\mathbb{C}} \to V_{\mathbb{C}}$ be a $\mathbb{C}$-linear operator of real type, $S = T_{\mathbb{C}} = T + iT$ for some $T : V \to V$. Let $W = W_{\mathbb{R}} + iW_{\mathbb{R}} \subseteq V_{\mathbb{C}}$ be a complex subspace of real type that is S-invariant. Verify that

$$\text{(a)}\ \ S(W_{\mathbb{R}} + i0) \subseteq (W_{\mathbb{R}} + i0) \qquad \text{and} \qquad \text{(b)}\ \ S|_{(W_{\mathbb{R}}+i0)} = (T|_{W_{\mathbb{R}}}) + i0 \ . \quad \Box$$

This will be the key to determining structure of an $\mathbb{R}$-linear operator $T : V \to V$ from that of its complexification $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$.

Consider now the situation not covered by Case 1 above.

Case 2: Nonreal conjugate pairs $\mu, \overline{\mu}$. The space $E_{\mu,\overline{\mu}} = E_\mu(T_{\mathbb{C}}) \oplus E_{\overline{\mu}}(T_{\mathbb{C}})$ is of real type and $T_{\mathbb{C}}$-invariant; $T_{\mathbb{C}}$ is an operator of real type on $V_{\mathbb{C}}$ by definition. Let us list the pairs of non-real eigenvalues $\mu, \overline{\mu}$ according to their multiplicities as in (28), and let

$$f_j^{(\mu)} = x_j^\mu + iy_j^\mu \qquad \text{with } (x_j, y_j \in V)$$

be a $\mathbb{C}$-basis for $E_\mu(T_{\mathbb{C}})$, so that

$$E_\mu(T_{\mathbb{C}}) = \bigoplus_{j=1}^{d} \mathbb{C} f_j^{(\mu)} \qquad \text{and} \qquad T_{\mathbb{C}}|_{E_\mu} = \mu \cdot I_{E_\mu} \ .$$

Since $J(E_\mu) = E_{\overline{\mu}}$, we get a matching $\mathbb{C}$-basis in $E_{\overline{\mu}} = \bigoplus_{j=1}^{d} \mathbb{C}J(f_j^{(\mu)})$ using (29).

$$J(f_j^{(\mu)}) = x_j^\mu - i x_j^\mu$$

Then for $1 \le i \le d = \dim_{\mathbb{C}}(E_\mu)$ define the 2-dimensional complex subspaces in $E_{\mu,\overline{\mu}}$

$$V_j^\mu = \mathbb{C}f_j^{(\mu)} \oplus \mathbb{C}J(f_j^{(\mu)}),\ j = 1, 2, \cdots, d = \dim_{\mathbb{C}}(E_\mu(T_{\mathbb{C}}))$$

These are $T_{\mathbb{C}}$-invariant and are of real type since they are J-invariant by definition. Clearly $E_\mu \oplus E_{\overline{\mu}} = \bigoplus_{j=1}^{d} V_j^\mu$. We claim that for each V_j^μ, we can find a $\mathbb{C}$-basis consisting of two *real* vectors in $(V_j^\mu)_{\mathbb{R}} = \big(V_j^\mu\big) \cap (V + i0)$ (something that cannot be done for the spaces $\mathbb{C}f_j^{(\mu)}$ or E_μ alone).

We prove $W = W_{\mathbb{R}} + iW_{\mathbb{R}}$. If $f_j^{(\mu)} = x_j + iy_j$, $J(f_j^{(\mu)}) = x_j - iy_j$, then $x_j = x_j + i0$ and $y_j = y_j + i0$ are in $(V_j^\mu)_{\mathbb{R}}$ but their $\mathbb{C}$-span includes $f_j^{(\mu)}$ and $J\big(f_j^{(\mu)}\big)$, and is obviously all of $V_j^{(\mu)}$; these real vectors are a $\mathbb{C}$-basis for $V_j^{(\mu)}$. They are also an $\mathbb{R}$-basis for the 2-dimensional space $(V_j^\mu)_{\mathbb{R}} = (\mathbb{R}x_j + \mathbb{R}y_j) + i0$ of real points in $V_j^{(\mu)}$.

Note that $x_j + i0 \in (V_j^\mu)_{\mathbb{R}}$ can be written as

$$\begin{aligned} x_j + i0 &= \tfrac{1}{2}\big(f_j^{(\mu)} + J(f_j^{(\mu)})\big), \text{ and similarly} \\ y_j + i0 &= \frac{1}{2i}\big(f_j^{(\mu)} - J(f_j^{(\mu)})\big)\ . \end{aligned}$$

As previously noted, $T_{\mathbb{C}}$ (resp. T) leaves V_j^μ (resp. $(V_j^\mu)_{\mathbb{R}}$) invariant.

We now determine the matrix of $T_j = T|_{(V_j^\mu)_{\mathbb{R}}}$ with respect to the ordered $\mathbb{R}$-basis $\mathfrak{X}_j = \{x_j^{(\mu)}, y_j^{(\mu)}\}$. If $\mu = a + ib$ with a, b real and $b \ne 0$, then $\overline{\mu} = a - ib$; suppressing the superscript "μ" for clarity, we then have

$$\begin{aligned} T_{\mathbb{C}}(x_j + iy_j) &= \mu(x_j + iy_j) \;=\; (a+ib)(x_j + iy_j) \;=\; (ax_j - by_j) + i(ay_j + bx_j) \\ T_{\mathbb{C}}(x_j - iy_j) &= \bar{\mu}(x_j - iy_j) \;=\; (a+ib)(x_j - iy_j) \;=\; (ax_j - by_j) - i(ay_j + bx_j) \end{aligned}$$

Write μ in polar form

$$\mu, \overline{\mu} = a \pm ib = re^{\pm i\theta} = r\cos(\theta) \pm ir\sin(\theta)\ .$$

$T_{\mathbb{C}}$ and J commute because $T_{\mathbb{C}}$ is of real type $T_{\mathbb{C}}$, and since $J(zw) = \overline{z}J(w)$ for $z \in \mathbb{C}$ we get

$$\begin{aligned} T(x_j) + i0 &= T_{\mathbb{C}}(x_j + i0) \;=\; T_{\mathbb{C}}\Big(\frac{f_j + J(f_j)}{2}\Big) \;=\; \frac{T_{\mathbb{C}}(f_j) + J(T_{\mathbb{C}}(f_j))}{2} \\ &= \tfrac{1}{2}\big[\,(a+ib)f_j + J\big((a+ib)f_j\big)\big] \\ &= \tfrac{1}{2}\big[\,(a+ib)(x_j + iy_j) + (a-ib)\cdot(x_j - iy_j)\,\big] \\ &= (ax_j - by_j) \\ &= x_j \cdot r\cos(\theta) - y_j \cdot r\sin(\theta) \end{aligned}$$

Similarly, we obtain

$$T(y_j) + i0 = T_{\mathbb{C}}\Big(\frac{w_j - J(w_j)}{2i}\Big) = (ay_j + bx_j) = x_j \cdot r\sin(\theta) + y_j \cdot r\cos(\theta)$$

We previously proved that $[T_{\mathbb{C}}]_{\{\tilde{e_i}\}} = [T]_{\{e_i\}}$ for any $\mathbb{R}$-basis in $\big(V_j^{(\mu)}\big)_{\mathbb{R}}$, so the matrix of $T_j : (V_j^\mu)_{\mathbb{R}} \to (V_j^\mu)_{\mathbb{R}}$ with respect to the $\mathbb{R}$-basis $\mathfrak{X}_j = \{x_j, y_j\}$ in $(V_j^{(\mu)})_{\mathbb{R}}$ is

$$\big[T|_{(V_j^\mu)_{\mathbb{R}}}\big]_{\mathfrak{X}_j} = r \cdot \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$$

Reversing order of basis vectors yields basis $\mathfrak{X}'i_j = \{y_j^\mu, x_j^\mu\}$ such that the matrix of $T|_{(V_j^\mu)_\mathbb{R}}$ is a scalar multiple $r \cdot R(\theta)$ of the rotation matrix $R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ that corresponds to a counterclockwise rotation about the origin by θ radians, with $\theta \neq \pi, 0 \pmod{2\pi}$ because $\mu = a + ib$ is nonreal ($b \neq 0$).

Of course, with respect to the complex basis $\mathfrak{Y} = \{J(f_j^{(\mu)}), f_j^{(\mu)}\} = \{x_j - iy_j, x_j + iy_j\}$ in V_j^μ the operator $T_\mathbb{C}|_{V_j^\mu}$ is represented by the complex diagonal matrix

$$[T_\mathbb{C}|_{V_j^\mu}]_\mathfrak{Y} = \begin{pmatrix} r\,e^{-i\theta} & 0 \\ 0 & r\,e^{i\theta} \end{pmatrix} = \begin{pmatrix} \overline{\mu} & 0 \\ 0 & \mu \end{pmatrix}$$

To summarize: we have

$$\begin{aligned} V_\mathbb{C} &= \Big[\bigoplus_{\lambda\,\text{real}} \big(\bigoplus_{j=1}^{d(\lambda)} \mathbb{C} \cdot f_j^{(\lambda)} \big) \Big] \oplus \Big[\bigoplus_{\mu \neq \overline{\mu}\ \text{nonreal}} \big(\bigoplus_{j=1}^{d(\mu)} V_j^{(\mu)} \big) \Big] \\ &= \Big[\bigoplus_{\lambda\,\text{real}} \big(E_\lambda(T)\big)_\mathbb{C} \Big] \oplus \Big[\bigoplus_{\mu \neq \overline{\mu}\ \text{nonreal}} \big(\bigoplus_{j=1}^{d(\mu)} \big((V_j^{(\mu)})_\mathbb{R}\big)_\mathbb{C} \Big] \end{aligned}$$

where

$$V_j^{(\mu)} = \mathbb{C} f_j^{(\mu)} \oplus \mathbb{C} \cdot J(f_j^{(\mu)}) = \mathbb{C}(x_j^{(\mu)} + i0) \oplus \mathbb{C}(y_j^{(\mu)} + i0) \ ,$$

and all the spaces $\mathbb{C} f_j^{(\lambda)}$, $V_j^{(\mu)}$ are of real type. Restricting attention to the real points in $V_\mathbb{C}$ we arrive at a direct sum decomposition of the original real vector space V into T-invariant $\mathbb{R}$-subspaces

$$V = \Big[\bigoplus_{\lambda\,\text{real}} E_\lambda(T) \Big] \oplus \Big[\bigoplus_{\mu \neq \overline{\mu}\ \text{nonreal}} \big(\bigoplus_{j=1}^{d(\mu)} (V_j^{(\mu)})_\mathbb{R} \big) \Big] \tag{30}$$

We have arrived at the following decomposition of the $\mathbb{R}$-linear operator $T : V \to V$, when $T_\mathbb{C}$ is diagonalizable. Note: For each complex pair $(\mu, \overline{\mu})$ in $\mathrm{sp}(T_\mathbb{C})$, $E_\mu \oplus E_{\bar{\mu}}$ is of real type and we claim that $(E_\mu \oplus E_{\overline{\mu}}) \cap (V + i0) = \bigoplus_{j=1}^{d(\mu)} (V_j^\mu)_\mathbb{R}$. The sum on the right is direct so $\dim_\mathbb{R}(\bigoplus_{\mu,\overline{\mu}} \cdots) = 2d(\mu)$. Since we also have

$$\dim_\mathbb{R} \big(E_\mu \oplus E_{\overline{\mu}}\big)_\mathbb{R} = \dim_\mathbb{C} \big(E_\mu \oplus E_{\overline{\mu}}\big) = 2d(\mu) \ ,$$

the spaces coincide.

1.20. Theorem ($T_\mathbb{C}$ Diagonalizable). *If $T : V \to V$ is $\mathbb{R}$-linear and $T_\mathbb{C}$ is diagonalizable with eigenvalues labeled $\mu_1, \bar{\mu_1}, \cdots, \lambda_{r+1}, \cdots, \lambda_s$ as in (28), there is an $\mathbb{R}$-basis $\mathfrak{X}$ such that $[T]_\mathfrak{X}$ has the block diagonal form*

$$\begin{pmatrix} \boxed{R_1} & & & & & & 0 \\ & \cdot & & & & & \\ & & \cdot & & & & \\ & & & \boxed{R_r} & & & \\ & & & & \boxed{D_1} & & \\ & & & & & \cdot & \\ & & & & & & \cdot \\ 0 & & & & & & \boxed{D_s} \end{pmatrix}$$

where

$$R_k = \begin{pmatrix} r_k R(\theta_k) & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & r_k R(\theta_k) \end{pmatrix}$$

for $1 \leq k \leq r$, *and*

$$D_k = \begin{pmatrix} \lambda_k & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \lambda_k \end{pmatrix}$$

for $r+1 \leq k \leq s$.

Here, $\mu_k = r_k e^{i\theta_k}$ are representatives for the non-real pairs $(\mu, \overline{\mu})$ in $\mathrm{sp}(T_{\mathbb{C}})$.

When $T_{\mathbb{C}}$ is not diagonalizable, we apply the Jordan Canonical form for $T_{\mathbb{C}}$.

1.21. Lemma. *If* $T : V \to V$ *is a linear operator on a vector space over* $\mathbb{R}$, *let* $\mu \in \mathbb{C}$ *and* $M_\mu = \{w \in V_{\mathbb{C}} : (T_{\mathbb{C}} - \mu I)^k w = 0$ *for some* $k \in \mathbb{N}\}$. *Then* $w \in M_\mu \Leftrightarrow J(w) \in M_{\overline{\mu}}$, *so that*

$$J(M_\mu) = M_{\overline{\mu}} \quad \text{and} \quad J(M_{\overline{\mu}}) = M_\mu$$

Proof: The map $\Phi : S \mapsto JSJ = JSJ^{-1}$ is an *automorphism* of the algebra of all $\mathbb{C}$-linear maps $\mathrm{Hom}_{\mathbb{C}}(V_{\mathbb{C}}, V_{\mathbb{C}})$: it preserves products, $\Phi(S_1S_2) = \Phi(S_1)\Phi(S_2)$ because $J^2 = I \Rightarrow JS_1S_2J = (JS_1J)(JS_2J)$, and obviously $\Phi(S_1 + S_2) = \Phi(S_1) + \Phi(S_2)$, $\Phi(cS) = c\Phi(S)$ for $c \in \mathbb{C}$. In particular, $\Phi(S^k) = \Phi(S)^k$ for $k \in \mathbb{N}$. Then

$$J\big((T_{\mathbb{C}} - \mu I)^k\big)J = (J(T_{\mathbb{C}} - \mu I)J)^k = \big(JT_{\mathbb{C}}J - J(\mu I)J\big)^k = (T_{\mathbb{C}} - \overline{\mu}I)^k$$

($JT_{\mathbb{C}} = T_{\mathbb{C}}J$ because $T_{\mathbb{C}}$ is of real-type by definition), and

$$J(\mu I)J(w) = J(\mu \cdot J(w)) = \bar{\mu}J^2(w) = \bar{\mu}w, \ \ for \ w \in V_{\mathbb{C}}$$

Finally, we have $(T_{\mathbb{C}} - \mu I)^k w = 0$ if and only if

$$\begin{aligned} J(T_{\mathbb{C}} - \mu I)^k w = 0 \quad &\Leftrightarrow \quad J(T_{\mathbb{C}} - \mu I)J(J(w)) = 0 \\ &\Leftrightarrow \quad (T_{\mathbb{C}} - \overline{\mu})^k J(w) = 0 \Leftrightarrow J(w) \in M_{\overline{\mu}}\ . \end{aligned}$$

Hence $J(M_\mu) = M_{\overline{\mu}}$, which implies $M_\mu = J(M_{\overline{\mu}})$ since $J^2 = I$. $\square$

1.22. Theorem ($T_{\mathbb{C}}$ not Diagonalizable). *Let* $T : V \to V$ *lie a linear operator on a finite dimensional vector space over* $\mathbb{R}$ *and let* $\mu_1, \overline{\mu_1}, \cdots, \mu_r, \overline{\mu_r}, \lambda_{r+1}, \cdots, \lambda_s$ *be the eigenvalues of* $T_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$, *listed as in* (28). *Then there is an* $\mathbb{R}$-*basis for* V *that puts* $[T]_{\mathfrak{X}}$ *in block diagonal form:*

$$[T]_{\mathfrak{X}} = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \cdot & \\ & & \cdot \\ 0 & & \boxed{A_m} \end{pmatrix},$$

in which each block A_j *has one of two possible block upper-triangular forms:*

$$A = \begin{pmatrix} \lambda & 1 & & 0 \\ & \cdot & \ddots & \\ & & \cdot & 1 \\ 0 & & & \lambda \end{pmatrix} \qquad \text{for real eigenvalues } \lambda \text{ of } T_{\mathbb{C}}\ ,$$

or

$$A = \begin{pmatrix} \boxed{R_{\mu,\overline{\mu}}} & I_2 & & 0 \\ & \cdot & \ddots & \\ & & \cdot & I_2 \\ 0 & & & \boxed{R_{\mu,\overline{\mu}}} \end{pmatrix} \qquad \textit{for conjugate pairs } \mu = e^{i\theta}, \overline{\mu} = e^{-i\theta}$$

where I_2 *is the* 2×2 *identity matrix and*

$$R_{\mu,\overline{\mu}} = r \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} .$$

Proof: The proof will of course employ the generalized eigenspace decomposition $V_{\mathbb{C}} = \bigoplus_{z \in \mathrm{sp}(T_{\mathbb{C}})} M_z(T_{\mathbb{C}})$. As above, we write $M_{\mu,\overline{\mu}} = M_\mu \oplus M_{\overline{\mu}}$ for each distinct pair of nonreal eigenvalues. The summands $M_\lambda, M_{\mu,\overline{\mu}}$ are of real type and are $T_{\mathbb{C}}$ invariant, so for each μ we may write $M_{\mu,\overline{\mu}}$ as the complexification of its space of real points

$$W_\mu = M_{\mu,\overline{\mu}} \cap (V + i0) . \tag{31}$$

Here $M_{\mu,\overline{\mu}} = W_\mu + iW_\mu$ is $T_{\mathbb{C}}$-invariant. Since $T_{\mathbb{C}}$ is of real type (by definition), it leaves invariant the $\mathbb{R}$-subspace $V + i0$; therefore the space of real points W_μ is also $T_{\mathbb{C}}$ invariant, with $T_{\mathbb{C}}|_{(W_\mu + i0)} = T|_{W_\mu}$. It follows that $T_{\mathbb{C}}|_{M_{\mu,\overline{\mu}}}$ is the complexification

$$(T|_{W_\mu}) + i(T|_{W_\mu})$$

of T_{W_μ}. If we can find an $\mathbb{R}$-basis in the space (31) of real points $W_\mu \subseteq M_{\mu,\overline{\mu}}$ for which $T|_{W_\mu}$ takes the form described in the theorem, then our proof is complete. For this we may obviously restrict attention to a single subspace $M_{\mu,\overline{\mu}}$ in $V_{\mathbb{C}}$.

Case 1: A Real Eigenvalue λ. If λ is real then

$$\begin{aligned} (T_{\mathbb{C}} - \lambda)^k (x + iy) &= (T_{\mathbb{C}} - \lambda)^{k-1} \big[(T - \lambda)x + i(T - \lambda)y\big] = \cdots \\ &= (T - \lambda)^k x + i(T - \lambda)^k y \quad \text{for } k \in \mathbb{N} . \end{aligned}$$

Thus, $x + iy \in M_\lambda(T_{\mathbb{C}})$ if and only if $x, y \in M_\lambda(T)$, and the subspace of real points in $M_\lambda(T_{\mathbb{C}}) = M_\lambda(T) + iM_\lambda(T)$ is precisely $M_\lambda(T) + i0$. There is a $\mathbb{C}$-basis $\mathfrak{X} = \{f_j\}$ of real vectors in $M_\lambda(T_{\mathbb{C}})$ that yields the same matrix for $T_{\mathbb{C}}$ and for its restriction to this subspace of real points, and we have

$$T_{\mathbb{C}}|_{W_\lambda} = T|_{W_\lambda} = T|_{M_\lambda(T)} . \qquad \text{and} \qquad (T_{\mathbb{C}})|_{M_\lambda(T)} = T|_{M_\lambda(T)} .$$

Case 2: A Conjugate Pair $\mu, \overline{\mu}$. The space $M_{\mu,\overline{\mu}} = M_\mu(T_{\mathbb{C}}) \oplus M_{\overline{\mu}}(T_{\mathbb{C}})$ is of real type because $J(M_\mu) = M_{\overline{\mu}}$. In fact, if $v \in M_\mu$ then $(T_{\mathbb{C}} - \mu)^k v = 0$ for some k. But $T_{\mathbb{C}} J = J T_{\mathbb{C}}$, so

$$(T_{\mathbb{C}} - \overline{\mu})^k J(v) = (T_{\mathbb{C}} - \overline{\mu})^{k-1} J(T_{\mathbb{C}} - \mu)(v) = \cdots = J(T_{\mathbb{C}} - \mu)^k (v) = 0 .$$

Thus $M_{\mu,\overline{\mu}}$ is the complexification of its subspace of real points $V_\mu = M_{\mu,\overline{\mu}}(T_{\mathbb{C}}) \cap (V + i0)$.

By the Cyclic Subspace Decomposiiton (Theorem 3.2) a generalized eigenspace M_μ for $T_{\mathbb{C}}$ is a direct sum of $T_{\mathbb{C}}$-invariant spaces C_j that are cyclic under the action of the nilpotent operator $(T_{\mathbb{C}} - \mu I)$. In each C_j take a basis $\mathfrak{X}_j = \{f_1^{(j)}, \cdots, f_{d_j}^{(j)}\}$ that puts $(T_{\mathbb{C}} - \mu I)|_{C_j}$ into elementary nilpotent form

$$\begin{pmatrix} 0 & 1 & \cdot & 0 \\ & \cdot & \cdot & \cdot \\ & & \cdot & 1 \\ 0 & & & 0 \end{pmatrix} \qquad \text{so that} \qquad [T_{\mathbb{C}}] = \begin{pmatrix} \mu & 1 & \cdot & 0 \\ & \cdot & \cdot & \cdot \\ & & \cdot & 1 \\ 0 & & & \mu \end{pmatrix}$$

For the basis $\mathfrak{X}_j$ we have

$$(T_{\mathbb{C}} - \mu I)f_1 = 0 \qquad \text{and} \qquad (T_{\mathbb{C}} - \mu I)f_j = f_{j-1} \text{ for } j > 1,$$

which implies that $T_{\mathbb{C}}(f_1) = \mu f_1$ and $T_{\mathbb{C}}(f_j) = \mu f_j + f_{j-1}$ for $j > 1$.

If $f_j = x_j + iy_j \in W_\mu + iW_\mu$, we have $\bar{f}_j = J(f_j) = x_j - iy_j$ and $T_{\mathbb{C}}(f_j) = \mu f_j + f_{j-1}$, hence

$$T_{\mathbb{C}}(\bar{f}_j) = T_{\mathbb{C}}(J(f_j)) = J(T_{\mathbb{C}}(f_j)) = J(\mu f_j + f_{j-1}) = \overline{\mu}J(f_j) + J(f_{j-1}) \ .$$

Since real and imaginary parts must agree we get $T_{\mathbb{C}}(\bar{f}_j) = \overline{\mu}\bar{f}_j + \bar{f}_{j-1}$, as claimed.

Writing $\mu = a + ib$ with $b \neq 0$ (or in polar form, $\mu = re^{i\theta}$ with $\theta \notin \pi\mathbb{Z}$), we get

$$\begin{aligned} T(x_j) + iT(y_j) &= T_{\mathbb{C}}(x_j + iy_j) \ = \ T_{\mathbb{C}}(f_j) \ = \ \mu f_j + f_{j-1} \\ &= (a+ib)(x_j + iy_j) + (x_{j-1} + iy_{j-1}) \\ &= (ax_j - by_j) + i(bx_j + ay_j) + (x_{j-1} + iy_{j-1}) \end{aligned}$$

Since $\mu = re^{i\theta}$ and $\overline{\mu} = re^{-i\theta}$ this means

$$\begin{aligned} T(x_j) &= ax_j - by_j + x_{j-1} \ = \ x_j \cdot r\cos(\theta) - y_j \cdot r\sin(\theta) + x_{j-1} \\ T(y_j) &= bx_j + ay_j + y_{j-1} \ = \ x_j \cdot r\sin(\theta) + y_j \cdot r\cos(\theta) + y_{j-1} \end{aligned}$$

with respect to the $\mathbb{R}$-basis

$$\{x_1^{(1)}, y_1^{(1)}, \cdots, x_{d_1}^{(1)}, y_{d_1}^{(1)}, x_1^{(2)}; y_1^{(2)}, \cdots\}$$

in $V_\mu = M_{\mu,\overline{\mu}} \cap (V + i0) = \bigoplus_{j=1}^{d} V_J^{(\mu)}$. Thus the matrix $[T]_{\mathfrak{X}}$ consists of diagonal blocks of size $2d_j \times 2d_j$ that have the form

$$\begin{pmatrix} R & I_2 & \cdot & 0 \\ & R & I_2 & \\ & & \ddots & I_2 \\ 0 & & & R \end{pmatrix} = \begin{pmatrix} rR(\theta) & I_2 & \cdot & 0 \\ & rR(\theta) & I_2 & \\ & & \ddots & I_2 \\ 0 & & & rR(\theta) \end{pmatrix}$$

in which I_2 is the 2×2 identity matrix and

$$R = r \cdot R(\theta) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} r\ \cos(\theta) & -r\ \sin(\theta) \\ r\ \sin(\theta) & r\ \cos(\theta) \end{pmatrix}$$

if $\mu = a + ib = re^{i\theta}$. $\square$

Chapter IX. Bilinear and Multilinear Forms.

IX.1. Basic Definitions and Examples.

1.1. Definition. *A* **bilinear form** *is a map* $B : V \times V \to \mathbb{C}$ *that is linear in each entry when the other entry is held fixed, so that*

$$\begin{aligned} B(\alpha x, y) &= \alpha B(x,y) = B(x, \alpha y) \\ B(x_1 + x_2, y) &= B(x_1, y) + B(x_2, y) \qquad \text{for all } \alpha \in \mathbb{F},\ x_k \in V,\ y_k \in V \\ B(x, y_1 + y_2) &= B(x, y_1) + B(x, y_2) \end{aligned}$$

(*This of course forces* $B(x,y) = 0$ *if either input is zero.*) *We say* B *is* **symmetric** *if* $B(x,y) = B(y,x)$, *for all* x, y *and* **antisymmetric** *if* $B(x,y) = -B(y,x)$.

Similarly a **multilinear form** (*aka a* k-**linear form** , *or a* **tensor of rank** k) *is a map* $B : V \times \cdots \times V \to \mathbb{F}$ *that is linear in each entry when the other entries are held fixed. We write* $V^{(0,k)} = V^* \otimes \ldots \otimes V^*$ *for the set of* k*-linear forms. The reason we use* V^* *here rather than* V, *and the rationale for the "tensor product" notation, will gradually become clear.*

The set $V^* \otimes V^*$ of bilinear forms on V becomes a vector space over $\mathbb{F}$ if we define

1. Zero element: $B(x,y) = 0$ for all $x, y \in V$;
2. Scalar multiple: $(\alpha B)(x,y) = \alpha B(x,y)$, for $\alpha \in \mathbb{F}$ and $x, y \in V$;
3. Addition: $(B_1 + B_2)(x,y) = B_1(x,y) + B_2(x,y)$, for $x, y \in V$.

When $k > 2$, the space of k-linear forms $V^* \otimes \ldots \otimes V^*$ is also a vector space, using the same definitions. The space of 1-linear forms (= *tensors of rank* 1 on V) is the dual space $V^* = \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$ of all $\mathbb{F}$-linear maps $\ell : V \to \mathbb{F}$. By convention the space of **0-forms** is identified with the ground field: $V^{(0,0)} = \mathbb{F}$; its elements are not mappings on V. It is also possible (and useful) to define multilinear forms of mixed type, mappings $\theta : V_1 \times \ldots \times V_k \to \mathbb{F}$ in which the components V_j are not all the same. These forms also constitute a vector space. We postpone any discussion of forms of "mixed type."

If $\ell_1, \ell_2 \in V^*$ we can create a bilinear form $\ell_1 \otimes \ell_2$ by taking a "*tensor product*" of these forms

$$\ell_1 \otimes \ell_2(v_1, v_2) = \langle \ell_1, v_1 \rangle \cdot \langle \ell_2, v_2 \rangle \qquad \text{for } v_1, v_2 \in V$$

Bilinearity is easily checked. More generally, if $\ell_1, \cdots, \ell_k \in V^*$ we obtain a k-linear map from $V \times \ldots \times V \to \mathbb{F}$ if we let

$$\ell_1 \otimes \ldots \otimes \ell_k(v_1, \cdots, v_k) = \prod_{j=1}^{k} \langle \ell_j, v_j \rangle \quad .$$

We will show that "*monomials*" of the form $\ell_1 \otimes \ldots \otimes \ell_k$ span the space $V^{(0,k)}$ of rank-k tensors, but they do not by themselves form a vector space except when $k = 1$.

1.2. Exercise. If $A : V \to V$ is any linear operator on a real inner product space verify that

$$\phi(v_1, v_2) = (Av_1, v_2) \qquad \text{for } v_1, v_2 \in V$$

is a bilinear form.
Note: This would not be true if $\mathbb{F} = \mathbb{C}$. Inner products on a complex vector space are

conjugate-linear in their second input, with $(x, z \cdot y) = \overline{z} \cdot (x, y)$ for $z \in \mathbb{C}$; for $\mathbb{C}$-linearity in the second entry we would need $(x, z \cdot y) = z \cdot (x, y)$. However, $\overline{c} = c$ for real scalars so an inner product on a *real* vector space is a linear function of each input when the other is held fixed. $\square$

1.3. Example. Let $A \in \mathrm{M}(n, \mathbb{F})$ and $V = \mathbb{F}^n$. Regarding elements of $\mathbb{F}^n$ as $n \times 1$ column vectors, define

$$B(x, y) = x^{\mathrm{t}} A y = \sum_{ij=1}^{n} x_i A_{ij} y_j$$

where x^{t} is the $1 \times n$ transpose of the $n \times 1$ column vector x. If we interpret the 1×1 product as a scalar in $\mathbb{F}$, then B is a typical bilinear form on $V = \mathbb{F}^n$. $\square$

The analogous construction for multilinear forms is more complicated. For instance, to describe a rank-3 linear form $B(x, y, z)$ on $V \times V \times V$ we would need a three-dimensional $n \times n \times n$ array of coefficients $\{B_{i_1, i_2, i_3} : 1 \leq i_k \leq n\}$, from which we recover the original multilinear form via

$$B(x, y, z) = \sum_{i_1, i_2, i_3 = 1}^{n} x_{i_1} y_{i_2} z_{i_3} B_{i_1, i_2, i_3} \qquad \text{for } (x, y, z) \in \mathbb{F}^3 \ .$$

The coefficient array is an *ntimesn* square matrix only for bilinear form ($k = 2$). For the time being we will focus on bilinear forms, which are quite important in their own right.

Many examples involve symmetric or antisymmetric bilinear forms, and in any case we have the following result.

1.4. Lemma. *Every bilinear form B is uniquely the sum $B = B_+ + B_-$ of a symmetric and antisymmetric form.*

Proof: $B_\pm$ are given by

$$B_+(v_1, v_2) = \frac{B(v_1, v_2) + B(v_2, v_1)}{2} \qquad \text{and} \qquad B_- = \frac{B(v_1, v_2) - B(v_2, v_1)}{2} \ .$$

As for uniqueness, you can't have $B = B'$ with B symmetric and B' antisymmetric without both being the zero form. $\square$

Variants. If V is a vector space over $\mathbb{C}$, a map $B : V \times V \to \mathbb{C}$ is **sesquilinear** if it is a linear function of its first entry when the other is held fixed, but is conjugate-linear in its second entry, so that

$$\begin{aligned} B(x_1 + x_2, y) = B(x_1, y) + B(x_2, y) \quad &\text{and} \quad B(x, y_1 + y_2) \ = \ B(x, y_1) + B(x, y_2) \\ B(\alpha x, y) = \alpha B(x, y) \quad &\text{and} \quad B(x, \alpha y) = B(x, y)\, \overline{\alpha} \ \text{ for all } \alpha \in \mathbb{C} \ . \end{aligned}$$

This is the same as bilinearity when $\mathbb{F} = \mathbb{R}$. The map is **Hermitian symmetric** if

$$B(y, x) = \overline{B(x, y)}$$

On a vector space over $\mathbb{R}$, an **inner product** is a special type of bilinear form, one that is *strictly positive definite* in the sense that

$$B(x, x) \geq 0 \text{ for all } x \in V \quad \text{and} \quad B(x, x) = \|x\|^2 = 0 \Rightarrow x = 0 \tag{32}$$

Over $\mathbb{C}$, an inner product is a map $B : V \times V \to \mathbb{C}$ that is sesquilinear, Hermitian symmetric, and satisfies the nondegeneracy condition (32).

A bilinear form $B \in V^* \otimes V^*$ is completely determined by its action on a basis $\mathfrak{X} = \{e_i\}$ via the matrix $[B]_{\mathfrak{X}} = [B_{ij}]$ with entries

$$B_{ij} = B(e_i, e_j) \quad \text{for } 1 \leq i, j \leq n$$

This matrix is symmetric/antisymmetric if and only if B has these properties. Given $[B]_{\mathfrak{X}}$ we recover B by writing $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$; then

$$\begin{aligned} B(x,y) &= B\big(\sum_i x_i e_i\,,\, \sum_j y_j e_j\big) \;=\; \sum_i x_i\, B\big(e_i, \sum_j y_j e_j\big) \\ &= \sum_{i,j} x_i B_{ij} y_j \;=\; [x]^{\mathrm{t}}_{\mathfrak{X}} [B]_{\mathfrak{X}} [y]_{\mathfrak{X}}\ , \end{aligned}$$

a 1×1 matrix regarded as an element of $\mathbb{F}$. Conversely, given a basis and a matrix $A \in \mathrm{M}(n, \mathbb{F})$ the previous equality determines a bilinear form B (symmetric if and only if $B = B^{\mathrm{t}}$ etc) such that $[B]_{\mathfrak{X}} = A$. Thus we have isomorphisms between vector spaces over $\mathbb{F}$:

1. The space of rank-2 tensors $V^{(0,2)} = V^* \otimes V^*$ is $\cong \mathrm{M}(n, \mathbb{F})$ via $B \to [B]_{\mathfrak{X}}$;
2. The space of *symmetric* bilinear forms is isomorphic to the space of symmetric matrices , etc.

We next produce a basis for $V^* \otimes V^*$ and determine its dimension.

1.5. Proposition. *If $\mathfrak{X} = \{e_i\}$ is a basis in a finite-dimensional vector space V, and $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis in V^* such that $\langle e_i^*, e_j\rangle = \delta_{ij}$, then the monomials $e_i^* \otimes e_j^*$ given by*

$$e_i^* \otimes e_j^*(v_1, v_2) = \langle e_i^*, v_1\rangle \cdot \langle e_j^*, v_2\rangle$$

are a bases on $V^* \otimes V^*$. Hence, $\dim(V^* \otimes V^*) = n^2$.

Proof: The monomials $e_i^* \otimes e_j^*$ span $V^* \otimes V^*$, for if B is any bilinear form and $B_{ij} = B(e_i, e_j)$, then $\tilde{B} = \sum_{i,j} B_{ij} e_i^* \otimes e_j^*$ has the same action on pairs $e_k, e_\ell \in V$ as the original tensor B.

$$\begin{aligned} \tilde{B}(e_k, e_l) &= \Big(\sum_{i,j} B_{ij} \cdot e_i^* \otimes e_j^*\Big)\langle e_k, e_\ell\rangle \;=\; \sum_{i,j} B_{ij} \langle e_i^*, e_k\rangle \cdot \langle e_j^*, e\ell\rangle \\ &= \sum_{i,j} B_{ij} \delta_{ik} \delta_{j\ell} = B_{k\ell} = B(e_k, e_\ell)\ , \end{aligned}$$

so $\tilde{B} = B \in \mathbb{F}\text{-span}\{e_i^* \otimes e_j^*\}$. As for linear independence, if $\tilde{B} = \sum_{i,j} b_{ij} e_i^* \otimes e_j^* = 0$ in $V^{(0,2)}$, then $\tilde{B}(x,y) = 0$ for all x, y, so $b_{k\ell} = \tilde{B}(e_k, e_\ell) = 0$ for $1 \leq k, \ell \leq n$. $\square$

A similar discussion shows that the space $V^{(0,r)}$ of rank-k tensors has dimension

$$\dim(V^{(0,r)}) = \dim(V^* \otimes \ldots \otimes V^*) = \dim(V)^r = n^r\ .$$

If $\mathfrak{X} = \{e_1, \ldots, e_n\}$ is a basis for V and $\{e_i^*\}$ is the dual basis in V^*, the monomials

$$e_{i_1}^* \otimes \ldots \otimes e_{i_r}^* \qquad 1 \leq i_1, \ldots, i_r \leq n$$

are a basis for $V^{(0,r)}$.

1.6. Theorem (Change of Basis) *Given $B \in V^* \otimes V^*$ and a basis $\mathfrak{X}$ in V, we describe B by its matrix via* (32). *If $\mathfrak{Y} = \{f_j\}$ is another basis, and if*

$$\mathrm{id}(f_j) = f_j = \sum_k s_{kj} e_k \quad \textit{for} \quad 1 \leq j \leq n\ , \tag{33}$$

then $S = [s_{ij}] = [\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}$ *is the transition matrix for basis vectors and we have*

$$\begin{aligned} \big([B]_{\mathfrak{Y}}\big)_{ij} &= B(f_i, f_j) \;=\; B\Big(\sum_{k,\ell} s_{ki} e_k\,,\, \sum_{\ell} s_{\ell j} e_\ell\Big) \\ &= \sum_{k,\ell} S_{ki} B_{k\ell} S_{\ell j} \;=\; \sum_{k,\ell} (S^{\mathrm{t}})_{ik} B_{k\ell} S_{\ell j} \\ &= (S^{\mathrm{t}}[B]_{\mathfrak{X}} S)_{ij} \end{aligned}$$

Note: We can also write this as $[B]_{\mathfrak{Y}} = P[B]_{\mathfrak{X}}P^{\mathrm{t}}$, taking $P = S^{\mathrm{t}} = [\mathrm{id}]^{\mathrm{t}}_{\mathfrak{X}\mathfrak{Y}}$. $\square$

Thus change of basis is effected by "*congruence*" of matrices $A \mapsto SAS^{\mathrm{t}}$, with $\det(S) \neq 0$. This differs considerably from the "*similarity transforms*" $A \mapsto SAS^{-1}$ that describe the effect of change of basis on the matrix of a linear operator $T : V \to V$. Notethet S^{t} is generally not equal to S^{-1}, so congruence and similarity are not the same thing. The difference between these concepts will emerge when we seek "normal forms" for various kinds of bilinear (or sesquilinear) forms.

1.7. Definition. *A bilinear form* B *is* **nondegenerate** *if*

$$B(v, V) = 0 \;\Rightarrow\; v = 0 \qquad \text{and} \qquad B(V, v) = 0 \;\Rightarrow\; v = 0$$

If B *is either symmetric or antisymmetric we only need the one-sided version. The* **radical** *of* B *is the subspace*

$$\mathrm{rad}(B) = \{v \in V : B(v, v') = 0 \text{ for all } v' \in V\}\ ,$$

which measures the degree of degeneracy of the form B *The* B-**orthocomplement** *of a subspace* $W \subseteq V$ *is defined to be*

$$W^{\perp,B} = \{v \in V : B(v, W) = (0)\}\ .$$

Obviously, $W^{\perp,B}$ is a subspace. When B is symmetric or antisymmetric the conditions $B(v, W) = 0$ and $B(W, v) = 0$ yield the same subspace $B^{\perp,B}$. Then nondegeneracy means that $V^{\perp,B} = \{0\}$, and in general $V^{\perp,B}$ is equal to the radical of B.

1.8. Exercise (Dimension Formula). If B is nondegenerate and either symmetric or antisymmetric, and if $W \subseteq V$ is a subspace, prove that

$$\dim(W) + \dim\big(W^{\perp,B}\big) = \dim(V) \qquad \square.$$

The notion of "nondegeneracy" is a little ambiguous when the bilinear form B is neither symmetric nor antisymmetric: Is there a difference between "right nondegenerate," in which $B(V, y) = 0 \Rightarrow y = 0$, and nondegeneracy from the left: $B(x, V) = 0 \Rightarrow x = 0$? The answer is no. In fact if we view vectors $x, y \in V$ as $n \times 1$ columns, we may write $B(x, y) = [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[y]_{\mathfrak{X}}$, and if $[B]_{\mathfrak{X}}$ is singular there would be some $y \neq 0$ such that $[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} = 0$, hence $B(V, y) = 0$. That can't happen if B is right nondegenerate so B right-nondegenerate implies $[B]_{\mathfrak{X}}$ is nonsingular. The same argument shows B left-nondegenerate also implies $[B]_{\mathfrak{X}}$ nonsingular.

But in fact, this works in both directions, so

1.9. Lemma. B *is right nondegenerate if and only if* $[B]_{\mathfrak{X}}$ *is non singular.*

Proof: We have already proved ($\Leftarrow$) for both left- and right nondegeneracy. Conversely, if $B(V, y) = 0$ for some $y \neq 0$, then $[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} \neq 0$ if $\det([B]_{\mathfrak{X}}) \neq 0$, and we would have

$$B(e_i, y) = e_i^{\mathrm{t}}\,[B]_{\mathfrak{X}}\,[y]_{\mathfrak{X}} \neq 0$$

for some i. This conflicts with the fact that $[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} \neq 0$. Contradiction. $\square$

Thus for any basis $\mathfrak{X}$, B is right-nondegenerate $\Leftrightarrow$ $[B]_{\mathfrak{X}}$ is nonsingular $\Leftrightarrow$ left-nondegenerate, and it is legitimate to drop the "left/right" conditions on nondegeneracy.

Hereafter we will often abbreviate $\dim(V) = |V|$, which is convenient in this and other situations.

1.10. Lemma. *If B is a nondegenerate bilinear form on a finite dimensional space V, and M is a vector subspace, we let $M^{\perp,B} = \{w : B(V,w) = 0\}$. Then*

$$|M| + |M^{\perp,B}| = |V| \ ,$$

even though we need not have $M \cap M^{\perp,B} = (0)$.

Proof: If $|V| < \infty$ any nondegenerate bilinear form B mediates a natural bijection $J : V \to V^*$ that identifies each vector $v \in V$ with a functional $J(v)$ in V^* such that

$$\langle J(v), \ell\rangle = \langle \ell, v\rangle \qquad \text{for all } v \in V, \ell \in V^* \ .$$

This map is clearly $\mathbb{F}$-linear and $J(w) = 0 \Rightarrow B(V,w) = 0 \Rightarrow w = 0$ by non degeneracy of B, so J is one-to-one and also a bijection because $|V| = |V^*|$.

In Section III.3 of the *Linear Algebra I Course Notes*, we defined the "**annihilator**" of a subspace $M \subseteq V$ to be

$$M^\circ = \{\ell \in V^* : \langle \ell, M\rangle = 0\}$$

and discussed its properties, indicating that

$$(M^\circ)^\circ = M \qquad \text{and} \qquad |V| = |M| + |M^\circ|$$

when $|V| < \infty$. The annihilator M° is analogous to the orthogonal complement $M^\perp$ in an inner product space, but it lives in the dual space V^* instead of V; it has the advantage that M° makes sense in *any* vector space V, whether or not it is equipped with an inner product or a nondegenerate bilinear form. (Also, orthogonal complements $M^\perp$ depend on the particular inner product on V, while the annihilator M° has an absolute meaning.)

1.11. Exercise. When V is equipped with a nondegenerate bilinear form B we may invoke the natural isomorphism $V \cong V^*$ it induces to identify an annihilator M° in V^* with a uniquely defined subspace $J^{-1}(M^\circ)$ in V. From the definitions, verify that $M^\circ \subseteq V^*$ becomes the B-orthocomplement $M^{\perp,B} \subseteq V$ under these identifications. $\square$

1.12. Exercise. *If B is a nondegenerate bilinear form on a finite dimensional vector space, and if M* is any subspace, prove that

$$|M| + |M^{\perp,B}| = |V| \qquad \text{and} \qquad \left(M^{\perp,B}\right)^{\perp,B} = M. \tag{34}$$

Hint: Identifiying B-orthocomplements with annihilators, apply the basic properties of annihilators mentioned in Exercise 1.12. $\square$

If B is degenerate, so the radical $\text{rad}(B)$ is nonzero, the role of the radical can be eliminated for most practical purposes, allowing us to focus on nondegenerate forms.

1.13. Exercise. Let $M = \text{rad}(B)$ and form the quotient space $\tilde{V} = V/M$. Show that

1. B induces a well-defined bilinear form $\tilde{B} : \tilde{V} \times \tilde{V} \to \mathbb{F}$ if we let

$$\tilde{B}(x + M, y + M) = B(x,y) \qquad \text{for all } x, y \in V$$

2. $\tilde{B}$ is symmetric (or antisymmetric) $\Leftrightarrow$ B is.

3. Prove that $\tilde{B}$ is now nondegenerate on V/M. $\square$

1.14. Exercise. Given $n \times n$ matrices A, B show that

$$x^{\mathrm{t}}By = x^{\mathrm{t}}Ay \text{ for all } x, y \in \mathbb{F}^n \text{ if and only if } A = B. \quad \square$$

IX.2. Canonical Models for Bilinear Forms.

Bilinear forms arise often in physics and many areas of mathematics are concerned with these objects, so it is of some importance to find natural "*canonical forms*" for B that reveal its properties. This is analogous to the diagonalization problem for linear operators, and we will even speak of "diagonalizing" bilinear forms, although these problems are quite different and have markedly different outcomes.

In doing calculations it is natural to work with the matrices $[B]_{\mathfrak{X}}$ that represent B with respect to various bases, and seek bases yielding the simplest possible form. If a bilinear form B is represented by $A = [B]_{\mathfrak{X}}$ we must examine the effect of a change of basis $\mathfrak{X} \to \mathfrak{Y}$, and describe the new matrix $[B]_{\mathfrak{Y}}$ in terms of the transition matrix $S = [\mathrm{id}]_{\mathfrak{Y}\mathfrak{X}}$ that tells us how to write vectors in the $\mathfrak{Y}$-basis in terms of vectors in $\mathfrak{X}$, as in (32). Thus if $\mathfrak{X} = \{e_i\}$ and $\mathfrak{Y} = \{f_j\}$, $S = [s_{ij}]$ is the matrix such that

$$f_j = \sum_k s_{kj} e_k \qquad \text{for } 1 \leq j \leq n \tag{35}$$

Obviously $\det(S) \neq 0$ because this system of vector equations must be invertible.

In Theorem 1.6 we worked out the effect of such a basis change: $[B]_{\mathfrak{Y}} = S^{\mathrm{t}}[B]_{\mathfrak{X}}S$, which takes the form

$$[B]_{\mathfrak{Y}} = P[B]_{\mathfrak{X}}P^{\mathrm{t}} \qquad \text{if we set } P = S^{\mathrm{t}} \ . \tag{36}$$

We now show that the matrix of a nondegenerate B has a very simple standard form, at least when B is either symmetric or antisymmetric, the forms of greatest interest in applications. We might also ask whether these canonical forms are unique. (Answer: not very.)

The Automorphism Group of a Form B. If a vector space is equipped with a nondegenerate bilinear form B, a natural (and important) *automorphism group* $\mathrm{Aut}(B) \subseteq \mathrm{GL}_{\mathbb{F}}(V)$ comes along with it. It consists of the invertible linear maps $T : V \to V$ that "*leave the form invariant*," in the sense that $B\big(T(x), T(y)\big) = B(x, y)$ for all vectors. We have encountered such automorphism groups before, by various names. For example,

1. The **real orthogonal group** $\mathrm{O}(n)$ consists of the invertible linear maps T on $\mathbb{R}^n$ that preserve the usual inner product,

$$B(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i y_i \quad \text{ for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \ .$$

As explained in Section VI.5 of the *Linear Algebra I Notes*, the automorphisms that preserve this symmetric bilinear form are precisely the linear *rigid motions* on Euclidean space, those that leave invariant lengths of vectors and distances between them, so that

$$\|T(\mathbf{x})\| = \|\mathbf{x}\| \quad \text{ and } \quad \|T(\mathbf{x}) - T(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\| \quad \text{ for } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

where $\|\mathbf{x}\| = \big(\sum_{i=1}^{n} |x_i|^2\big)^{1/2}$ (Pythagoras' formula).

2. The **unitary group** $\mathrm{U}(n)$ is the group of invertible linear operators on $V = \mathbb{C}^n$ that preserve the (Hermitian, sesquilinear) standard inner product

$$B(\mathbf{z},\mathbf{w}) = \sum_{k=1}^{n} z_k \overline{w_k}$$

on complex n-space. For these operators the following conditions are equivalent (see *Linear Algebra I Notes*, Section VI.4).

$$\begin{aligned} T \in \mathrm{U}(n) \quad &\Leftrightarrow \quad B\big(T(\mathbf{z})\,,\,T(\mathbf{w})\big) = B(\mathbf{z},\mathbf{w}) \\ &\Leftrightarrow \quad \|T(\mathbf{z})\| = \|\mathbf{z}\| \\ &\Leftrightarrow \quad \|T(\mathbf{z}) - T(\mathbf{w})\| = \|\mathbf{z}-\mathbf{w}\| \end{aligned}$$

for $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$, where

$$\|\mathbf{z}\| = B(\mathbf{z},\mathbf{z})^{1/2} = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2} \quad \text{(Pythagoras' formula for complex } n\text{-space).}$$

2.1. Exercise. Explain why $\mathrm{U}(n)$ is a closed and bounded subset in matrix space $\mathrm{M}(n,\mathbb{C}) \cong \mathbb{C}^{n^2}$ $\square$

3. The **complex orthogonal group** $\mathrm{O}(n,\mathbb{C})$ is the automorphism group of the bilinear form on complex n-space $\mathbb{C}^n$

$$B(\mathbf{z},\mathbf{w}) = \sum_{k=1}^{n} z_k w_k \qquad (\mathbf{z},\mathbf{w} \in \mathbb{C}^n)$$

This is bilinear over $\mathbb{F} = \mathbb{C}$, but is *not* an inner product because it is not conjugate-linear in the entry $\mathbf{w}$ because w_k appears in B instead of $\overline{w_k}$; furthermore, not all vectors have $B(\mathbf{z},\mathbf{z}) \geq 0$ (try $\mathbf{z} = (1,i)$ in $\mathbb{C}^2$).

In the present section we will systematically examine the canonical forms and associated automorphism groups for nondegenerate symmetric or antisymmetric forms over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. The number of possibilities is surprisingly small.

2.1A. Definition. *The* **automorphism group** *of a nondegenerate symmetric or antisymmetric form* $B : V \times V \to \mathbb{F}$ *is*

$$\mathrm{Aut}(B) = \{T \in \mathrm{GL}_{\mathbb{F}}(V) : B\big(T(v), T(w)\big) = B(v,w) \ \ \textit{for all } v, w \in V\}\ , \tag{37}$$

where $\mathrm{GL}_{\mathbb{F}}(V) = \{T : \det(T) \neq 0\}$ *is the* **general linear group** *consisting of all invertible* $\mathbb{F}$*-linear operators* $T : V \to V$.

$\mathrm{Aut}(B)$ is a *group* because it contains: the identity $I = \mathrm{id}_V$; the composition product $S \circ T$ of any two elements; and the inverse T^{-1} of any element.

Given a basis $\mathfrak{X}$ for V, each element $T \in \mathrm{Aut}(B)$ corresponds to an invertible matrix $[B]_{\mathfrak{X}} = [\,B(e_i, e_j)]$, and these matrices form a group

$$G_{B,\mathfrak{X}} = \{[T]_{\mathfrak{X}} : T \in \mathrm{Aut}(B)\}$$

under matrix multiplication $(\cdot)$. The group $(\mathrm{Aut}(B), \circ)$ and the matrix group $(G_{B,\mathfrak{X}}, \cdot\,)$ are isomorphic and are often identified.

Matrices in $G_{B,\mathfrak{X}}$ are characterized by their special algebraic properties,

$$G_{B,\mathfrak{X}} = \{E \in \mathrm{GL}(n,\mathbb{F}) : E^{\mathrm{t}}[B]_{\mathfrak{X}}E = [B]_{\mathfrak{X}}\}\ , \tag{38}$$

This identification follows because

$$\begin{aligned} T \in \mathrm{Aut}(B) \quad &\Leftrightarrow \quad B\big(T(x), T(y)\big) = B(x,y) \qquad \text{for all } x,y \in V \\ &\Leftrightarrow \quad [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} \;=\; [T(x)]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}\,[T(y)]_{\mathfrak{X}} \\ &\qquad\qquad = [x]^{\mathrm{t}}_{\mathfrak{X}}\big([T]^{\mathrm{t}}_{\mathfrak{X}}\,[B]_{\mathfrak{X}}[T]_{\mathfrak{X}}\big)[y]_{\mathfrak{X}} \\ &\Leftrightarrow \quad [B]_{\mathfrak{X}} \;=\; [T]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}\,[T]_{\mathfrak{X}} \qquad \text{for all } x,y \in V \ . \end{aligned}$$

Given basis $\mathfrak{X}$, T is an automorphism of the bilinear form B if and only if the matrix $[T]_{\mathfrak{X}}$ satisfies the identity $[B]_{\mathfrak{X}} = [T]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[T]_{\mathfrak{X}}$, and this must be true for any basis $\mathfrak{X}$. Matrices in $G_{B,\mathfrak{X}}$ are precisely the matrix realizations (with respect to basis $\mathfrak{X}$) of all the automorphisms in $\mathrm{Aut}(B)$.

2.2. Exercise. If B is a non degenerate bilinear form, show that $G_B = \mathrm{Aut}(B)$ is a subgroup in the general linear group $\mathrm{GL}_{\mathbb{F}}(V)$ – i.e. that (i) $I \in G_B$, (ii) $T_1, T_2 \in G_B \Rightarrow T_1, T_2 \in G_B$, and (iii) $T \in G_B \Rightarrow T^{-1} \in G_B$. $\square$

We can also assess the effect of change of basis $\mathfrak{X} \rightarrow \mathfrak{Y}$: $G_{B,\mathfrak{Y}}$ is a conjugate of $G_{B,\mathfrak{X}}$ under the action of $\mathrm{GL}(n,\mathbb{F})$.

2.3. Exercise. If $\mathfrak{X}$, $\mathfrak{Y}$ are bases in V, define $G_{B,\mathfrak{X}}$ and $G_{B,\mathfrak{Y}}$ as in (38) and prove that

$$G_{B,\mathfrak{Y}} = S^{-1} G_{B,\mathfrak{X}} S \qquad \text{where } S = [\mathrm{id}]_{\mathfrak{Y},\mathfrak{X}}$$

(or equivalently $G_{B,\mathfrak{Y}} = \tilde{S} G_{B,\mathfrak{X}} \tilde{S}^{-1}$ where $\tilde{S} = [\mathrm{id}]_{\mathfrak{Y},\mathfrak{X}}$ since $[\mathrm{id}]_{\mathfrak{Y},\mathfrak{X}} \cdot [\mathrm{id}]_{\mathfrak{X},\mathfrak{Y}} = I$). $\square$

Recall that S is the matrix such that $f_i = \sum_{k=1}^{n} s_{ji} e_j$ if $\mathfrak{X} = \{e_i\}$, $\mathfrak{Y} = \{f_j\}$.

The general linear group $\mathrm{GL}_{\mathbb{F}}(V)$ in which all these automorphism groups live is defined by the condition $\det(T) \neq 0$, which makes no reference to a bilinear form. The **special linear group** $\mathrm{SL}_{\mathbb{F}}(V) = \{T \in \mathrm{GL}_{\mathbb{F}}(V) : \det(T) = 1\}$ is another “classical group” that does not arise as the automorphism group of a bilinear form B. All the other classical groups of physics and geometry are automorphism groups, or their intersections with $\mathrm{SL}_{\mathbb{F}}(V)$

Canonical Forms for Symmetric and Antisymmetric B. We classify the congruence classes of nondegenerate bilinear forms according to whether B is symmetric or antisymmetric, and whether the ground field is $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, always assuming B is nondegenerate. The analysis is the same for antisymmetric forms over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, so there are really only three cases to deal with.

Canonical Forms. Case 1: B symmetric, $\mathbb{F} = \mathbb{R}$.

If B is a nondegenerate symmetric bilinear form on a vector space over $\mathbb{R}$ with $\dim(V) = n$, there are $n+1$ possible canonical forms.

2.4. Theorem (B symmetric; $\mathbb{F} = \mathbb{R}$). *is an $\mathbb{R}$-basis $\mathfrak{X} \subseteq V$ such that the matrix describing B has the form*

$$[B]_{\mathfrak{X}} = \begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} \qquad \textit{with } p + q = n = \dim(V) \ . \tag{39}$$

In this case, we say B has **signature** (p,q).

Proof: First observe that we have a polarization identity for symmetric B that determines $B(v,w)$ from homogheneous expressions of the form $B(u,u)$, just as with inner products over $\mathbb{R}$.

(40) Polarization Identity: $$B(v,w) = \tfrac{1}{2}\big[\,B(v+w, v+w) - B(v,v) - B(w,w)\,\big]$$

for all $v, w \in V$.

2.5. Definition. *The map $Q(v) = B(v,v)$ from V to $\mathbb{R}$ is the* **quadratic form** *associated with a symmetric is bilinear form. Note that $B(\lambda v, \lambda v) = \lambda^2 B(v,v)$, and the quadratic form $Q : V \to \mathbb{R}$ determines the full bilinear form $B : V \times V \to \mathbb{F}$ via the polarization identity* (40).

Therefore, since $B \not\equiv 0$ there is some $v_1 \neq 0$, such that $B(v_1, v_1) \neq 0$, and after scaling v_1 by some $a \neq 0$ we can insure that $B(v_1, v_1) = \pm 1$. But because $\mathbb{F} = \mathbb{R}$ we can't control whether the outcome will be $+1$ or -1.

Let $M_1 = \mathbb{R}{\cdot}v_1$ and

$$M_1^{\perp,B} = \{v \in V : B(V, v_1) = 0\}\ .$$

We have $M_1 \cap M_1^{\perp,B} = \{0\}$ because any w in the intersection must have the form $w = c_1 v_1$, $c_1 \in \mathbb{R}$. But $w \in M_1^{\perp,B}$ too, so $0 = B(w,w) = c_1^2 B(v_1, v_1) = \pm c_1^2$, hence, $c_1 = 0$ and $w = 0$. Therefore $M_1 \oplus M_1^{\perp,B} = V$ because $|W| + |W^{\perp,B}| = |V|$ for any $W \subseteq V$ (Exercise 1.12). $\big[$For an alternative proof: recall the general result about the dimensions of subspaces W_1, W_2 in a vector space V: $|W_1 + W_2| = |W_1| + |W_2| - |W_1 \cap W_2|$.$\big]$

If B_1 is the restriction of B to $M_1^{\perp}$ we claim that $B_1 : M_1^{\perp,B} \times M_1^{\perp,B} \to \mathbb{R}$ is nondegenerate on the lower-dimensional subspace $M_1^{\perp,B}$. Otherwise, there would be an $x \in M_1^{\perp,B}$ such that $B\big(x, M_1^{\perp,B}\big) = 0$. But since $x \in M_1^{\perp,B}$ too, we also have $B(x, M_1) = 0$, and therefore by additivity of B in each entry,

$$B(x, V) = B(x\,,\, M_1^{\perp,B} + M_1) = 0\ .$$

Nondegenerancy of B on V then forces $x = 0$.

We may therefore continue by induction on $\dim(V)$. Choosing a suitable basis $\mathfrak{X}' = \{v_2, \cdots, v_n\}$ in $M_1^{\perp,B}$ and $\mathfrak{X} = \{v_1, v_2, \ldots, v_n\}$ in V we get

$$[B]_{\mathfrak{X}} = \begin{pmatrix} \boxed{\pm 1} & \cdots & 0 \\ 0 & \boxed{I_{p\times p}} & \\ 0 & & \boxed{-I_{q\times q}} \end{pmatrix} \qquad \text{with } p + q = n - 1.$$

If the top left entry is -1, we may switch vectors $e_1 \leftrightarrow e_p$, which replaces $[B]_{\mathfrak{X}}$ with $[B]_{\mathfrak{Y}} = E^{\mathrm{t}} [B]_{\mathfrak{X}} E$, where E is the following permutation matrix (the zero on the diagonal is at the position p)

$$E = \begin{pmatrix} 0 & 0 & \cdot & \cdot & 1 & \cdot & \cdot & 0 \\ 0 & +1 & & & & & & 0 \\ \cdot & & \ddots & & & & & \\ \cdot & & & +1 & & & & \\ 1 & & & & 0 & & & \\ \cdot & & & & & -1 & & \\ \cdot & & & & & & \ddots & \\ 0 & 0 & & & & & & -1 \end{pmatrix}$$

(Note that $E^{\mathrm{t}} = E$ for this particular permutation matrix). Then $[B]_{\mathfrak{Y}}$ has the block-diagonal form (39), completing the proof. $\square$

Later on, we will describe an algorithmic procedure for putting B into canonical form $\mathrm{diag}(+1, \cdots, +1, -1, \cdots, -1)$; these algorithms work the same way over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We will also see that an antisymmetric B cannot be diagonalized by any congruence, but they do have a different (and equally useful) canonical form.

The Real Orthogonal Groups $\mathbf{O(p,q)}$**,** $\mathbf{p+q = n.}$ The outcome in Theorem 2.4 breaks into $n+1$ possibilities. If $\mathfrak{X}$ is a basis such that $[B]_{\mathfrak{X}}$ has the standard form (39), then $A \in G_{B,\mathfrak{X}}$ if and only if

$$A^{\mathrm{t}} \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right) A = \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right) \tag{41}$$

This condition can be written concisely as $A^{\mathrm{t}}JA = J$ where $J = \left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right)$.

The members of this family of classical matrix groups over $\mathbb{R}$ are denoted by $\mathrm{O}(p,q)$, and each one contains as a subgroup the **special orthogonal group of signature** (p,q),

$$\mathrm{SO}(p,q) = \mathrm{O}(p,q) \cap \mathrm{SL}(n,\mathbb{R}) \ .$$

Several of the groups $\mathrm{O}(p,q)$ and $\mathrm{SO}(p,q)$, are of particular interest.

The real Orthogonal Groups $\mathrm{O}(n,0) = \mathrm{O}(n)$ and $\mathrm{SO}(n)$. With respect to the standard basis in $\mathbb{R}^n$ we have $B_{\mathfrak{X}} = I_{n\times n}$, so $J = I_{n\times n}$ in (41) and

$$\mathrm{O}(n,0) = G_{B,\mathfrak{X}} = \{A : A^{\mathrm{t}}A = A^{\mathrm{t}}IA = I\} \ .$$

Thus $\mathrm{O}(n,0)$ is the familar group of orthogonal transformations on $\mathbb{R}^n$, traditionally denoted $\mathrm{O}(n)$. This group is a closed and bounded set in matrix space $\mathrm{M}(n,\mathbb{R}) \cong \mathbb{R}^{n^2}$. □

The Lorentz Group $\mathrm{O}(n-1,1)$. This is the group of space-time symmetries at the center of Einstein's theory of special relativity for $n-1$ space dimensions $x_1,\ldots,x_{n-1}$ and one time dimension x_n which is generally labeled "t" by physicists. For a suitably chosen basis $\mathfrak{X}$ in $\mathbb{R}^n$ the matrix describing an arbitrary nondegenerate symmetric bilinear form B of signature $(n-1,1)$ becomes

$$[B]_{\mathfrak{X}} = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & -1 \end{pmatrix} , \tag{42}$$

and the associated quadratic form is

$$B(x,x) = [x]^{\mathrm{t}}_{\mathfrak{X}}[B]_{\mathfrak{X}}[x]_{\mathfrak{X}} = x_1^2 + \ldots + x_{n-1}^2 - x_n^2$$

Note: The physicists' version of this is a little different:

$$B(x,x) = x_1^2 + \ldots + x_{n-1}^2 - c^2t^2 \ ,$$

where c is the speed of light. But the numerical value of c depends on the physical units used to describe it – *feet per second*, etc – and one can always choose the units of (*length*) and (*time*) to make the experimentally measured speed of light have numerical value $c = 1$. For instance we could take $t =$ (*seconds*) and measure lengths in (*light seconds*) = the distance a light ray travels in one second; or, we could measure t in (*years*) and lengths in (*light years*). Either way, the numerical value of the speed of light is $c = 1$. □

From (41) it is clear that A is in $\mathrm{O}(n-1,1)$ if and only if

$$A^{\mathrm{t}} \left(\begin{array}{c|c} I_{n-1} & 0 \\ \hline 0 & -1 \end{array}\right) A = \left(\begin{array}{c|c} I_{n-1} & 0 \\ \hline 0 & -1 \end{array}\right) \tag{43}$$

$O(n-1,1)$ contains the subgroup $SO(n-1,1) = O(n-1,1) \cap SL(n,\mathbb{R})$ of "*proper*" Lorentz transformations, those having determinant $+1$. Within $SO(n-1,1)$ we find a copy $\widetilde{SO}(n-1)$ of the standard orthogonal group $SO(n-1) \subseteq M(n-1,\mathbb{R})$, embedded in $M(n,\mathbb{R})$ via the one-to-one homomorphism

$$A \in \widetilde{SO}(n-1) \subseteq M(n-1,\mathbb{R}) \;\mapsto\; \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right) \in SO(n-1,1) \subseteq M(n,\mathbb{R}) \;.$$

The subgroup $\widetilde{SO}(n-1)$ acts only on the "space coordinates" $x_1, \cdots, x_{n-1}$ in $\mathbb{R}^n$, leaving the time coordinate $t = x_n$ fixed.

The following family of matrices in $O(n-1,1)$ is of particular interest in understanding the meaning of special relativity.

$$A = \begin{pmatrix} 1/\sqrt{1-v^2} & 0 & & 0 & -v/\sqrt{1-v^2} \\ 0 & 1 & \ddots & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & 1 & 0 \\ -v/\sqrt{1-v^2} & 0 & \cdots & 0 & 1/\sqrt{1-v^2} \end{pmatrix} \tag{44}$$

When we employ units that make the speed of light $c = 1$, the parameter v must have values $|v| < 1$ to prevent the corner entries in this array from having physically meaningless imaginary values; as $v \to 1$ these entries blow up, so $SO(n-1,1)$ is indeed an unbounded set in matrix space $M(n,\mathbb{R})$.

In special relativity, an *event* is described by a point $(\mathbf{x}, t)$ in space-time $\mathbb{R}^{n-1} \times \mathbb{R}$ that specifies the location $\mathbf{x}$ and the time t at which the event occurred. Now suppose two observers are moving through space at constant velocity with respect to one another (*no acceleration* as time passes). Each will use his or her own frame of reference in observing an event to assign space-time coordinates to it. The matrix A in (44) tells us how to make the (relativistic) transition from the values $(\mathbf{x}, t)$ seen by Observer #1 to those recorded by Observer #2:[1]

$$\begin{pmatrix} \mathbf{x}' \\ t' \end{pmatrix} = A \cdot \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}$$

2.6. Exercise. Verify that the matrices in (44) all lie in $SO(n-1,1)$. Be sure to check that $\det(A) = +1$.
Note: Show that (41) $\Rightarrow \det(A)^2 = 1$, so $\det(A) = \pm 1$, and then argue that $\det(I) = 1$ and $\det(A)$ is a continuous function of the real-valued parameter $-1 < v < +1$. $\Box$.

2.7. Exercise. Show that

$$B = \begin{pmatrix} \cosh(y) & 0 & 0 & \sinh(y) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(y) & 0 & 0 & \cosh(y) \end{pmatrix}$$

is in $SO(3,1)$ for all $y \in \mathbb{R}$. $\Box$

A Final Remark about (44). If we work with physical units that do not make $c = 1$, as assumed in (44), we must replace "$\sqrt{1-v^2}$" everywhere it appears with

$$\sqrt{1 - \left(\frac{v}{c}\right)^2}$$

[1] To keep things simple, the transition matrix (44) describes what happens when Observer #2 is moving with velocity v in the positive x_1-direction, as seen by Observer #1, so that $x_1' = x_1 - vt, x_2' = x_2, \ldots, x_{n-1}' = x_{n-1}$. The general formula is more complicated.

in which the speed of light c appears explicitly $\quad\square$.

Invariance of the Signature for $A \in \mathrm{O}(p,q)$. One way to compute the signature would be to find a basis that puts $[B]_{\mathfrak{X}}$ into the block-diagonal form (39), but how do we know the signature does not depend on the basis used to compute it? That it does not is the subject of the next theorem. Proving this amounts to showing that the signature is a **congruence invariant**: you cannot transform

$$\begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} \quad \text{to} \quad S^{\mathrm{t}}\begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} S = \begin{pmatrix} \boxed{I_{p'\times p'}} & 0 \\ 0 & \boxed{-I_{q'\times q'}} \end{pmatrix}$$

unless $p' = p$ and $q' = q$. This fact is often referred to as "*Sylvesters's Law of Intertia.*"

2.8. Theorem (Sylvester). *If A is a nondegenerate real symmetric $n\times n$ matrix, then there is some $P \in \mathrm{GL}(n,\mathbb{R})$ such that $P^{\mathrm{t}}AP = \mathrm{diag}(1,\cdots,1,-1,\cdots,-1)$. The number p of $+1$ entries and the canonical form* (39) *are uniquely determined.*

Proof: The existence of a diagonalization has already been proved. If $B(\mathbf{x},\mathbf{y}) = \sum_{i,j} x_i A_{ij} y_j = \mathbf{x}^{\mathrm{t}}A\mathbf{y}$ is a nondegenerate symmetric bilinear form on $\mathbb{R}^n$, so $[B] = [A_{ij}]$ with respect to the standard basis, then there is a basis $\mathfrak{X}$ such that $[B]_{\mathfrak{X}} = \mathrm{diag}(1,\cdots,1,-1,\cdots,-1)$. Suppose $p = \#(\text{entries} = +1)$ for $\mathfrak{X}$, and that there is another diagonalizing basis $\mathfrak{Y}$ such that $p' = \#(\text{entries} = +1)$ is $\neq p$. We may assume $p < p'$. Writing $\mathfrak{X} = \{v_1,\cdots,v_p,v_{p+1},\cdots,v_n\}$ and $\mathfrak{Y} = \{w_1,\cdots,w_{p'},w_{p'+1},\cdots,w_n\}$, define $L : V \to \mathbb{R}^{p-p'+n}$ via

$$L(\mathbf{x}) = \big(B(\mathbf{x},v_1),\cdots,B(\mathbf{x},v_p),B(\mathbf{x},w_{p'+1}),\cdots,B(\mathbf{x},w_n)\big)$$

The rank $\mathrm{rk}(L)$ of this linear operator is at most $\dim(\mathbb{R}^{p-p'+n}) = p - p' + n < n$, hence $\dim(\ker(L)) = \dim(V) - \mathrm{rk}(L) > 0$ and there is some $v_0 \neq 0$ in V such that $L(v_0) = 0$. That means

$$B(v_0,v_i) = 0 \quad \text{for } 1 \le i \le p \qquad \text{and} \qquad B(v_0,w_i) = 0 \quad \text{for } p'+1 \le i \le n\ .$$

Writing v_0 in terms of the two bases we have $v_0 = \sum_{j=1}^n a_j v_j = \sum_{k=1}^n b_k w_k$.

For $i \le p$ we get

$$\begin{aligned} 0 = B(v_0,v_i) &= B\big(\sum_j a_j v_j\ ,\ v_i\big) = \sum_j a_j B(v_j,v_i) \\ &= \sum_j a_j \delta_{ij} = a_i = a_i B(v_i,v_i)\ , \end{aligned}$$

since $[B]_{\mathfrak{X}} = \mathrm{diag}(1,\cdots,1,-1,\cdots,-1)$. But $B(v_i,v_i) > 0$ for $i \le p$ while $B(v_0,v_i) = 0$, so we conclude that $a_i = 0$ for $0 \le i \le p$. Similarly, $b_j = 0$ for $p'+1 \le j \le n$.

It follows that $a_i \neq 0$ for some $p' < i \le n$, and hence

$$\begin{aligned} B(v_0,v_0) &= B\big(\sum_{j=1}^n a_j v_j\ ,\ \sum_{\ell=1}^n a_\ell v_\ell\big) = \sum_{j=1}^n a_j^2 B(v_j,v_j) \\ &= \sum_{j=p+1}^n a_j^2 B(v_j,v_j) < 0\ . \end{aligned}$$

Furthermore,

$$\begin{aligned} B(v_0,v_0) &= B\big(\sum_{j=1}^n b_j w_j\ ,\ \sum_{\ell=1}^n b_\ell w_\ell\big) = \sum_{j=1}^n b_j^2 B(w_j,w_j) \\ &= \sum_{j=1}^{p'} b_j^2 B(w_j,w_j) > 0\ . \end{aligned}$$

Thus $B(v_0, v_0) < 0$ and $B(v_0, v_0) \geq 0$, which is a contradiction. $\square$

2.9. Corollary. *Two non singular symmetric matrices in $M(n, \mathbb{R})$ are congruent via $A \to P^{\mathrm{t}}AP$ for some $P \in \mathrm{GL}(n, \mathbb{R})$ if and only if they have the same signature (p, q).*

Let A be a symmetric $n \times n$ matrix with entries from a field $\mathbb{F}$ not of characteristic two. We know that there are matrices $Q, D \in \mathrm{M}(n, \mathbb{F})$ such that Q is invertible and $Q^{\mathrm{t}}AQ = D$ is diagonal. We now give a method for computing suitable Q and diagonal form D via elementary row and column operations; a short additional step then yields the signature (p, q) when $\mathbb{F} = \mathbb{R}$.

The Diagonalization Algorithm. Recall that the effect of an elementary row operation on A is obtained by right multiplication $A \mapsto AE$ by a suitable "elementary matrix" E, as explained in *Linear Algebra I Notes*, Sections I-1 and IV-2. Furthermore, the same elementary operation on columns is effected by a left multiplication $A \mapsto E^{\mathrm{t}}A$ using the same E. If we perform an elementary operation on rows followed by the same elementary operation on columns, this is effected by

$$A \mapsto E^{\mathrm{t}}AE$$

(The order of the operations can be reversed because matrix multiplication is associative.)

Now suppose that Q is an invertible matrix such that $Q^{\mathrm{t}}AQ = D$ is diagonal. Any invertible Q is a product of elementary matrices, say $Q = E_1E_2 \cdots E_k$, hence

$$D = Q^{\mathrm{t}}AQ = E_k^{\mathrm{t}}E_{k-1}^{\mathrm{t}} \cdot \ldots \cdot E_1^{\mathrm{t}}AE_1E_2 \cdot \ldots \cdot E_k$$

Putting these observations together we get

2.10. Lemma. *A sequence of paired elementary row and column operations can transform any real symmetric matrix A into a diagonal matrix D. Furthermore, if $E_1, \cdots, E_k$ are the appropriate elementary matrices that yield the necessary row operations (indexed in the order performed), then $Q^{\mathrm{t}}AQ = D$ if we take $Q = E_1E_2 \cdots E_k$.*

2.11. Example. *Let A be the symmetric matrix in* $\mathrm{M}(3, \mathbb{R})$

$$A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

We apply the procedure just described to find an invertible matrix Q such that $Q^{\mathrm{t}}AQ = D$ is diagonal.

Discussion: We begin by eliminating all of the nonzero entries in the first row and first column except for the entry a_{11}. To this end we start by performing the column operation $Col(2) \to Col(2) + Col(1)$; this yields a new matrix to which we apply the same operation on rows, $Row(2) \to Row(2) + Row(1)$. These first steps yield

$$A = \begin{pmatrix} 1 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 \\ -1 & 1 & 1 \\ 3 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 3 & 4 & 1 \end{pmatrix} = E_1^{\mathrm{t}}AE_1$$

where

$$E_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

The second round of moves is: $Col(3) \to Col(3) - 3 \cdot Col(1)$ followed by $Row(3) \to Row(3) - 3 \cdot Row(1)$, which yields

$$\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 4 \\ 3 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 3 & 4 & -8 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 4 & -8 \end{pmatrix} = E_2^{\mathrm{t}} E_1^{\mathrm{t}} A E_1 E_2$$

where

$$E_2 = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Finally we achieve a diagonal form by applying $Col(3) \to Col(3) - 4 \cdot Col(2)$ and then the corresponding operation on rows to get

$$E_3^{\mathrm{t}} E_2^{\mathrm{t}} E_1^{\mathrm{t}} A E_1 E_2 E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -24 \end{pmatrix} \qquad \text{where} \qquad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} .$$

Since the outcome is a diagonal matrix, the process is complete. To summarize: taking

$$Q = E_1 E_2 E_3 = \begin{pmatrix} 1 & 1 & -7 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{we get a diagonal form} \qquad D = Q^{\mathrm{t}} A Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -24 \end{pmatrix}$$

To obtain the canonical form (39) we need one more pair of operations

$$Row(3) \to \frac{1}{\sqrt{24}} \cdot Row(3) \qquad \text{and} \qquad Col(3) \to \frac{1}{\sqrt{24}} \cdot Col(3) \ ,$$

both of which correspond to the (diagonal) elementary matrix

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{24}} \end{pmatrix} .$$

The canonical form is

$$\mathrm{diag}(1, 1, 1, -1) = \tilde{Q}^{\mathrm{t}} A \tilde{Q} \qquad \text{where} \qquad \tilde{Q} = E_4 \cdot \ldots \cdot E_1 \quad \square$$

This example also shows that the diagonal form of a real symmetric matrix achieved through congruence transformations $A \to Q^{\mathrm{t}} A Q$ is not unique; both $\mathrm{diag}(1, 1, 1, -24)$ and $\mathrm{diag}(1, 1, 1, -1)$ are congruent to A. Only the signature $(3, 1)$ is a true congruence invariant.

In Section IV-2 of the *Linear Algebra I Notes* we showed that the inverse A^{-1} of an invertible matrix can be obtained multiplying on the left by a sequence of elementary matrices (or equivalently, by executing the corresponding sequence of elementary row operations). We also developed the *Gauss-Seidel Algorithm* does this efficiently.

> GAUSS-SEIDEL ALGORITHM. *Starting with the $n \times 2n$ augmented matrix $\left[A : I_{n\times n}\right]$, apply row operations to bring the left-hand block into reduced echelon form, which must equal $I_{n\times n}$ since A is invertible. Applying the same moves to the entire $n \times 2n$ augmented matrix we arrive at a matrix $\left[I_{n\times n} : A^{-1}\right]$ whose right-hand block is the desired inverse.*

An algorithm similar to Gauss-Seidel yields a matrix Q such that $Q^{\mathrm{t}}AQ = D$ is diagnonal; the signature (r, s) can then be determined by inspection as in the last steps of Example 2.11. The reader should justify the method, illustrated below, for computing an appropriate Q without recording each elementary matrix separately. Starting with an augmented $n \times 2n$ matrix $\left[A : I_{n\times n}\right]$, we apply paired row and column operations to drive the left-hand block into diagonal form; but we apply them to the *entire augmented matrix.* When the left-hand block achieves diagonal form D the right-hand block in $\left[D : Q^{\mathrm{t}}\right]$ is a matrix such that $Q^{\mathrm{t}}AQ = D$. The steps are worked out below; we leave the reader to verify that $Q^{\mathrm{t}}AQ = D$.

Details: Starting with $Col(2) \rightarrow Col(2) + Col(1)$ and then the corresponding operation on rows, we get

$$
\begin{aligned}
[A : I\,] \;=\; &\left(\begin{array}{ccc|ccc} 1 & -1 & 3 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{paired R/C opns.}} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 1 & 0 \\ 3 & 4 & 1 & 0 & 0 & 1 \end{array}\right) \\
&\xrightarrow{\text{paired R/C opns.}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 4 & 1 & 1 & 0 \\ 0 & 4 & -8 & -3 & 0 & 1 \end{array}\right) \\
&\xrightarrow{\text{paired R/C opns.}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & -24 & -7 & -4 & 1 \end{array}\right) \rightarrow \; [\,D : Q^{\mathrm{t}}]
\end{aligned}
$$

Therefore,

$$
Q^{\mathrm{t}} \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -7 & -4 & 1 \end{pmatrix} \qquad\qquad Q \;=\; \begin{pmatrix} 1 & 1 & -7 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix}
$$

and a diagonalized form $Q^{\mathrm{t}}AQ$ is

$$
D \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -24 \end{pmatrix}
$$

We now turn to the the next type of bilinear form to be analyzed.

Canonical Forms. Case 2: B symmetric, $\mathbb{F} = \mathbb{C}$.

In this case there is just one canonical form.

2.12. Theorem (B symmetric; $\mathbb{F} = \mathbb{C}$). *If B is a nondegenerate, symmetric bilinear form over $\mathbb{F} = \mathbb{C}$ there is a basis $\mathfrak{X}$ such that $[B]_{\mathfrak{X}} = I_{n\times n}$. In coordinates, for this basis we have*

$$
B(x, y) = \sum_{j=1}^{n} x_j y_j \qquad (\textit{no conjugate, even though } \mathbb{F} = \mathbb{C})\ .
$$

Proof: We know (by our discussion of $\mathbb{F} = \mathbb{R}$), we can put B in diagonal form $[B]_{\mathfrak{X}} = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, with each $\lambda_i \neq 0$ since B is nondegenerate. Now take square roots in $\mathbb{C}$ and let $P = \mathrm{diag}(1/\sqrt{\lambda_1}, \cdots, 1/\sqrt{\lambda_n})$ to get $P^{\mathrm{t}}[B]_{\mathfrak{X}}P = I_{n\times n}$. $\square$

There is just one matrix automorphism group, modulo conjugations in $\mathrm{GL}(n, \mathbb{C})$. Taking a basis such that $[B]_{\mathfrak{X}} = I$, we get the **complex orthogonal group** in $\mathrm{M}(n, \mathbb{C})$,

$$
\mathrm{O}(n, \mathbb{C}) = G_{B,\mathfrak{X}} = \{A \in \mathrm{M}(n, \mathbb{C}) : \det(A) \neq 0 \text{ and } A^{\mathrm{t}}A = I\}
$$

(Note our use of the transpose A^{t} here, not the adjoint $A^* = \overline{A^{\mathrm{t}}}$, even though $\mathbb{F} = \mathbb{C}$. As a subgroup we have the **special orthogonal group** over $\mathbb{C}$,

$$
\mathrm{SO}(n, \mathbb{C}) = \mathrm{O}(n, \mathbb{C}) \cap \mathrm{SL}(n, \mathbb{C})
$$

These are closed unbounded subsets in and $\mathrm{M}(n, \mathbb{C})$.

2.13. Exercise.

1. Show that $\mathrm{SO}(2, \mathbb{C})$ is abelian and isomorphic to the direct product group $S^1 \times \mathbb{R}$ where $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and the product operation is
$$(z, x) \cdot (z'x') = (zz', x + x')$$
2. Show that $A \in \mathrm{SO}(2, \mathbb{C})$ if and only if
$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$
with $a, b \in \mathbb{C}$ and $a^2 + b^2 = 1$.
3. Show that $\mathrm{SO}(2, \mathbb{C})$ is an unbounded subset in $\mathrm{M}(2, \mathbb{C})$, and hence that $\mathrm{SO}(n, \mathbb{C})$ is unbounded in $\mathrm{M}(n, \mathbb{C})$ because we may embed $\mathrm{SO}(2, \mathbb{C})$ in $\mathrm{SO}(n, \mathbb{C})$ via
$$A \in \mathrm{SO}(2, \mathbb{C}) \mapsto \left(\begin{array}{c|ccc} A & 0 & \cdot & 0 \\ \hline 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \end{array}\right)$$
if $n \geq 2$.

Hints: For (1.) you must produce an explicit bijection $\Phi : S^1 \times \mathbb{R} \to \mathrm{SO}(2, \mathbb{C})$ such that $\Phi(q_1, q_2) = \Phi(q_1) \cdot \Phi(q_2)$ (matrix product of elements in $\mathrm{M}(2, \mathbb{C})$). In (2.), if we write $A = [a, b; c, d]$ the identities $A^{\mathrm{t}}A = I = AA^{\mathrm{t}}$ plus $\det(A) = 1$ yield 9 equations in the complex unknowns a, b, c, d, which reduce to 7 when duplicates are deleted. There is a lot of redundancy in the remaining system, and it can actually be solved by algebraic elimination despite its nonlinearity . In (3.) use the sup-norm $\|A\| = \max_{i,j}\{|A_{ij}|\}$ to discuss bounded sets in matrix space. $\square$

Note: A similar problem was posed in the *Linear Algebra I Notes* regarding the group of real matrices $\mathrm{SO}(3) \subseteq \mathrm{M}(3, \mathbb{R})$ – see *Notes*, Section VI-5, especially Euler's Theorem VI-5.6. The analog for $\mathrm{SO}(3)$ of the problem posed above for $\mathrm{SO}(2, \mathbb{C})$ is crucial in understanding the geometric meaning of the corresponding linear operators $L_A : \mathbb{R}^3 \to \mathbb{R}^3$. By Euler's Theorem $\mathrm{SO}(3)$ gets identified as the group of all *rotations* $R_{\ell,\theta} : \mathbb{R}^3 \to \mathbb{R}^3$, by any angle θ about any oriented axis ℓ through the origin. $\square$

2.14. Exercise. Is $\mathrm{SO}(n, \mathbb{C})$ a *closed subset* in $\mathrm{M}(n, \mathbb{C}) \simeq \mathbb{C}^{n^2}$? Prove or disprove. Which scalar matrices λI lie in $\mathrm{SO}(n, \mathbb{C})$ or $\mathrm{O}(n, \mathbb{C})$?

Canonical Forms. Case 3: B Antisymmetric; $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

In the antisymmetric case, the same argument applies whether $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Note that $B(v, v) = 0$ for all v, and if $W \subseteq V$ the B-annihilator $W^{\perp,B} = \{v : B(v, W) = 0\}$ need not be complementary to W. We might even have $W^{\perp,B} \supseteq W$, although the identity $\dim(W) + \dim(W^{\perp,B}) = \dim(V)$ remains valid.

2.15. Theorem (B antisymmetric; $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$). *If B is a nondegenerate antisymmetric form over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, there is a basis $\mathfrak{X}$ such that*
$$[B]_{\mathfrak{X}} = J = \begin{pmatrix} 0 & I_{m\times m} \\ -I_{m\times m} & 0 \end{pmatrix}$$
In particular $\dim_{\mathbb{F}}(V)$ *must be even if V carries a nondegenerate skew-symmetric bilinear form.*

Proof: Recall that $\dim(W) + \dim\left(W^{\perp,B}\right) = \dim(V)$ for any nondegenerate bilinear form B on V. Fix $v_1 \neq 0$. Then $M_1 = (\mathbb{F}v_1)^{\perp,B}$ has dimension $n-1$ if $\dim(V) = n$, but it includes $\mathbb{F}v_1 \subseteq (\mathbb{F}v_1)^{\perp,B}$. Now take any $v_2 \notin M_1$ (so $v_2 \neq 0$) and scale it to get $B(v_1, v_2) = -1$. Let $M_2 = (\mathbb{F}v_2)^{\perp,B}$; again we have $\dim(M_2) = n-1 = \dim(M_1)$. But $M_2 \neq M_1$ since $v_2 \in M_2$ and $v_2 \notin M_1$, so $\dim(M_1 \cap M_2) = n-2$. The space $M = M_1 \cap M_2$ is B-orthogonal to $\mathbb{F}\text{-span}\{v_1, v_2\}$ by definition of these vectors. Furthermore, $B|_M$ is antisymmetric and nondegenerate. [In fact, we already know that $B(w, w_1) = B(w, v_2) = 0$ and $V = \mathbb{F}v_1 \oplus \mathbb{F}v_2 \oplus M$, so if $B(w, M) = 0$ for some $w \in M$, then $B(w, V) = B(w, \mathbb{R}v_1 + \mathbb{R}v_2 + M) = 0$ and $w = 0$ by nondegeneracy.] Furthermore, if $N = \mathbb{F}\text{-span}\{v_1, v_2\}$ we have $V = N \oplus M$. (Why?)

We can now argue by induction on $n = \dim(V)$: $\dim(M)$ must be even and there is a basis $\mathfrak{X}_0 = \{v_3, \cdots, v_n\}$ in M such that

$$[\,B|_M\,]_{\mathfrak{X}_0} = \begin{pmatrix} \boxed{R} & & 0 \\ & \ddots & \\ 0 & & \boxed{R} \end{pmatrix}$$

with

$$R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Hence, $\mathfrak{X} = \{v_1, v_2\} \cup \mathfrak{X}_0$ is a basis for V such that

$$[B]_{\mathfrak{X}} = \left(\begin{array}{c|c} R & 0 \\ \hline 0 & \left[B|_M\right]_{\mathfrak{X}_0} \end{array}\right) = \begin{pmatrix} \boxed{R} & & & 0 \\ & \boxed{R} & & \\ & & \ddots & \\ 0 & & & \boxed{R} \end{pmatrix}$$

A single permutation of basis vectors (corresponding to some permutation matrix E such that $E^{\mathrm{t}} = E^{-1}$) gives the standard form

$$E^{\mathrm{t}}[B]_{\mathfrak{X}}E = [B]_{\mathfrak{Y}} = \left(\begin{array}{c|c} 0 & I_{m\times m} \\ \hline -I_{m\times m} & 0 \end{array}\right)$$

where $m = \frac{1}{2}\dim(V)$. $\square$

A skew-symmetric nondegenerate form B is called a **symplectic structure** on V. The dimension $\dim_{\mathbb{F}}(V)$ must be even, and as we saw earlier there is just one such nondegenerate structure up to congruence of the representative matrix.

2.16. Definition. *The automorphism group* $\mathrm{Aut}(B)$ *of a nondegenerate skew-symmetric form on* V *is called a* **symplectic group**. *If* $\mathfrak{X}$ *is a basis that puts* B *into standard form, we have*

$$B(x, y) = [x]_{\mathfrak{X}}^{\mathrm{t}}[B]_{\mathfrak{X}}[y]_{\mathfrak{X}} = [x]_{\mathfrak{X}}^{\mathrm{t}}J\,[y]_{\mathfrak{X}} \qquad \text{where} \qquad J = \begin{pmatrix} 0 & I_{m\times m} \\ -I_{m\times m} & 0 \end{pmatrix}.$$

By (38)*, elements of* $\mathrm{Aut}(B)$ *are determined by the condition*

$$A \text{ is in } G_{\mathfrak{X},B} \Leftrightarrow A^{\mathrm{t}}JA = J.$$

on $V \simeq \mathbb{R}^{2m}$. *The corresponding matrix group*

$$\mathrm{Sp}(n, \mathbb{F}) = G_{B,\mathfrak{X}} = \{A \in \mathrm{M}(n, \mathbb{F}) : A^{\mathrm{t}}JA = J\}$$

is the classical **symplectic group of degree** $m = \frac{1}{2}\dim(V)$.

The related matrix

$$J' = \begin{pmatrix} \boxed{R} & & 0 \\ & \ddots & \\ 0 & & \boxed{R} \end{pmatrix} \qquad \textit{with} \qquad R = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

is a GL*-conjugate of J, with $J' = CJC^{-1}$ for some $C \in \mathrm{GL}(2m, \mathbb{R})$, and the algebraic condition*

$$A^{\mathrm{t}} J' A = J'$$

determines a subgroup $G' \subseteq \mathrm{GL}(n, \mathbb{F})$ that is conjugate (hence isomorphic to) the matrix group $G_{B,\mathfrak{X}} = \mathrm{Sp}(n, \mathbb{F})$.

Both versions of the commutation relations determining matrix versions of $\mathrm{Aut}(B)$ are used in the literature.

Note: $\det(A) \neq 0$ automatically because $\det(J) = (-1)^m \neq 0$. In fact, $A \in \mathrm{Sp}(n, \mathbb{F})$ implies $\det(J) = \det(A^{\mathrm{t}} JA) \Rightarrow (\det(A))^2 = 1$, so $\det(A) = \pm 1$ whether the underlying field $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. $\square$

The only scalar matrices λI in $\mathrm{Sp}(n, \mathbb{F})$ are those such that $\lambda^2 = 1$. The fact that $\det(J) = (-1)^m$ follows because m row transpositions send $J \to I_{2m \times 2m}$.

IX-3. Sesquilinear Forms ($\mathbb{F} = \mathbb{C}$).

Finally we take up sesquilinear forms $B : V \times V \to \mathbb{C}$ (over complex vector spaces), which are linear functions of the first entry in $B(v, w)$, but conjugate-linear in the second, so that $B(x, \lambda y) = \overline{\lambda} B(x, y)$, $B(\lambda x, y) = \lambda B(x, y)$. There are only a limited number of possibilities.

3.1. Lemma. *A sesquilinear form on V cannot be symmetric or antisymmetric unless it is zero.*

Proof: We know that $\lambda B(x, y) = B(\lambda x, y)$, and if B is (anti-)symmetric this would be equal to $\pm B(x, \lambda y) = \pm \overline{\lambda} B(x, y)$ for all $\lambda \in \mathbb{C}$, $x, y \in V$. This is impossible if $B(x, y) \neq 0$. $\square$

Thus the only natural symmetry properties for sesquilinear forms over $\mathbb{C}$ are

1. Hermitian symmetry: $B(x, y) = \overline{B(y, x)}$
2. Skew-hermitian symmetry: $B(x, y) = -\overline{B(y, x)}$.

However, if B is Hermitian then iB (where $i = \sqrt{-1}$) is skew-Hermitian and vice-versa, so once we analyze Hermitian sesquilinear forms there is nothing new to say about skew-Hermitian forms.

The sesquilinear forms on V are a vector space over $\mathbb{C}$. Every such form is uniquely a sum $B = B_H + B_S$ of a Hermitian and skew-Hermitian form

$$B(v, w) = \frac{B(v, w) + \overline{B(w, v)}}{2} + \frac{B(v, w) - \overline{B(w, v)}}{2} \qquad \text{for all } v, w \in V$$

As usual, a sesquilinear form B is determined by its matrix representation relative to a basis $\mathfrak{X} = \{e_1, \ldots, e_n\}$ in V, given by

$$[B]_{\mathfrak{X}} = [B_{ij}] \quad \text{where } B_{ij} = B(e_i, e_j) \; .$$

Given any basis $\mathfrak{X}$, the form B is

1. Nondegenerate if and only if $[B]_{\mathfrak{X}}$ is nonsingular (nonzero determinant).

2. Hermitian symmetric if and only if $[B]_{\mathfrak{X}}$ is self-adjoint ($= [B]^*_{\mathfrak{X}}$).

3. The correspondence $B \mapsto [B]_{\mathfrak{X}}$ is a $\mathbb{C}$-linear isomorphism between the vector space of sesquilinear forms on V and matrix space $\mathrm{M}(n, \mathbb{C})$.

The change of basis formula is a bit different from that for bilinear forms. If $\mathfrak{Y} = \{f_j\}$ is another basis, related to $\mathfrak{X} = \{e_i\}$ via

$$f_i = \sum_{j=1}^{n} s_{ji} e_j \qquad \text{where} \qquad S = [\mathrm{id}]_{\mathfrak{X},\mathfrak{Y}} \ .$$

we then have

$$\begin{aligned}
\big([B]_{\mathfrak{Y}}\big)_{ij} &= B(f_i, f_j) \ = \ B\Big(\sum_k s_{ki} e_k \,,\, \sum_\ell s_{\ell j} e_\ell\Big) \\
&= \sum_{k,l} s_{ki} \, \overline{s_{\ell j}} \, ([B]_{\mathfrak{X}})_{k\ell} \\
&= (S^{\mathrm{t}} [B]_{\mathfrak{X}} \overline{S})_{ij} \quad \text{where } \overline{S} \text{ is the complex conjugate matrix: } \overline{S}_{ij} = \overline{s_{ij}}
\end{aligned}$$

Letting $P = \overline{S}$, we may rewrite the result of this calculation as

$$[B]_{\mathfrak{Y}} = P^* [B]_{\mathfrak{X}} P \tag{45}$$

where $\det(P) \neq 0$, $P^* = (\overline{P})^{\mathrm{t}}$. In terms of the transition matrix S between bases, we have $P = \overline{S} = \overline{[\mathrm{id}]_{\mathfrak{X}\mathfrak{Y}}}$.

Note that P^* need not to be equal to P^{-1}, so P need not be a unitary matrix in $\mathrm{M}(n, \mathbb{C})$. Formula (45) differs from that for orthogonal matrices in that P^{t} has been replaced by P^*.

3.2. Exercise. If B is sesquilinear, $\mathfrak{X}$ is a basis in V, and $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$ in V, show that

$$B(x, y) = [x]^{\mathrm{t}}_{\mathfrak{X}} [B]_{\mathfrak{X}} [y]^{-}_{\mathfrak{X}}, \quad \text{so that } B(x, y) = \sum_{ij} x_i B_{ij} \overline{y_j} \ . \quad \square$$

3.3. Definition. *A non degenerate sesquilinear form is an* **inner product** *if*

1. Hermitian: *$B(x, y) = \overline{B(y, x)}$;*

2. Positive Definite: *$B(x, x) \geq 0$, $\forall x$*

3. Nondegenerate:: *$B(x, V) = (0) \Leftrightarrow x = 0$.*

Conditions 2. + 3. amount to saying $B(x, x) \geq 0$ and $B(x, x) = 0 \Rightarrow x = 0$ – i.e. the form *strictly* positive definite. This equivalence follows from the polarization identity for Hermitian sesquilinear forms.

3.4. Lemma (Polartization Identity). *If B is a Hermitian sesquilinear form then*

$$B(v, w) = \frac{1}{4}\Big[\sum_{k=0}^{3} i^k B(v + i^k w, v + i^k w)\Big], \qquad \textit{where } i = \sqrt{-1}$$

Proof: Trivial expansion of the sum. $\square$

If B is a nondegenerate Hermitian sesquilinear form and $v \neq 0$ there must be some $w \in V$ such that $B(v, w) \neq 0$, but by the polarization identity nondegeneracy of B

implies that there is some $v \neq 0$ such that $B(v,v) \neq 0$ (and if B is positive definite it must be strictly positive definite). If v_1 is such a vector and $M_1 = \mathbb{R}v_1$, we obviously have $M_1 \cap M_1^{\perp} = (0)$ because $w \in M_1 \cap M_1^{\perp} \Rightarrow w = cv_1$ and also $0 = (w, v_1) = c(v_1, v_1)$, which implies $c = 0$. The restricted form $B|_{M_1^{\perp}}$ is again Hermitian symmetric; it is also nondegenerate because if $B(w, M_1^{\perp}) = 0$ for some nonzero $w \in M_1^{\perp}$, then $B(w, V) = B(w, M_1 + M_1^{\perp}) = (0)$ too, contrary to nondegeneracy of B on V. So, by an induction argument there is a basis $\mathfrak{X} = \{e_1 = v_1, e_2, \cdots, e_n\}$ in V such that

$$[B]_{\mathfrak{X}} = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_n \end{pmatrix}$$

where $\mu_k \in \mathbb{C}$ and $\mu_k \neq 0$ (B being non degenerate).

Since $B(e_i, e_j) = \overline{B(e_j, e_i)}$ we get $\mu_k = \overline{\mu_k}$, so all entries are real and nonzero. Taking $P = \mathrm{diag}\big(1/\sqrt{|\mu_1|}, \cdots, 1/\sqrt{|\mu_n|}\,\big)$, we see that

$$P^*[B]_{\mathfrak{X}}P = \begin{pmatrix} \pm 1 & & 0 \\ & \ddots & \\ 0 & & \pm 1 \end{pmatrix}$$

$= [B]_{\mathfrak{Y}}$ for some new basis $\mathfrak{Y}$; recall the change of basis formula.) Finally apply a permutation matrix (relabel basis vectors) to get

$$[B]_{\mathfrak{Y}} = E^*P^*[B]_{\mathfrak{X}}PE = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \tag{46}$$

where $P = I_{p\times p}$, $Q = -I_{q\times q}$, and $p + q = n = \dim_{\mathbb{C}}(V)$. We have proved

3.5. Proposition. *Every nondegenerate Hermitian sesquilinear form B can be put into the canonical form* (46) *by a suitable choice of basis in V. If $x = \sum_i x_i e_i$, $y = \sum_j y_j e_j$ with respect to a basis such that $[B]_{\mathfrak{X}}$ has canonical form, we get*

$$B(x, y) = \sum_{i=1}^{p} x_i\overline{y_i} - \sum_{i=p+1}^{n} x_i\overline{y_i}$$

In particular, if $p = n$ and $q = 0$ we obtain the standard inner product $(x, y) = \sum_{j=1}^{n} x_j\overline{y_j}$ in $\mathbb{C}^n$ when we identify V with $\mathbb{C}^n$ using the basis $\mathfrak{X}$ such that $[B]_{\mathfrak{X}}$ has the form (46).

There are just $n+1$ $\mathfrak{X}$-congruence classes of nondegenerate Hermitian sesquilinear forms on a complex vector space of dimension n; they are distinguished by their signatures (p, q). The possible automorphism groups

$$\mathrm{Aut}(B) = \{T \in \mathrm{Hom}_{\mathbb{C}}(V,V) : \det(T) \neq 0 \text{ and } B(T(v), T(w)) = B(v,w) \text{ for all } v, w\}$$

are best described as matrix groups $G_{B,\mathfrak{X}}$ with respect to a basis that puts B into canonical form. This yields the **unitary groups of type (p,q)**. $\mathrm{Aut}(B)$ is isomorphic to the matrix group

$$\mathrm{U}(p,q) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^*JA = J\} \quad \text{where} \quad J = \begin{pmatrix} \boxed{I_{p\times p}} & 0 \\ 0 & \boxed{-I_{q\times q}} \end{pmatrix} \tag{47}$$

There is a slight twist in the correspondence between operators $T \in \mathrm{Aut}(B)$ and matrices $A \in \mathrm{U}(p,q)$.

3.6. Exercise. Let B be nondegenerate Hermitian sesquilinear and let $\mathfrak{X} = \{e_i\}$ be a basis such that $[B]_{\mathfrak{X}}$ is in canonical form. If $[T]_{\mathfrak{X}}$ is the matrix associated with $T \in \mathrm{Aut}(B)$, verify that the complex conjugate $A = \big([T]_{\mathfrak{X}}\big)^-$ satisfies the identity (47), and conversely if $A \in \mathrm{U}(p,q)$ then $A = \big([T]_{\mathfrak{X}}\big)^-$ for some $T \in \mathrm{Aut}(B)$. $\square$

Thus the correspondence $\Phi : T \mapsto A = \big([T]_{\mathfrak{X}}\big)^-$ (rather than $T \mapsto A = [T]_{\mathfrak{X}}$) is a bijection between $\mathrm{Aut}(B)$ and the matrix group $\mathrm{U}(p,q) \subseteq \mathrm{GL}(n,\mathbb{C})$ such that $\Phi(T_1 \circ T_2) = \Phi(T_1) \cdot \Phi(T_2)$ (matrix product), and Φ is a group isomorphism between $\mathrm{Aut}(B)$ and $\mathrm{U}(p,q)$.

When $p = n$, we get the classical group of *unitary operators* on an inner product space, and when we identify $V \simeq \mathbb{C}^n$ via a basis such that $[B]_{\mathfrak{X}} = I_{n\times n}$, we get the group of **unitary matrices** in $\mathrm{M}(n,\mathbb{C})$,

$$\mathrm{U}(n) = \mathrm{U}(n,0) = \{A \in \mathrm{GL}(n,\mathbb{C}) : A^*A = I\} \quad (\text{because } A^*IA = A^*A)$$

As a closed subgroup of $\mathrm{U}(n)$ we have the **special unitary group**

$$\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n,\mathbb{C}) \subseteq \mathrm{U}(n)\ .$$

There are also special unitary group of type (p,q), the matrix groups

$$\mathrm{SU}(p,q) = \mathrm{U}(p,q) \cap \mathrm{SL}(n,\mathbb{C})\ .$$

For $A \in \mathrm{U}(p,q)$ the identity (46) implies

$$\det(A^*) \cdot \det\left(\begin{array}{c|c} I_{p\times p} & 0 \\ \hline 0 & -I_{q\times q} \end{array}\right) \cdot \det(A) \;=\; (-1)^q$$

so $|\det(A)|^2 = (-1)^q$ (remember: $\mathbb{F} = \mathbb{C}$ so this could be negative). In particular, $|\det(A)|^2 = 1$ if $A \in \mathrm{U}(n)$, so $\det(A)$ always lies on the unit circle $S^1 = \{z : |z| = 1\}$ in the complex plane.

We already know that unitary matrices are orthogonally diagonalizable since they are normal operators ($A^*A = AA^*$, so $A^*A = I \Leftrightarrow AA^* = I$). Since $\|Ax\|^2 = \|x\|^2$ for all x, all eigenvalues λ_i have absolute value 1, so the spectrum $\mathrm{sp}_{\mathbb{C}}(A)$ is a subset of the unit circle $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ for unitary matrices (or operators). Furthermore, $\mathrm{U}(n)$ contains a copy of the unit circle (which is a group under the usual multiplication of complex number because $|zw| = |z| \cdot |w|$ and $|z| = 1 \Rightarrow |1/z| = 1$); in fact $(S^1,\cdot) \cong \{\lambda I_{n\times n} : |\lambda| = 1\}$. In $\mathrm{SU}(n)$, however, the only scalar matrices are of the form λI where λ is an n^{th} root of unity, $\lambda = e^{2\pi i k/n}$ with $0 \le k \le n$.

Notice the parallel between certain groups over $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$.

1. $\mathrm{SO}(p,q)$ and $\mathrm{O}(p,q)$ over $\mathbb{R}$ are the "real parts" of $\mathrm{SU}(p,q)$ and $\mathrm{U}(p,q)$. In fact we have
$$\mathrm{O}(p,q) = \mathrm{U}(p,q) \cap \big(\mathrm{M}(n,\mathbb{R}) + i0\big)\ .$$
when we identity $\mathrm{M}(n,\mathbb{C}) = \mathrm{M}(n,\mathbb{R}) + \sqrt{-1}\,\mathrm{M}(n,\mathbb{R})$ by splitting a complex matrix $A = [z_{ij}]$ as $[x_{ij}] + \sqrt{-1}\,[y_{ij}]$ if $z_{ij} = x_{ij} + \sqrt{-1}\,y_{ij}$.

2. We also recognize $\mathrm{SO}(n)$ and $\mathrm{O}(n)$ as the real parts of the complex natrix groups $\mathrm{SO}(n,\mathbb{C})$ and $\mathrm{O}(n,\mathbb{C})$, as well as being the real parts of $\mathrm{SU}(n)$ and $\mathrm{U}(n)$.

3.7. Exercise. Prove that $\mathrm{U}(n)$ is a closed bounded subset when we identify $\mathrm{M}(n,\mathbb{C}) \approx \mathbb{C}^{n^2}$; hence it is a compact matrix group. $\square$

3.8. Exercise. If $p \neq n$, prove that $\mathrm{U}(p,q)$ and $\mathrm{SU}(p,q)$ are closed but unbounded subsets in $\mathrm{M}(n,\mathbb{C})$ when $q \neq 0$. $\square$

Chapter X. Matrix Lie Groups.

X.1. Matrix Groups and Implicit Function Theorem.

The rank of a linear operator $T : V \to W$ is $\dim(\text{range}(T)) = \dim(V) - \dim(\ker(T))$. If $\mathfrak{X}$, $\mathfrak{Y}$ are bases in V, W the rank of T can be determined from the matrix $A = [T]_{\mathfrak{Y},\mathfrak{X}}$ as follows. A $k \times k$ submatrix is obtained by designating k rows and k columns and extracting from A the $k \times k$ array where these meet. To describe the outcome we must specify the row indices $I = \{i_1 < \cdots < i_k\}$ and column indices $J = \{j_1 < \cdots < j_k\}$, and then we might indicate how the submatrix was constructed by writing A_{IJ}. Note that A itself is not necessarily square; it is $n \times m$ if $\dim(V) = m$, $\dim(W) = n$.

1.1. Lemma. *Given a nonzero $n \times m$ matrix A its rank* $\text{rk}(A)$ *is equal to*

$$k_{max} = \max\{k \in \mathbb{N} : A \text{ has a nonsingular } k \times k \text{ submatrix } A_{IJ}\}$$

Proof: Obviously $\text{rk}(A) \geq k_{max}$: if A_{IJ} is nonsingular its columns $\{C'_{j_1}, \cdots, C'_{j_k}\}$ are truncated versions of the corresponding columns $\{C_{j_1}, \cdots, C_{j_k}\}$ of A, which forces the latter to be linearly independent. Hence $|J| \leq k_{max} \leq \text{rk}(A)$.

We also have $\text{rk}(A) \leq k_{max}$, for if $\text{rk}(A) = k$ there is some set of column indices with $|J| = k$ such that $\{C_j : i \in J\}$ are linearly independent. If B is the $n \times k$ matrix $[C_{j_1}; \cdots; C_{j_k}]$, it is well known that

$$\text{row rank}(B) = \text{ column rank}(B) ,$$

so we can find a set I of row indices with $|I| = |J| = k$ such that the rows $\{R_i(B) : i \in I\}$ are linearly independent. The rows in the $n \times k$ matrix $B = |C_{j_1}; \cdots; C_{j_k}|$ are truncated versions of the corresponding rows in A, and those with row indices in J are precisely the rows of the $k \times k$ submatrix A_{IJ}. Obviously, this submatrix is nonsingular, so $k_{max} \geq k = \text{rk}(A_{IJ}) = \text{rk}(A)$. $\square$

Note that various choices I, J of row and column indices may yield nonsingular square submatrices A_{IJ} of maximal size.

Smooth Mappings and their Differentials. Now consider a mapping $\mathbf{y} = \phi(\mathbf{x}) = \big(\phi_1(\mathbf{x}), \ldots, \phi_n(\mathbf{x})\big)$ from $\mathbb{F}^m \to \mathbb{F}^n$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, but mostly $\mathbb{R}$ in our discussion). We say that ϕ is a $\mathcal{C}^\infty$ **map** (or **smooth map**) if the scalar components $\phi_k(\mathbf{x})$ have continuous partial derivatives of all orders.

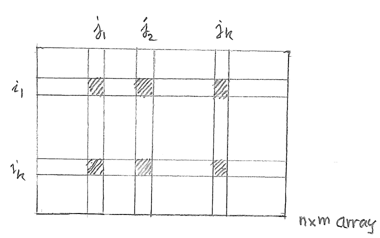

Figure 10.1. A square $k \times k$ submatrix A_{IJ} is extracted from an $n \times m$ matrix by specifying row indices $I = \{i_1 < \cdots < i_k\}$ and column indices $J = \{j_1 < \cdots < j_k\}$. The rank $\text{rk}(A)$ is equal to k if A_{IJ} is nonsingular and all nonsingular square submatrices have size $r \leq k$

The **Jacobian matrix** for ϕ at base point p is the $n \times m$ matrix

$$(d\phi)_p = \begin{pmatrix} \frac{\partial y_1}{\partial x_1}(p) & \cdot & \cdot & \cdot & \frac{\partial y_1}{\partial x_m}(p) \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \frac{\partial y_n}{\partial x_1}(p) & \cdot & \cdot & \cdot & \frac{\partial y_n}{\partial x_m}(p) \end{pmatrix}_{n \times m}$$

whose entries are smooth scalar-valued functions of $\mathbf{x} \in \mathbb{F}^m$. We will be concerned with the rank $\text{rk}(d\phi)_\mathbf{x}$ of the Jacobian matrix at and near various base point. We assign a linear operator, the **differential of ϕ at** p,

$$(d\phi)_p : \mathbb{F}^m \to \mathbb{F}^n \quad \text{ such that } \quad (d\phi)_p(\mathbf{v}) = (d\phi)_p \cdot \text{col}(v_1, \ldots, v_m)$$

at each base point in $\mathbb{F}^m$ where ϕ is smooth. The operator $(d\phi)_p$ is the unique *linear* operator $\mathbb{F}^m \to \mathbb{F}^n$ that "closely approximates" the behavior of the (nonlinear) map $\phi : \mathbb{F}^m \to \mathbb{F}^n$ near p, in the sense that

$$\Delta\phi = \phi(p + \Delta\mathbf{x}) - \phi(p) = (d\phi)_p \cdot (\Delta\mathbf{x}) + E(\Delta\mathbf{x}) \ ,$$

in which the "error term" $E(\Delta\mathbf{x})$ becomes very small compared to $\Delta\mathbf{x}$ for small increments away from the base point p:

$$(48) \qquad \text{Error Estimate:} \quad \frac{\|E(\Delta\mathbf{x})\|}{\|\Delta\mathbf{x}\|} \longrightarrow 0 \text{ in } \mathbb{F}^n \quad \text{ as } \quad \|\Delta\mathbf{x}\| \to 0 \text{ in } \mathbb{F}^m.$$

As a function of the base point $p \in \mathbb{F}^m$, the linear operator (matrix) $(d\phi)_p$ is a $\mathcal{C}^\infty$ map from $\mathbb{F}^m$ into the matrix space $\text{M}(n \times m, \mathbb{F})$.

We define the **rank of ϕ at** p to be the rank $r = \text{rk}(d\phi)_p$ of its Jacobian matrix. As above, we have $\text{rk}(d\phi)_p = r \Leftrightarrow \dim(\text{range}(d\phi)_p) \Leftrightarrow m - \dim(\ker(d\phi)_p) = m - r \Leftrightarrow$ there are r row indices $I = \{i_1 < \cdots < i_r\}$ and column indices $J = \{j_1 < \cdots < j_r\}$, such that

1. The submatrix $(d\phi_p)_{IJ}$ is non singular, and
2. No larger square submatrix (with $k > r$) can be nonsingular, so $r \times r$ is the maximum size of any nonsingular square submatrix.

Note the following points:

1. Various choices of I, J may yield nonsingular submatrices $(d\phi)_{IJ}$ of maximal size $r \times r$. The valid choices of I, J may also vary with the base point p.
2. For fixed choices of indices I, J the entries in the $r \times r$ matrix $(d\phi_\mathbf{x})_{IJ}$ vary smoothly with $\mathbf{x}$, and so does the determinant $\det(d\phi)_\mathbf{x}$, so if the determinant is nonzero at p it must also be nonzero for all x near p. Hence for fixed choice of I, J we have

$$\text{rk}(d\phi_\mathbf{x})_{IJ} \geq r \quad \text{ for all } \mathbf{x} \text{ near } p \text{ if } \text{rk}(d\phi_p)_{IJ} = r.$$

   Now let r_{max} be the largest value $\text{rk}(d\phi_\mathbf{x})$ achieves on $\mathbb{F}^m$. If $\text{rk}(d\phi)_p = r_{max}$ it follows that $\text{rk}(d\phi)_\mathbf{x} = r_{max}$ (constant rank) on some open neighborhood of p in $\mathbb{F}^m$. Quite often, as when $\phi : \mathbb{F}^m \to \mathbb{F}^n$ has scalar components $\phi = (\phi_1(x), \ldots, \phi_n(x))$ that are polynomials in $\mathbf{x} = (x_1, \cdots, x_m)$, this open set is dense in $\mathbb{F}^m$ and its complement has Lebesgue measure zero – i.e. maximal (constant) rank is achieved at "almost all" points in $\mathbb{F}^m$.

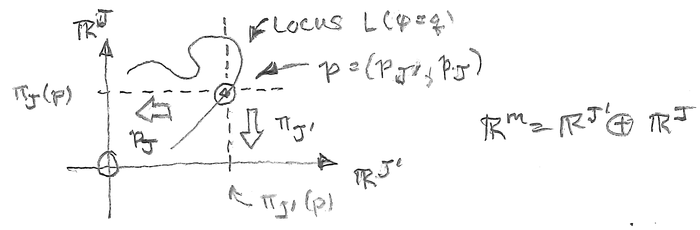

Figure 10.2. The projection maps $\pi_{J'}, \pi_J$ and direct sum decomposition $\mathbb{F}^m = \mathbb{F}^{J'} \oplus \mathbb{F}^J$ associated with a partition of column indices $[1, m] = J' \cup J$, $J' \cap J = \emptyset$. Near any base point $p \in M$ this splits the variables in $\mathbf{x} = (x_1, \ldots, x_m)$ into two groups so $\mathbf{x} = \left(x_{J'}, x_J\right)$ with $x_J = (x_{k_1}, \ldots, x_r)$ and $x_{J'} = (x_{\ell_1}, \ldots, x_{\ell_s})$, where $r = |J|, s = |J'|$, and $r+s = m = \dim(\mathbb{F}^m)$. In our narrative such partitions of variables arise in discussing the rank $r = \mathrm{rk}(d\phi)_p$ of the the $n \times m$ Jacobian matrix $[\partial f_i / \partial x_j]$ of a differentiable map $\phi : \mathbb{F}^m \rightarrow \mathbb{F}^n$ at points on a typical level set $L(\phi = q)$.

3. For any choice of indices I, J with $|I| = |J| = k \leq \min\{m, n\}$ define $J' = [1, m] \sim J$ and let $\mathbb{F}^J$, $\mathbb{F}^{J'} \subseteq \mathbb{F}^m$ be the subspaces

$$\mathbb{F}^J = \mathbb{R}\text{-span}\{e_i : i \in J\}, \qquad \mathbb{F}^{J'} = \mathbb{R}\text{-span}\{e_i : i \in J'\} \ ,$$

where $\{e_i\}$ is the standard basis in $\mathbb{F}^m$. Then $\mathbb{F}^m$ splits as a direct sum $\mathbb{F}^{J'} \oplus \mathbb{F}^J$, and this decomposition determines projections

$$\pi_{J'} : \mathbb{F}^m \rightarrow F^{J'} \quad \text{and} \quad \pi_J : \mathbb{F}^m \rightarrow \mathbb{F}^J$$

onto these subspaces.

Note: By composing ϕ with translations in $\mathbb{F}^m$ and $\mathbb{F}^n$ we can assume ϕ maps the origin in $\mathbb{F}^m$ to the origin in $\mathbb{F}^n$. This will not change $\mathrm{rk}(d\phi_p)$, but greatly simplifies the notation. The following exercise shows that this maneuver does not affect the Jacobian matrices or their determinants. $\square$

1.2. Exercise. If $p \in \mathbb{R}^m$

(a) Consider a translation operator $\mathbf{y} = \phi(\mathbf{x}) = (x_1 + p_1, \ldots, x_m + p_m)$ from $\mathbb{R}^m \rightarrow \mathbb{R}^m$. Prove that $(d\phi)_p = I_{m \times m}$ at every base point.

(b) Given smooth maps $\mathbb{R}^m \xrightarrow{\phi} \mathbb{R}^n \xrightarrow{\psi} \mathbb{R}^k$ and base points $p \in \mathbb{R}^m, q = \phi(p) \in \mathbb{R}^n$, explain why the differential of a composite map $\psi \circ \phi : \mathbb{R}^m \rightarrow \mathbb{R}^k$ is the matrix product of their differentials

$$d(\psi \circ \phi)_p = (d\psi)_{\phi(p)} \cdot (d\phi)_p \quad \square$$

Smooth Hypersurfaces and the Implicit Function Theorem. The Implicit Function Theorem (IFT) concerns itself with **level sets**

$$L(\phi = q) = \{x \in \mathbb{F}^m : \phi(x) = q\} \subseteq \mathbb{F}^m \ ,$$

on which a smooth mapping $\phi : \mathbb{F}^m \rightarrow \mathbb{F}^n$ has constant (vector) value $\phi(\mathbf{x}) = q \in \mathbb{F}^n$. In essence, the IFT says that if p lies in a level set $L(\phi = q)$, and if $(d\phi)_{\mathbf{x}}$ has constant rank $= r$ at and near p, then the locus $L(\phi = q)$ can be described locally as a smooth surface of dimension $m - r$ in $\mathbb{F}^m$. That is to say, near p the level set coincides with the graph

$$\Gamma = \{(\mathbf{x}, f(\mathbf{x}) : \mathbf{x} \in \mathbb{F}^{m-r}\} \subseteq \mathbb{F}^m = \mathbb{F}^{m-r} \times \mathbb{F}^r$$

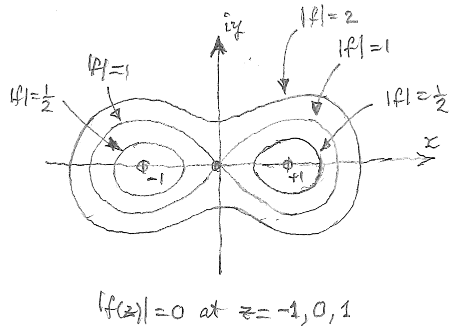

Figure 10.3. Level sets for the map $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ with $f(x, y) = |z^2 - 1|^2$, identifying $z = x + iy$ with $(x, y) \in \mathbb{R}^2$. If $c < 0$ the level set $L_c = L(f = c)$ is empty; when $c = 0$ it consists of three isolated points $z = -1, 0, +1$; and for $c > 0$ the locus is usually a smooth curve (perhaps with more than one connected component, as when $0 < c < 1$). But when $c = 1$ the locus has a singularity at the origin. It cannot be represented near the origin as the graph of any smooth function $y = h(x)$ or $x = g(y)$. The origin is a "branch point" for the locus.

of a smooth map $f : \mathbb{F}^{m-r} \rightarrow \mathbb{F}^r$. The idea is illustrated in the following example (see also Figure 10.2.) The map $\mathbf{y} = f(\mathbf{x})$ is the "implicit function" of the IFT.

1.3. Example. Define $\phi(z) = |z^2 - 1|^2$ ifor $z \in \mathbb{C}$ and regard it as a map $\mathbb{R}^2 \rightarrow \mathbb{R}$ by identifying $z = x+iy \in \mathbb{C}$ with $\mathbf{x} = (x, y) \in \mathbb{R}^2$. Then ϕ becomes a 4^{th} degree polynomial in x and y,

$$\phi(x, y) = x^4 + 2x^2y^2 + y^4 - 2x^2 + 2y^2 + 1 \ .$$

The level sets $L_c = L(\phi = c)$ are empty if $c < 0$; reduce to the isolated points $\{-1, 0, +1\}$ if $c = 0$; and for $c > 0$ are smooth curves (sometimes with more than one connected component if $0 < c < 1$). However there is one exception. When $c = 1$ the locus $L(\phi = 1)$, shown in Figure 10.3, has a singularity at the origin. Near $z = 0 + i0$ it cannot be described locally as the graph of a smooth function $y = h(x)$ or $x = g(y)$. The 1×2 Jacobian matrix $J\phi(z) = \big[\partial\phi/\partial x(z), \partial\phi/\partial y(z)\big] = (d\phi)_z$ has $\text{rk}(d\phi)_z \equiv 1$ (constant) throughout $\mathbb{R}^2$ except at $z = 0$, $z = -1$ and $z = +1$ on the real axis (the "critical points" where both partial derivatives of f are zero). $\square$

1.4. Exercise. Let $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ be the function in Example 1.3

(a) Verify that the Jacobian matrix

$$(d\phi)_{\mathbf{x}} = \Big[\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y}\Big]$$

has rank zero (both components = 0) if and only if $\mathbf{x} = (-1, 0), (0, 0)$, or $(+1, 0)$ in $\mathbb{R}^2$, by solving the system of equations

$$\frac{\partial\phi}{\partial x}(\mathbf{x}) = 0 \qquad \frac{\partial\phi}{\partial y}(\mathbf{x}) = 0$$

(b) At which points $\mathbf{x}$ is one of the derivatives $\partial\phi/\partial x$ and $\partial\phi/\partial y$ zero, while the other is nonzero? Draw pictures of the sets

$$\begin{aligned} S_1 &= \{\mathbf{x} : \partial\phi/\partial x = 0 \text{ and } \partial\phi/\partial y \neq 0\} \\ S_2 &= \{\mathbf{x} : \partial\phi/\partial x \neq 0 \text{ and } \partial\phi/\partial y = 0\} \end{aligned}$$

(c) Verify that $\mathrm{rk}(d\phi)_{\mathbf{x}} = 1$ at all points of S_1 and S_2 identified in (b). How are these points related to the pattern of level curves shown in Figure 10.3?

Note: The point sets in (b) are infinite. $\square$

In discussing the IFT we will discover that a level set $L_c = L(\phi = c)$ in Example 1.3 can be described as the graph of a smooth function $y = f(x)$ near any point p on the locus where $\partial\phi/\partial x(p) \neq 0$, and similarly we can write $x = g(y)$ if $\partial\phi/\partial y(p) \neq 0$. In Example 1.3, at least one of these condition is satisfied at every base point, except the origin $\mathbf{x} = (0, 0)$, which lies on the locus $L(\phi = 1)$, and the points $\mathbf{x} = (-1, 0)$ and $(1, 0)$ which make up the degenerate locus $L(\phi = 0)$. Consequently, for $c \neq 0$ or 1 the non-empty level curves $L_c = L(\phi = c)$ can be described locally as smooth curves (the graphs of smooth functions $y = f(x)$ or $x = g(y)$). Furthermore, L_c can be described *both* ways (with $y = f(x)$ or with $x = g(y)$) near most points $p \in L_c$, but at a few points only one such description is possible – these are the points on the curves in Figure 10.3 at which L_c has either a horizontal or vertical tangent line.

We will apply the IFT to show that the "classical matrix groups" $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{U}(n)$, $\mathrm{SO}(n, \mathbb{C})$, etc. are actually smooth "hypersurfaces" in matrix space $\mathrm{M}(n, \mathbb{F}) \simeq \mathbb{F}^{n^2}$, and hence have well-defined "dimensions," "tangent spaces," etc.

If $\phi : \mathbb{F}^m \rightarrow \mathbb{F}^n$ is a $\mathcal{C}^\infty$ map and if $\mathrm{rk}(d\phi)_{\mathbf{x}} \equiv r$ (constant) on some open neighborhood of p in $\mathbb{F}^m$, the IFT asserts that the level set $S_p = \{\mathbf{x} \in \mathbb{F}^m : \phi(\mathbf{x}) = \phi(p)\}$ passing through p can be described near p as a smooth hypersurface of dimension $m - r$ in $\mathbb{F}^m$. For simplicity we state the result taking $\mathbb{F} = \mathbb{R}$, though it remains true almost verbatim for $\mathbb{F} = \mathbb{C}$.

1.5. Theorem (Implicit Function Theorem). *Let $\phi : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a C^∞ map defined near $p \in \mathbb{R}^m$ and let $M = L(\phi = \phi(p))$ be the level set containing p. Assume $\mathrm{rk}(d\phi)_{\mathbf{x}} \equiv r$ (constant) for $\mathbf{x}$ near p in $\mathbb{R}^m$. Given index sets $I = \{i_1 < \cdots < i_r\} \subseteq [1, n]$, $J = \{j_1 < \cdots < j_r\} \subseteq [1, m]$ such that the square submatrix $[(d\phi)_p]_{IJ}$ is non singular, let $J' = [1, m] \sim J$ and let $\pi_{J'}, \pi_J$ be the projections of $\mathbb{R}^m$ onto $\mathbb{R}^{J'}, \mathbb{R}^J$ associated with the decomposition $\mathbb{R}^m = \mathbb{R}^{J'} \oplus \mathbb{R}^J$, in which $|J| = r$, $|J'| = m - r$. Then there is an open rectangular neighborhood $B_1 \times B_2$ of p in $\mathbb{R}^m = \mathbb{R}^{J'} \oplus \mathbb{R}^J$ such that*

1. *On the relatively open neighborhood $U_p = (B_1 \times B_2) \cap M$ of p in M, the restriction $\pi_{J'}|_{U_p} : U_p \rightarrow B_1$ of the linear projection $\pi_{J'} : \mathbb{R}^m \rightarrow \mathbb{R}^{J'} \cong \mathbb{R}^{m-r}$ is a bicontinuous bijection between the open set $U_p \subseteq M$ and the open set $B_1 \subseteq \mathbb{R}^{m-r}$. It assigns unique Euclidean coordinates $\mathbf{x} = (x_1, \cdots, x_{m-r})$ to every point in U_p.*

2. *The inverse map*
$$\Psi = \left(\pi_{J'}|_{U_p}\right)^{-1} : B_1 \rightarrow U_p \subseteq \mathbb{R}^m$$
is a $\mathcal{C}^\infty$ map from the open set $B_1 \subseteq \mathbb{R}^{J'}$ into all of $\mathbb{R}^m$, and maps B_1 onto the relatively open neighborhood U_p of p in the level set M.

Then the map $f : \mathbb{R}^r \rightarrow \mathbb{R}^n$ obtained by following Ψ with the "horizontal" projection π_I shown in Figure 10.4
$$f = \pi_I \circ \Psi : B_1 \rightarrow B_2 \subseteq \mathbb{R}^n$$
is a $\mathcal{C}^\infty$ map. Furthermore Ψ is the graph map for the smooth function f because
$$\Psi(x) = \left(\pi_{J'}(\Psi(x)), \pi_I(\Psi(x))\right) = \left(x, \pi_I \circ \Psi(x)\right) = (x, f(x)) \qquad \textit{for } x \in B_1 .$$
In particular the open neighborhood U_p in M is the graph of the $\mathcal{C}^\infty$ map $f : \mathbb{R}^{J'} \rightarrow \mathbb{R}^I$.

Conclusion: Near p the locus $M = L(\phi = \phi(p))$ passing through p looks like part of a

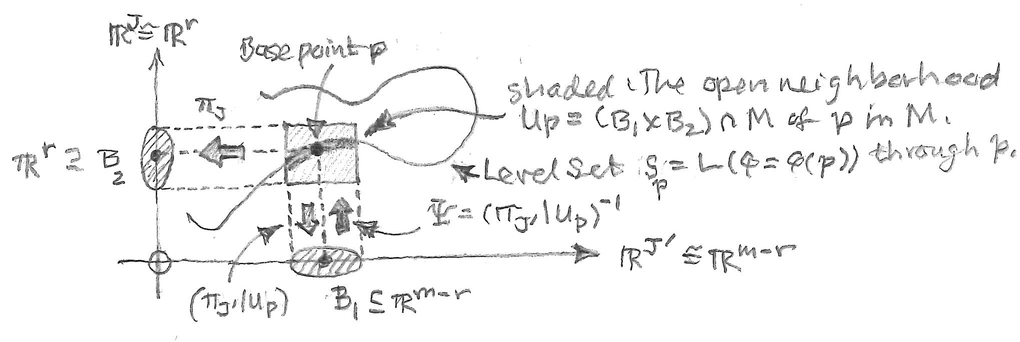

Figure 10.4. A diagram showing the players in the Implicit Function Theorem. Here S_p is the level set passing through p for a $\mathcal{C}^\infty$ map $\phi : \mathbb{R}^m \to \mathbb{R}^n$. By splitting coordinates in $\mathbb{R}^m$ into two groups we get a decomposition $\mathbb{R}^m = \mathbb{R}^{J'} \oplus \mathbb{R}^J$ and associated projections $\pi_{J'}, \pi_J$ from $\mathbb{F}^m$ to $\mathbb{F}^{J'}, \mathbb{F}^J$ such that (i) the restriction $\pi_{J'}|_{U_p}$ becomes a *bijective* bicontinuous map to an open set B_1 in $\mathbb{R}^{J'}$ for a suitably chosen open neighborhood U_p of p in S_p, and (ii) the inverse $\Psi = \left(\pi^{J'}|_{U_p}\right)^{-1} : B_1 \to U_p$ is a $\mathcal{C}^\infty$ map from the open set $B_1 \subseteq \mathbb{R}^{J'}$ into the entire Euclidean space $\mathbb{R}^m$ in which the level set S_p lives.

smooth hypersurface in $\mathbb{R}^m$ of dimension $k = m - n$. The situation described in the IFT is shown in Figure 10.4.

A rough general principle is at work here. If $\phi : \mathbb{R}^m \to \mathbb{R}$ is a scalar valued $\mathcal{C}^\infty$ function we often find that the solution set $L(\phi = c),\ c \in \mathbb{R}$, is a smooth hypersurface of dimension $m - 1$. A level set of a vector valued map $\phi : \mathbb{R}^m \to \mathbb{R}^n$ with $\phi = (\phi_1, \ldots, \phi_n)$ is the intersection of the solution sets for a system of scalar constraint equations

$$\phi_1(\mathbf{x}) = c_1, \quad \ldots \quad , \phi_n(\mathbf{x}) = c_n$$

The solution set tends to lose one degree of freedom for each imposed constraint, so the outcome is usually a smooth hypersurface in $\mathbb{R}^m$ of dimension $m - n$, but that is not always the case and the point of the IFT is to make clear when it is true. This principle also suggests why it is often natural to restrict attention to the case $m \geq n$, in which "maximal rank" means $\mathrm{rk}(d\phi)_p = n$. If the number of constraints n exceeds the dimension m of the space $\mathbb{R}^m$ in which the level set lives, the locus may be degenerate with solutions at all, or it may reduce to a set of isolated points in $\mathbb{R}^m$. $\square$

The "Maximal Rank" Case: As a particular example, if ϕ maps $\mathbb{F}^m \to \mathbb{F}^n$ and $m \geq n$, the maximum possible value for the rank of $(d\phi_p)$ is n. If this maximal rank is achieved at some base point $p \in \mathbb{F}^m$, ϕ will automatically have the same (maximal) rank at all points $\mathbf{x}$ near p in $\mathbb{F}^m$. The maximal rank case is often encountered, but the IFT is proved in the more general "*constant rank*" case, in which we do not assume $m \geq n$, or that the "constant rank" is the maximum possible rank of $(d\phi)_p$ on all of $\mathbb{F}^m$. $\square$

1.6. Exercise. Consider the $\mathcal{C}^\infty$ map $\phi : \mathbb{R}^4 \to \mathbb{R}^2$ given by

$$\mathbf{y} = f(\mathbf{x}) = \phi(x_1, x_2, x_3, x_4) = \big(y_1(\mathbf{x}), y_2(\mathbf{x})\big) = \big(x_1^2 + x_2^2\,,\, x_3^2 - x_4^2 + x_1x_4\big)$$

(a) Show that the locus $M = L(\phi = q)$ can be described as a smooth two-dimensional hypersurface in $\mathbb{R}^4$ near $p = (1, 2, -1, 3)$, at which $q = \phi(p) = (5, -5)$. Identify all pairs of variables x_i, x_j, $(1 \leq i < j \leq 4)$ that can be used to smoothly parametrize this hypersurface near p

(b) Is this locus smooth near *all* of its points?

Hint: In (b), start by showing $x_4 \neq 0$ for every point $\mathbf{x} \in M$, so you can assume $x_4 \neq 0$ in calculations involving points on this locus, even if you cannot draw a picture. In answering (b) you will have to compute $\mathrm{rk}(A_{IJ})$ for various square submatrices of the 2×4 Jacobian matrix $\big[\partial y_i/\partial x_j\big]$, which has variable coefficients. Do this using symbolic row operations. $\square$

1.7. Exercise. Consider the $\mathcal{C}^\infty$ scalar-valued function $f : \mathbb{R}^2 \to \mathbb{R}$,

$$\phi(x, y) = x^3 y + 2e^{xy}$$

(a) Find all *critical points*, where both partial derivatives $\partial\phi/\partial x$ and $\partial\phi/\partial y$ are zero.

At a critical point p there is *no* way to represent the level set $S_p = L(\phi = \phi(p))$ passing through p as the graph of a smooth function $y = f(x)$ or $x = g(y)$.

(b) Locate all points p where one of the partial derivatives is zero but the other is not (two cases to consider). Find the value of ϕ at each such base point to determine which level sets $L(\phi = c), c \in \mathbb{R}$, contain such points.

(c) The locus $M = L(\phi = 2)$ obviously contains the horizontal and vertical axes. Prove that there are no other points on this locus. (Thus the origin is a singularity for the locus M, and there are no others.)

Hint: In (c): Quadrant-by-quadrant, what is the sign of $\partial\phi/\partial y$ off of the x- and y-axes? $\square$

1.8. Exercise. Let $\mathbf{y} = \phi(\mathbf{x}) = \big(y_1(\mathbf{x}), y_2(\mathbf{x})\big)$ be a $\mathcal{C}^\infty$ map from $\mathbb{R}^3 \to \mathbb{R}^2$ such that

$$\begin{array}{ll} (i) & \phi(p) = q = (0,0) \text{ at base point } p = (3,-1,2) \\ (ii) & \text{At } p \text{ the Jacobian matrix is } \big[\partial y_i/\partial x_j\big] = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \end{pmatrix} \end{array}$$

Answer the following questions without knowing anything more about ϕ.

(a) Can the level set $M = L(\phi = q)$ be described near $p = (3, -1, 2)$ as a smooth hypersurface in $\mathbb{R}^3$? Of What dimension k?

(b) Which of the variables x_1, x_2, x_3 can be legitimately be used to parametrize M near $p = (3, -1, 2)$ as the graph of a smooth map $f : \mathbb{R}^k \to \mathbb{R}^3$? List all valid choices of the parametrizing variables $x_{i_1}, \ldots, x_{i_k}$. $\square$

1.9. Exercise. If $f : \mathbb{F}^r \to \mathbb{F}^s$ is a $\mathcal{C}^\infty$ map defined on open set $B \subseteq \mathbb{F}^r$, its **graph** $\Gamma = \{(\mathbf{x}, f(\mathbf{x})) \in \mathbb{F}^r \times \mathbb{F}^s : \mathbf{x} \in B\}$ is the range of the **graph map** $F(\mathbf{x}) = (\mathbf{x}, f(\mathbf{x}))$ from $\mathbb{F}^r \to \mathbb{F}^{r+s}$ Show that the graph map is $\mathcal{C}^\infty$ for $\mathbf{x} \in B \subseteq \mathbb{F}^r$ and that $\mathrm{rk}(df)_x \equiv r$ (constant) for all $\mathbf{x} \in B$. $\square$

Smooth Submanifolds in $\mathbb{F}^m$. A space M is **locally Euclidean** of dimension d if it can be covered by a family of charts $\{(x_\alpha, U_\alpha) : \alpha \text{ in some index set } I\}$, where $x_\alpha : U_\alpha \to V_\alpha \subseteq \mathbb{R}^d$ is a bicontinuous map from an open subset $U_\alpha \subseteq M$ to an open set $V_\alpha = x_\alpha(U_\alpha)$ in the Euclidean space $\mathbb{R}^d$. The "chart maps" x_α assign locally defined Euclidean coordinates $x_\alpha(u) = \big(x_1^{(\alpha)}(u), \ldots, x_k^{(\alpha)}(u)\big)$ for $u \in U_\alpha$. Thus M looks locally like Euclidean coordinate space $\mathbb{R}^d$.

Where the domains of two charts $(x_\alpha, U_\alpha), (x_\beta, U_\beta)$ overlap we have the situation shown in Figure 10.5. The intersection $U_\alpha \cap U_\beta$ is an open set in M, the images $N_\alpha = x_\alpha(U_\alpha \cap U_\beta)$, $N_\beta = x_\beta(U_\alpha \cap U_\beta)$ are open sets in coordinate space $\mathbb{R}^k$, and we have

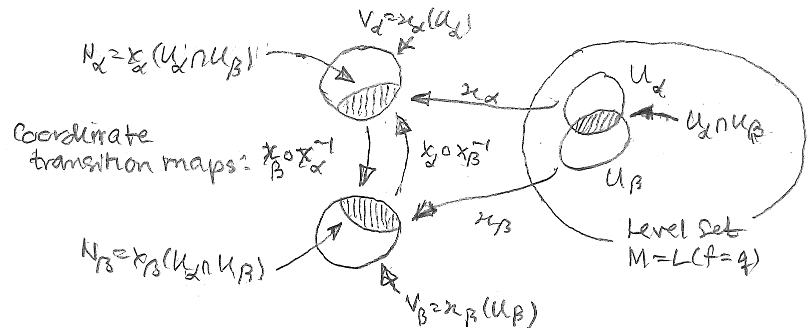

Figure 10.5. The coordinate transition maps $x_\alpha \circ x_\beta^{-1}$ and $x_\beta \circ x_\alpha^{-1}$ between two charts $(x_\alpha, U_\alpha), (x_\beta, U_\beta)$ and their (shaded) domains of definition in $\mathbb{R}^r$ are shown. Both shaded domains N_α, N_β in $\mathbb{R}^r$ correspond to the intersection $U_\alpha \cap U_\beta$ of the chart domains, which is an open set in the locally Euclidean space M.

induced **coordinate transition maps** that tell us how the coordinates $\mathbf{x} = x_\alpha(u)$ and $\mathbf{y} = x_\beta(u)$ assigned to $u \in M$ by the chart maps are related. These transition maps

$$\begin{aligned} \mathbf{x} &= \big(x_\alpha \circ x_\beta^{-1}\big)(\mathbf{y}) \qquad \text{from } N_\beta \rightarrow N_\alpha \\ \mathbf{y} &= \big(x_\beta \circ x_\alpha^{-1}\big)(\mathbf{x}) \qquad \text{from } N_\alpha \rightarrow N_\beta \end{aligned}$$

are bicontinuous bijections between the open sets N_α and N_β in $\mathbb{R}^k$

1.10. Definition. *A locally Euclidean space M is a* **smooth manifold** *of dimension* $\dim(M) = d$ *if the charts (U_α, x_α) that cover M map it into $\mathbb{R}^d$ and are* $\mathcal{C}^\infty$**-related**, *so the coordinate transition maps are $\mathcal{C}^\infty$ between the open sets $N_\alpha, N_\beta \subseteq \mathbb{R}^d$ that correspond to the* (*open*) *intersection $U_\alpha \cap U_\beta$ of chart domains in M.*

This allows us to make sense of "smooth manifolds" without requiring that they be embedded in some surrounding Euclidean space. IT is the starting point for modern differential geometry.

Once we make M a $\mathcal{C}^\infty$ manifold by introducing $\mathcal{C}^\infty$ related covering charts, we can begin to do Calculus on M. The following concepts now make sense:

1. Given any chart (x_α, U_α) on M, a scalar function $f : M \rightarrow \mathbb{F}$ on M becomes a function of local chart coordinates if we write

$$y = F(\mathbf{x}) = \big(f \circ x_\alpha^{-1}\big)(\mathbf{x}) \ ,$$

   which is defined on the open set $V_\alpha = x_\alpha(U_\alpha)$ in coordinate space $\mathbb{F}^m$. We say that f is a $\mathbf{C}^\infty$ **function on** M if $y = F(\mathbf{x})$ has continuous partial derivatives of all orders, for each of the covering charts that determine the manifold structure of M.

2. A map $\phi : M \rightarrow N$ is a $\mathbf{C}^\infty$ **mapping** between manifolds M and N of dimensions m and n if it becomes a $\mathcal{C}^\infty$ map from $\mathbb{F}^m \rightarrow \mathbb{F}^n$ when described in local coordinates on M and N. Thus if (x_α, U_α) and (y_β, U_β) are charts on M and N respectively, the composite $\mathbf{y} = \Phi(\mathbf{x}) = y_\beta \circ \phi \circ x_\alpha^{-1}(\mathbf{x})$ is a $\mathcal{C}^\infty$ map from $\mathbb{F}^m \rightarrow \mathbb{F}^n$ wherever it is well-defined.

3. A *parametric curve* in M is any continuous map $y = \gamma(t)$ from some interval $[a, b] \subseteq \mathbb{R}$ into M. It is a $\mathbf{C}^\infty$**-curve** in M if it becomes a $\mathcal{C}^\infty$ vector-valued map

$$\mathbf{x} = \big(y_1(t), \ldots, y_m(t)\big) = x_\alpha \circ \gamma(t) \quad \text{ for } t \in [a, b]$$

for every chart (x_α, U_α) on M.

1.11. Definition. Suppose $M \xrightarrow{h} N \xrightarrow{f} R$ are $\mathcal{C}^\infty$ maps between $\mathcal{C}^\infty$ manifolds. In terms of the preceding definitions, explain why the composite $f \circ h : M \to R$ is a $\mathcal{C}^\infty$ map wherever it is well-defined.
Hint: If M, N, R are Euclidean coordinate spaces $\mathbb{F}^m, \mathbb{F}^n, \mathbb{F}^r$ this follows by the Chain Rule of multivariate Calculus. $\square$

Smooth Manifolds and the IFT. One consequence of the IFT is this: If ϕ is a $\mathcal{C}^\infty$ map $\mathbb{R}^m \to \mathbb{R}^n$ and if $\mathrm{rk}(d\phi)_{\mathbf{x}} \equiv r$ (constant) near *every* point in a level set $M = L(\phi = q), q \in \mathbb{R}^n$, we can use the IFT to create a family of $\mathcal{C}^\infty$-related charts (x_α, U_α) that cover M. The resulting **standard $\mathcal{C}^\infty$ structure** makes M into a smooth r-dimensional manifold. The crucial fact that the charts are $\mathcal{C}^\infty$-related follows directly from the way the standard charts are constructed (see Proposition 1.13 below).

1.12. Constructing "Standard Charts" on a Level Set M. The IFT and the "constant rank" condition allow us to construct a chart (x_α, U_α) about a typical base point $p \in M$.

1. Write ϕ as $\mathbf{v} = \phi(\mathbf{u})$ in terms of the standard coordinates $\mathbf{u} = (u_1, \ldots, u_m)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ in $\mathbb{R}^m$ and $\mathbb{R}^n$. By Lemma 1.1 and the "constant rank" condition we can, for each $\mathbf{u} \in M$, choose row and column indices $I \subseteq [1, n]$ and $J \subseteq [1, m]$ with $|I| = |J| = r = \mathrm{rk}(d\phi)_{\mathbf{u}}$ such that the square submatrices $\big[\partial v_i/\partial u_j(\mathbf{u})\big]_{IJ}$ are nonsingular for $\mathbf{u}$ near p in $\mathbb{R}^m$. By Lemma 1.1 this cannot be done for any larger square submatrix.

2. Using the column indices determined in Step 1, let $J' = [1, m] \sim J$, split $\mathbb{R}^m = \mathbb{R}^{J'} \oplus \mathbb{R}^J$ and let $\pi_{J'}, \pi_J$ be the projection maps from $\mathbb{R}^m$ to $\mathbb{R}^{J'}$ or $\mathbb{R}^J$. By the IFT there is a rectangular open neighborhood $B_1 \times B_2$ of p in $\mathbb{F}^m$ such that the projection
$$\pi_{J'} : (B_1 \times B_2) \to B_1$$
maps the relatively open neighborhood $U_p = (B_1 \times B_2) \cap M$ in M onto the open set $B_1 \subseteq \mathbb{R}^{J'} \cong \mathbb{R}^{m-r}$. To get a chart (x_α, U_α) that imposes Euclidean coordinates on M near p we take $U_\alpha = U_p$ and bijective chart map $x_\alpha = \big(\pi_{J'}|_{U_p}\big) : U_p \to B_1$ (an open set in $\mathbb{R}^{m-r}$). The charts (x_α, U_α) obviously cover M owing to constancy of $\mathrm{rk}(d\phi)$ near every point in M.

3. The inverse map $\Psi = \big(\pi_{J'}|_{U_p}\big)^{-1} : B_1 \to U_p \subseteq M \subseteq \mathbb{R}^m$ is $\mathcal{C}^\infty$ from $B_1 \subseteq \mathbb{R}^{J'}$ into all of $\mathbb{R}^m$, and its range is precisely the chart domain U_p. $\square$

We now show that charts created this way, perhaps about different base points, are always $\mathcal{C}^\infty$-related wherever the chart domains overlap.

1.13. Proposition. *If $\phi : \mathbb{R}^m \to \mathbb{R}^n$ is a $\mathcal{C}^\infty$ map and $M = L(\phi = q)$ a level set such that* $\mathrm{rk}(d\phi)_x \equiv r$ *(constant) on an open neighborhood of every point in M, then all standard charts on M obtained by the preceding construction are $\mathcal{C}^\infty$ related where they overlap. This determines the* **standard $\mathcal{C}^\infty$ structure** *on M. The dimension of the resulting $\mathcal{C}^\infty$ manifold is $k = m - r$.*

Proof: Consider two standard charts (x_α, U_α) and (x_β, U_β) about a typical point p in $U_\alpha \cap U_\beta$. The chart (x_α, U_α) is determined by a partition of column indices $[1, m] = J'(\alpha) \cup J(\alpha)$ and a choice of row indices $I(\alpha) \subseteq [1, n]$ with $|J(\alpha)| = r$ and $|J'(\alpha)| = m - r$, such that $[(d\phi)_p]_{IJ}$ is nonsingular. In the notation of the IFT we then have

$$U_\alpha = (B_1^\alpha \times B_2^\alpha) \cap M \qquad x_\alpha = \big(\pi_{J'(\alpha)}|_{U_\alpha}\big) \quad \text{and} \quad V_\alpha = x_\alpha(U_\alpha) = B_1^\alpha \subseteq \mathbb{R}^{m-r}$$

The chart map is the restriction to U_α of a *linear* projection map $\mathbb{R}^m \to \mathbb{R}^{J'(\alpha)} \cong \mathbb{R}^{m-r}$,

$$x_\alpha = \left(\pi_{J'(\alpha)}|_{U_\alpha}\right) : U_\alpha \to B_1^\alpha \subseteq \mathbb{R}^{m-r} \ .$$

Its inverse $\Psi = x_\alpha^{-1}$ is the graph map

$$\Psi_\alpha = \left(\pi_{J'(\alpha)}|_{U_\alpha}\right)^{-1} : V_\alpha \to U_\alpha \ ,$$

which is actually a $\mathcal{C}^\infty$ map from B_1^α into *all of* $\mathbb{R}^m$.

The chart (x_β, U_β) corresponds to some other choice of column and row indices $[1,m] = J'(\beta) \cup J(\beta)$ and $I(\beta) \subseteq [1,n]$, and a corresponding rectangular open neighborhood $B_1^\beta \times B_2^\beta$ of p in $\mathbb{R}^m$. The chart map on $U_\beta = \left(B_1^\beta \times B_2^\beta\right) \cap M$ is just the restriction to U_β of a linear projection $\pi_{J'(\beta)}$, and by the IFT its inverse is a $\mathcal{C}^\infty$ map from B_1^β into all of $\mathbb{R}^m$. Therefore the coordinate transition map

$$x_\beta \circ x_\alpha^{-1} = \pi_{J'(\beta)} \circ \left(\pi_{J'(\alpha)}|_{U_\alpha}\right)^{-1} = \pi_{J'(\beta)} \circ \Psi_\alpha$$

is the composite of a linear map and a $\mathcal{C}^\infty$ map

$$\mathbb{R}^{m-r} \cong \mathbb{R}^{J'(\alpha)} \xrightarrow{\ x_\alpha^{-1}\ } \mathbb{R}^m \xrightarrow{\ x_\beta\ } \mathbb{R}^{J'(\beta)} \cong \mathbb{R}^{m-r}$$

and is certainly $\mathcal{C}^\infty$. Likewise for the transition map in the reverse direction. $\square$

The preceding proof is burdened by the complicated notation needed to label all the players. Here is a shorter proof that emphasizes the intuition behind the proof.

Alternative Proof of Proposition 1.13: Suppose p is any point in $U_\alpha \cap U_\beta$ and $\mathbf{x}_0 = x_\alpha(p), \mathbf{y}_0 = x_\beta(p)$ in $\mathbb{R}^{m-r}$. To show $\mathbf{y} = x_\beta \circ x_\alpha^{-1}(\mathbf{x})$ is $\mathcal{C}^\infty$ near $\mathbf{x}_0$ we observe that

- Near $\mathbf{x}_0$ the chart map x_α^{-1} coincides with the map $\left(\pi_{J'(\alpha)}|_{U_\alpha}\right)^{-1}$, which by the IFT is a $\mathcal{C}^\infty$ map from an open set in $\mathbb{R}^{m-r}$ into all of $\mathbb{R}^m$ that sends $\mathbf{x}_0 \to p$, and whose range is contained in M.

- Near p the chart map x_β coincides with the globally defined linear projection map $\pi_{J'(\beta)}$, which is certainly $\mathcal{C}^\infty$.

Therefore the transition map $x_\beta \circ x_\alpha^{-1}$ is the composite of a linear map and a $\mathcal{C}^\infty$ map

$$\mathbb{R}^{m-r} \cong \mathbb{R}^{J'(\alpha)} \xrightarrow{x_\alpha^{-1}} \mathbb{R}^m \xrightarrow{x_\beta} \mathbb{R}^{J'(\beta)} \cong \mathbb{R}^{m-r}$$

and is $\mathcal{C}^\infty$. Likewise for the transition map in the reverse direction. $\square$

1.14. Example. Let $\phi : \mathbb{R}^3 \to \mathbb{R}^1$ with $\phi(\mathbf{x}) = x_3^2 - x_1^2 - x_2^2$. At any $p = (x_1, x_2, x_3)$ the 1×3 Jacobian matrix

$$(d\phi)_p = \Big[\frac{\partial \phi}{\partial x_1}, \cdots, \frac{\partial \phi}{\partial x_3}\Big] = \big[-2x_1 \, , \, -2x_2 \, , \, 2x_3 \big]$$

is just the classical "gradient" vector $\nabla\phi(p)$. The rank $\mathrm{rk}(d\phi)_{\mathbf{x}}$ is constant $\equiv 1$ unless all three entries are zero, which happens only at the origin $p = (0,0,0)$. The level set $M_0 = L(\phi = 0)$ is the double cone shown in Figure 10.6(b). This two-dimensional hypersurface has a singularity at the origin in $\mathbb{R}^3$, where it fails to be locally Euclidean. Thus the locus $L(\phi = 0)$ cannot be made into a smooth manifold by covering it with suitably defined coordinate charts. All other level sets M_c $(c \neq 0)$ are smooth two-dimensional manifolds; a few of these level surfaces are shown in Figure 10.6(a).

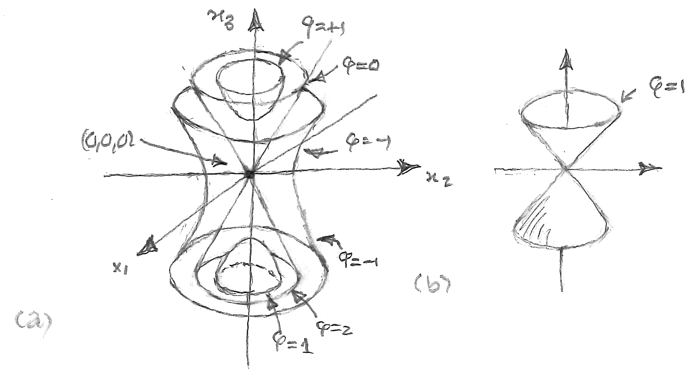

Figure 10.6. In (a) we show some level sets $L_c = L(\phi = c)$ for the map $\phi(x_1, x_2, x_3) = x_3^2 - x_2^2 - x_1^2$ from $\mathbb{R}^3 \to \mathbb{R}^1$. For $c = 0$ the level set where $x_1^2 - x_2^2 - x_3^2 = 0$, shown in (b), is a double cone with a singularity at the origin, where it fails to be locally Euclidean. All other level sets are smooth two-dimensional hypersurfaces in $\mathbb{R}^3$, but the geometry of L_c changes as we pass from $c < 0$ to $c > 0$. For $c < 0$, we get a single connected surface; for $c > 0$ there are two isolated pieces, both smooth.

Consider the possible charts we might impose near the point $p = (1, 1, \sqrt{3})$ on the particular level set $M = L(\phi = 1)$. Entries in $(d\phi)_p$

$$(d\phi)_p = [-2x_1, -2x_2, 2x_3] = (-2\sqrt{2}, -2\sqrt{2}, 2\sqrt{3}) \qquad \text{at } p\ ,$$

are all nonzero near p, so we have constant rank $\mathrm{rk}(d\phi)_{\mathbf{x}} \equiv 1$ near p,]and may apply the IFT to define standard charts about p. Each nonzero entry in $(d\phi)_p$ corresponds to a nonsingular 1×1 submatrix, so several legitimate groupings of variables are available to parametrize M near p:

$$(49) \qquad I = \{1\}, J = \{2, 3\} \quad \text{or} \quad I = \{2\}, J = \{1, 3\} \quad \text{or} \quad I = \{3\}, J = \{1, 3\}$$

Thus $L(\phi = 1)$ can be described as the graph in $\mathbb{R}^3$ of various smooth functions $x_k = f_k(x_i, x_j)$ by solving

$$1 = \phi(\mathbf{x}) = x_3^2 - x_2^2 - x_1^2$$

for one variable in terms of the other two.

1. $x_1 = f_1(x_2, x_3) = +\sqrt{x_3^2 - x_2^2 - 1}$ near $(1, \sqrt{3})$ in the (x_2, x_3)-plane.
2. $x_2 = f_2(x_1, x_3) = +\sqrt{x_3^2 - x_1^2 - 1}$ near $(1, \sqrt{3})$ in the (x_1, x_3)-plane.
3. $x_3 = f_3(x_1, x_2) = +\sqrt{1 + (x_1^2 + x_2^2)}$ near $(1, 1)$ in the (x_1, x_2)-plane.

The coordinate transition map $(x_1, x_3) = y_\beta \circ x_\alpha^{-1}(x_2, x_3)$ can be computed directly by writing $x_1 = f_1(x_2, x_3)$ to get (x_1, x_3) in terms of (x_2, x_3). The resulting transition map

$$\begin{aligned} (x_1, x_3) \;&=\; \Phi(x_2, x_3) \;=\; y_\beta \circ x_\alpha^{-1}(x_2, x_3) \\ &=\; (x_1, x_3)\Big|_{x_1 = f_1(x_2, x_3)} \\ &=\; \Big(+ \sqrt{x_3^2 - x_2^2 - 1}\,,\, x_3 \Big) \end{aligned}$$

is clearly a $\mathcal{C}^\infty$ map from (x_2, x_3) to (x_1, x_3). So is its inverse. $\square$

1.15. Exercise. In Example 1.14,

(a) Compute the inverse $(x_2, x_3) = \Phi^{-1}(x_1, x_3)$.

(b) One of the valid splittings $[1,3] = J' \cup J$ of column indices listed in (49) is $J' = \{1,2\}$ $J = \{3\}$. Find an explicit formula for the corresponding projection map $(x_1, x_2) = \pi_{_{J'}}(x_1, x_2, x_3)$ that assigns Euclidean coordinates to points $\mathbf{x} = (x_1, x_2, x_3)$ on M near $p = (1, 1, \sqrt{3})$.

(c) Give an explicit formula for the inverse

$$(x_1, x_2, x_3) = \big(\pi_{_{J'}}|_{\mathbb{R}^{J'}}\big)^{-1}(x_1, x_3)$$

of the projection map in (b). $\Box$

1.16. Exercise. Consider the points

$$\text{(a)}\;\; p = \big(1, 0, \sqrt{2}\big) \qquad\qquad \text{(b)}\;\; p = (0, 0, -1)$$

on the two-dimensional hypersurface $M = L(\phi = 1)$ of Example 1.14. In each case determine all pairs of coordinates $x_{_{J'}} = (x_i, x_j)$ that give a legitimate parametrization of M near the prescribed base point p. $\Box$

1.17. Exercise. Verify that the unit sphere $S^2 = L(\phi = 1)$ for $\phi(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$ is a $\mathcal{C}^\infty$ manifold in $\mathbb{R}^3$ by showing that $\mathrm{rk}(d\phi)_{\mathbf{x}} \equiv 1$ near *every* point $\mathbf{x} \in S^2$. $\Box$

1.18. Exercise. Describe a set of standard charts covering the unit sphere $S^2 = L(\phi = 1)$ where $\phi(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2$, taking for your chart domains the relatively open hemispheres (boundary circles excluded)

$$U_k^+ = \{\mathbf{x} \in S^2 : x_k > 0\} \qquad U_k^- = \{\mathbf{x} \in S^2 : x_k < 0\}$$

for $k = 1, 2, 3$. All six hemispheres are required to fully cover S^2.

The chart maps $x_k^\pm : U_k^\pm \to \mathbb{R}^2$ project points $\mathbf{x} \in U_k^\pm$ onto the open unit disc $x_1^2 + x_3^2 < 1$ in the (x_2, x_3)-plane when $k = 1$; project $U_2^\pm$ onto the open disc in the (x_1, x_3)-plane when $k = 2$; and project onto the disc in the (x_1, x_2)-plane when $k = 3$.

(a) Give explicit formulas for the chart maps on the particular domains U_1^+ and U_3^-.

(b) Compute the coordinate transition maps in both directions for these two charts, noting that they have the form $(x_i, x_j) = x_\alpha(\mathbf{x}) = x_\alpha(x_1, x_2, x_3)$ for $\mathbf{x} \in M$.

Note: These are examples of standard charts on the level set $L(\phi = 1)$. $\Box$

1.19. Exercise (Stereographic Projection). Let H^+ be the two-dimensional hyperplane in $\mathbb{R}^3$ that is tangent to the unit sphere $M = S^2$ at its "north pole" $N = (0, 0, +1)$, and consider the "*punctured sphere*" $U_\alpha = S^2 \sim \{S\}$ obtained by deleting the south pole $S = (0, 0, -1)$ from the sphere. Each point $\mathbf{u} \in U_\alpha$ determines a unique straight line in $\mathbb{R}^3$ that passes through S and the point $\mathbf{u}$; continuing along this line, we will meet the hyperplane H^+ in a unique point with coordinates $\big(x(\mathbf{u}), y(\mathbf{u}), +1\big)$. The resulting bijection $\Phi^+ : U_\alpha \to H^+$ is an example of *stereographic projection.* Dropping the redundant coordinate entry "1" we obtain the **stereographic projection map** $x_\alpha : U_\alpha \to \mathbb{R}^2$,

$$x_\alpha(\mathbf{u}) = \big(x_1(\mathbf{u}), x_2(\mathbf{u})\big) \in \mathbb{R}^2$$

which is is bicontinuous from the open subset $U_\alpha^+ \in S^2$ onto all of coordinate space $\mathbb{R}^2$.

Similarly we may stereographically project the punctured sphere $U_\beta = S^2 \sim \{N\}$ onto the hyperplane H^- tangent to the sphere at the south pole $S = (0, 0, -1)$, to define a second chart map $x_\beta(\mathbf{v}) = (x', y') \in \mathbb{R}^2$ for $\mathbf{v} \in U_\beta$.

(a) Give an explicit formula for the stereographic projection map $(x_1, x_2) = x_\alpha(\mathbf{u}) = x_\alpha(u_1, u_2, u_3)$. Note carefully that Φ^+ maps triples $\mathbf{u}$ with $u_1^2+u_2^2+u_3^3 = 1$ to pairs $(x_1, x_2) \in \mathbb{R}^2$.

(b) Compute the coordinate transition map $\mathbf{v} = x_\beta \circ x_\alpha^{-1}(\mathbf{u})$ and its inverse, and check that it is $\mathcal{C}^\infty$ where defined. $\square$

Stereographic projection allows us to cover the sphere S^2 with just two $\mathcal{C}^\infty$-related charts, the minimum number possible since it is well known that S^2 cannot be mapped bicontinuously to the plane $\mathbb{R}^2$. But for many purposes the covering with hemispheres leads to simpler computations.
Note: These charts are not of the standard form in Proposition 1.12 but they are $\mathcal{C}^\infty$-related to all standard charts (which by 1.13 are $\mathcal{C}^\infty$-related to each other).
Hint: In (a) use similar triangles, and rotational symmetry of the problem. $\square$

X.2. Matrix Lie Groups.[1]

The classical groups are the level sets of certain polynomial maps $\mathbb{F}^m \to \mathbb{F}^n$, except for the general linear group $\mathrm{GL} = \mathrm{GL}(n, \mathbb{F}) = \{A \in \mathrm{M}(n, \mathbb{F}) : \det(A) \neq 0\}$, which is an open subset in matrix space $\mathrm{M}(n, \mathbb{F}) \simeq \mathbb{F}^{n^2}$. This is a smooth manifold and it is covered by a single chart with chart domain $U_\alpha = \mathrm{GL}$ and chart map $x_\alpha =$ the identity map of $\mathrm{GL}(n, \mathbb{F}) \to \mathbb{F}^{n^2}$, which we shall write as

$$x_\alpha(A) = (A_{11}, \ldots, A_{in}; A_{21}, \ldots, A_{2n}; \ldots; A_{n1}, \ldots, A_{nn})$$

in what follows. Obviously $\dim_{\mathbb{F}}(\mathrm{GL}) = n^2$ since GL is an open set in $\mathrm{M}(n, \mathbb{F})$. All other classical groups are closed lower-dimensional subsets in $\mathrm{GL}(n, \mathbb{F})$ and in $\mathrm{M}(n, \mathbb{F}) \cong \mathbb{F}^{n^2}$.

2.1. Definition. *A smooth manifold G is an* **abstract Lie group** *if*

1. *It is a group under some product operation $P : G \times G \to G$ and under the inversion map $J : G \to G$ that sends $x \to x^{-1}$.*

2. *The product operation and inverse operation are both C^∞ maps.*

In particular if (U_α, x_α), (U_β, y_β) are coordinate charts the product operation becomes a C^∞ map from $\mathbb{F}^m \times \mathbb{F}^m \to \mathbb{F}^m$ when expressed in these coordinates. Thus if $x \in U_\alpha$, $y \in U_\beta$, and (U_γ, z_γ) is a chart containing $z = P(x, y) = x \cdot y$, the composite map

$$z_\gamma \circ P \circ (x_\alpha^{-1} \times y_\beta^{-1}) : \mathbb{F}^m \times \mathbb{F}^m \to \mathbb{F}^n \qquad \textit{is a } \mathcal{C}^\infty \textit{ map.}$$

Similarly, if $z \in U_\alpha$ and $z^{-1} \in U_\beta$,

$$x_\beta \circ J \circ x_\alpha^{-1} : \mathbb{F}^m \to \mathbb{F}^m \qquad \textit{is a } \mathcal{C}^\infty \textit{ map}$$

The dimension $d = \dim_{\mathbb{F}}(G)$ is the dimension of the charts that cover G. If $\mathbb{F} = \mathbb{C}$, we regard $\mathbb{F} \simeq \mathbb{R}^2$ and view G as a real manifold of dimension $\dim_{\mathbb{R}}(G) = 2 \cdot \dim_{\mathbb{C}}(G)$, with charts $\tilde{x}_\alpha : U_\alpha \to \mathbb{R}^{2d}$. $\square$

The general theory of Lie groups has become a vast subject. To keep things simple we restrict attention to **matrix Lie groups**, subsets $G \subseteq \mathrm{M}(n, \mathbb{F})$ such that G is:

(i) A group under matrix multiplication, as in Definition 2.1.

(ii) A $\mathcal{C}^\infty$ manifold (smooth hypersurface) in matrix space, as in Definition 1.10.

[1] That's pronounced "*Lee*" *Groups.* Sophus Lie was a Norwegian mathematician who pioneered the study of these structures toward the end of the 1800s. Esoteric concepts then, they are ubiquitous in modern physics and differential geometry.

The simplest example of a Lie group is $G = (\mathbb{R}^n, +)$ with structure given by the single identity chart $(U_\alpha, x_\alpha) = (\mathbb{R}^n, \mathrm{id})$. Clearly, the $(+)$ operation is a $\mathcal{C}^\infty$ map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, as is the inverse map $J(x) = -x$ on $\mathbb{R}^n$. But this in not a matrix Lie group because its elements are not matrices.

The general linear group $\mathrm{GL}(n, \mathbb{F})$ is a (noncommutive) matrix Lie group of $\dim_{\mathbb{F}}(G) = n^2$. The group operation, expressed in chart coordinates

$$x_\alpha(A) = (A_{11}, \ldots, A_{1n}; \ldots; A_{n1}, \ldots, A_{nn}) \in \mathbb{F}^{n^2} ,$$

is a polynomial map of $\mathbb{F}^{n^2} \to \mathbb{F}^{n^2}$ with

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} ,$$

and inversion is the rational map

$$A^{-1} = \frac{1}{\det(A)} \mathrm{Cof}(A)^{\mathrm{t}}$$

This involves the transpose of the cofactor matrix $\mathrm{Cof}(A)$, whose entries are polynomials in the entries of A; $\det(A)$ is also a polynomial in the entries of A.

Other matrix Lie groups are level sets in matrix space for various $\mathcal{C}^\infty$ (polynomial, actually) maps $\phi : \mathrm{M}(n, \mathbb{F}) \to \mathbb{F}^k$ with $k \leq n^2 = \dim_{\mathbb{F}} \mathrm{M}(n, \mathbb{F})$. We will soon indicate why they are all smooth manifolds. To see that they are also Lie groups we prove:

2.2. Theorem. *Suppose M is the level set through p for some $\mathcal{C}^\infty$ map $\phi : \mathrm{M}(n, \mathbb{F}) \to \mathbb{F}^k$, and that $\mathrm{rk}(d\phi)_x \equiv r$ (constant) on some open neighborhood of each $p \in M$. If M is also a subgroup of $\mathrm{GL}(n, \mathbb{F})$ under matrix multiplication, then M is a Lie group in the standard $\mathcal{C}^\infty$ structure it inherits as a smooth submanifold in matrix space $\mathrm{M}(n, \mathbb{F}) \simeq \mathbb{F}^{n^2}$.*

Proof: If $a, b \in M$, let (U_α, x_α), (U_β, x_β) be standard charts about a, b and let (U_γ, z_γ) be a chart about $c = ab = P(a, b)$. By the IFT x_γ is the restriction to U_γ of a linear projection $\pi_{J'(\gamma)}$ from $\mathbb{F}^{n^2}$ to an open set in $\mathbb{F}^d \cong F^{J'(\gamma)}$. The product map $P(a, b) = a \cdot b$ is defined and $\mathcal{C}^\infty$ on all of matrix space $\mathrm{M}(n, \mathbb{F}) \times \mathrm{M}(n, \mathbb{F}) \to \mathrm{M}(n, \mathbb{F})$; and in fact it is a polynomial map. To verify that it is a $\mathcal{C}^\infty$ map on the manifold G we must show that

$$z_\gamma \circ P \circ \big(x_\alpha^{-1} \times y_\beta^{-1}\big) : \mathbb{F}^d \times \mathbb{F}^d \to \mathbb{F}^d = \big(\pi_{J'(\gamma)}|_{U_\gamma}\big) \circ P \circ \big(x_\alpha^{-1} \times y_\beta^{-1}\big) : \mathbb{F}^d \times \mathbb{F}^d \to \mathbb{F}^d$$

is $\mathcal{C}^\infty$, where $d = \dim_{\mathbb{F}}(G) = n^2 - r$ and $r = \mathrm{rk}(d\phi)$, by breaking this composite into steps

$$\mathbb{F}^{d^2} \cong \mathbb{F}^d \times \mathbb{F}^d \xrightarrow{x_\alpha^{-1} \times y_\beta^{-1}} U_\alpha \times U_\beta \xrightarrow{P} U_\gamma \xrightarrow{z_\gamma = \pi_{J'(\gamma)}} \mathbb{F}^d$$

The matrix product operation $P : \mathbb{F}^{n^2} \times \mathbb{F}^{n^2} \to U_\gamma$ is defined and $\mathcal{C}^\infty$ on all of $\mathrm{M}(n, \mathbb{F})$. In particular points near (a, b) map to points near $a \cdot b$ in U_γ; furthermore $P\big(U_\alpha \times U_\beta\big) \subseteq G$ because G is a group. The map $z_\gamma : U_\gamma \to \mathbb{F}^d$ is the restriction to $U_\gamma \subseteq G$ of a globally defined linear projection map $\pi_{J'(\gamma)} : \mathbb{F}^{n^2} \to \mathbb{F}^{J'(\gamma)} \cong \mathbb{F}^d$, so it is the restriction to U_γ of a globally $\mathcal{C}^\infty$ map. Clearly then,

$$z_\gamma \circ P \circ (x_\alpha^{-1} \times y_\beta^{-1}) = \pi_{J'(\gamma)} \circ P \circ (x_\alpha^{-1} \times y_\beta^{-1}) : \mathbb{F}^{d^2} \to \mathbb{F}^d$$

is $\mathcal{C}^\infty$ too. Similarly, the inversion map $J : G \to G$ is $\mathcal{C}^\infty$ when expressed in standard chart coordinates because $J : \mathrm{GL}(n, \mathbb{F}) \to \mathrm{GL}(n, \mathbb{F})$ is a rational function of matrix coordinates. $\square$

The first example where we really need the Implicit Function Theorem is

2.3. Example. The special linear group $G = \mathrm{SL}(n, \mathbb{F})$ has dimension $n^2 - 1$ if $\mathbb{F} = \mathbb{C}$. But if we regard $\mathbb{F} \cong \mathbb{R}^2$ and $\mathrm{SL}(n, \mathbb{C}) \subseteq \mathrm{M}(n, \mathbb{C}) \cong \mathbb{R}^{2n^2}$, then $\dim_{\mathbb{R}}(G) = 2n^2 - 2$. On the other hand, if $\mathbb{F} = \mathbb{R}$ we have $\dim_{\mathbb{R}} \mathrm{SL}(n, \mathbb{R}) = n^2 - 1$. (Geometrically, we have $\mathrm{SL}(n, \mathbb{R}) = \mathrm{SL}(n, \mathbb{C}) \cap \mathrm{M}(n, \mathbb{R}) + i0$ when we identify $\mathrm{M}(n, \mathbb{C}) = \mathrm{M}(n, \mathbb{R}) + \sqrt{-1}\, \mathrm{M}(n, \mathbb{R})$.) We will restrict attention is the case $\mathbb{F} = \mathbb{R}$.

Discussion: $\mathrm{SL}(n, \mathbb{R})$ is the level set $L(\phi = 1)$ where $\phi : \mathbb{R}^{n^2} \cong \mathrm{M}(n, \mathbb{R}) \rightarrow \mathbb{R}$ is $\phi(A) = \det(A)$. To see that SL is a smooth manifold in matrix space we must show that the $1 \times n^2$ Jacobian matrix has $\mathrm{rk}(d\phi)_X = 1$ near every $X \in \mathrm{SL}(n, \mathbb{R})$. If we write coordinates of A as $(a_{11}, \ldots, a_{1n}; \ldots; a_{n1}, \ldots, a_{nn}) \in \mathbb{R}^{n^2}$, then

$$\phi = \det(A) = \sum_{\sigma \in S_n} sgn(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a_{n,\sigma(n)} \ ,$$

and by the product formula for derivatives we get

$$\begin{aligned} \frac{\partial \phi}{\partial a_{k,\ell}} &= \sum_{\sigma} sgn(\sigma) \cdot \Big[\sum_{j=1}^{n} a_{1,\sigma(1)} \cdot \ldots \cdot \frac{\partial a_{j,\sigma(j)}}{\partial a_{k,\ell}} \cdot \ldots \cdot a_{n,\sigma(n)} \Big] \\ &= \sum_{\sigma : \sigma(k) = \ell} sgn(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot \widehat{a_{k,\ell}} \cdot \ldots \cdot a_{n,\sigma(n)} \end{aligned}$$

(Here we use the standard math notation $b_1 \cdots \widehat{b_i} \cdots b_n = \prod_{j \neq i} b_j$.) Indeed, $\partial a_{j,\sigma(j)} / \partial a_{k,\ell} = 0$ unless $(j, \sigma(j)) = (k, \ell)$ which happens if and only if $j = k$ and $\sigma(k) = \ell$, in which case it is equal to 1. Therefore we have

$$a_{k\ell} \frac{\partial \phi}{\partial a_{k,\ell}} = \sum_{\sigma : \sigma(k) = \ell} sgn(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a_{k,\ell} \cdot \ldots \cdot a_{n,\sigma(n)}$$

But the $1 \times n^2$ matrix $\big[\partial \phi / \partial a_{k\ell} \big]$ has rank $= 1$ (maximal rank) unless all entries are zero, in which case we get

$$\begin{aligned} \det(A) &= \sum_{\sigma(1)=1} (\cdots) + \sum_{\sigma(1)=2} (\cdots) + \ldots + \sum_{\sigma(1)=n} (\cdots) \\ &= \sum_{\sigma \in S_n} sgn(\sigma) \cdot a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} = 0 \end{aligned}$$

That contradicts the hypothesis $\det(A) = 1$, and cannot occur. Thus $(d\phi)_X$ has maximal rank ($= 1$) at each point in $\mathrm{SL}(n, \mathbb{R})$, and in fact at every point in $\mathrm{GL}(n, \mathbb{R})$. $\square$

2.4. Example. The (real) orthogonal groups $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, for which $\mathbb{F} = \mathbb{R}$, both have dimension $\frac{1}{2}(n^2 - n)$.

Discussion: Both are closed subgroups of $\mathrm{M}(n, \mathbb{R})$. On $\mathrm{O}(n)$, $\det(A)$ can only achieve values ± 1 (and -1 is actually achieved); the value is $+1$ precisely on the subgroup $\mathrm{SO}(n)$. The full orthogonal group $\mathrm{O}(n)$ consists of two disjoint closed cosets, the subgroup $\mathrm{SO}(n)$ and the coset $J \cdot \mathrm{SO}(n)$, where J is any orthogonal matrix with $\det(J) = -1$ such as $J = \mathrm{diag}(1, \cdots, 1, -1)$. The coset $J \cdot \mathrm{SO}(n)$ is the set of orientation-reversing orthogonal maps on $\mathbb{R}^n$, which is not a subgroup in $\mathrm{O}(n)$, while $\mathrm{SO}(n)$ is the group of invertible orientation-preserving maps in $\mathrm{O}(n)$.

To show $\mathrm{O}(n)$ is smooth manifold recall that $A \in \mathrm{O}(n) \Leftrightarrow A^{\mathrm{t}} A = I \Leftrightarrow$ the rows R_i are an *ON* basis in $\mathbb{R}^n$, so $(R_i, R_j) = \delta_{ij}$ (Kronecker delta) for $1 \leq i, j \leq n$. Eliminating redundant identities in this list of n^2 identities by requiring $i \leq j$, we define the map

$$\phi : \mathbb{R}^{n^2} \cong \mathrm{M}(n, \mathbb{R}) \rightarrow \mathbb{R}^{(n^2+n)/2} = \mathbb{R}^d$$

given by

$$\phi(A) = \big((R_1, R_2), \ldots, (R_1, R_n); (R_2, R_2), \ldots, (R_2, R_n); \cdots; (R_n, R_n)\big) .$$

Then $\mathrm{O}(n)$ is the level set $L(\phi = \mathbf{x})$ where

$$\mathbf{x} = (1, 0, \cdots, 0; \cdots; 1, 0; 1)$$

We must show $\mathrm{rk}(d\phi)_A = d$ for all $A \in \mathrm{M}(n, \mathbb{R})$ near an arbitrary point $p \in \mathrm{O}(n)$. The idea is clearly revealed by the case $n = 3$, where

$$\phi(A) = \big(\sum_{i=1}^{3} a_{1i}^2 \,,\, \sum_i a_{1i}a_{2i} \,,\, \sum_i a_{1i}a_{3i} \,;\, \sum_i a_{2i}^2 \,,\, \sum_i a_{2i}a_{3i} \,;\, \sum_i a_{3i}^2\big)$$

Writing $\mathbf{x} = \big(a_{11}, a_{12}, a_{13}; a_{22}, a_{23}; a_{33}\big)$ we have

$$\Big[\frac{\partial \phi_i}{\partial a_{ij}}\Big] = \left(\begin{array}{ccc|ccc|ccc} 2a_{11} & 2a_{12} & 2a_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{11} & a_{12} & a_{13} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 & a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 & 2a_{21} & 2a_{22} & 2a_{23} & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{31} & a_{32} & a_{33} & a_{21} & a_{22} & a_{23} \\ 0 & 0 & 0 & 0 & 0 & 0 & 2a_{31} & 2a_{32} & 2a_{333} \end{array}\right)_{6\times 9}$$

Recall that row rank and column rank of any matrix are equal. Symbolic row/column operations show that the row rank $\mathrm{rk}(d\phi)_A$ of this matrix is $6 = \frac{1}{2}(n^2+n)$, so $\dim_{\mathbb{R}} \mathrm{O}(3) = \dim_{\mathbb{R}} \mathrm{SO}(3) = 9 - 6 = 3$ (see next exercise for hints on this calculation). $\square$

2.5. Exercise. Verify that the row rank of the matrix above is 6, at *every* $A \in \mathrm{O}(3)$.
Note: The row rank is equal to the number of linearly independent rows. But $\mathrm{rk}(d\phi)_A$ is always less than or equal to 6, so the rank must in fact be $\equiv 6$ (constant) in a neighborhood of every $A \in \mathrm{O}(3)$.
Hint: The original rows and columns of A are independent so we can transform $A \rightarrow I_{3\times 3}$ by column operations. If we scale the first row of $(d\phi)_A$ by $\frac{1}{2}$ the matrix A appears as a 3×3 submatrix in the upper left corner, with zeros below it. What happens if you apply the same column operations that worked for A to the full Jacobian matrix? Etc. (In later steps you may have to use both row and column operations.) $\square$

All the classical groups mentioned in the previous section can be shown to be smooth closed manifolds over $\mathbb{F} = \mathbb{R}$; we will not do that here. However, if $G \subseteq \mathrm{M}(n, \mathbb{C})$ we should identify $\mathbb{C} \cong \mathbb{R}^2$ and $\mathrm{M}(n, \mathbb{C}) \cong \mathbb{R}^{2n^2}$ in discussing G as a real submanifold. For example $\mathrm{SU}(2)$ is $\subseteq \mathrm{M}(2, \mathbb{C})$ but $\dim_{\mathbb{R}} \mathrm{SU}(2) = 3$, so there is no way complex coordinate charts $x_\alpha : U_\alpha \rightarrow \mathbb{C}^n$ can be introduced into $\mathrm{SU}(2)$ (or $\mathrm{SU}(n)$ for that matter). The classical groups are generally viewed as smooth hypersurfaces in $\mathbb{R}^{n^2}$ or $\mathbb{R}^{2n^2}$, even if they are defined as subsets of $\mathrm{M}(n, \mathbb{C})$. However, a few actually are complex manifolds. For instance, $\dim_{\mathbb{C}} \mathrm{GL}(n, \mathbb{C}) = n^2$ and $\dim_{\mathbb{C}} \mathrm{SL}(n, \mathbb{C}) = n^2 - 1$, although these groups can also be regarded as real manifolds with $\dim_{\mathbb{R}} = 2n^2$ or $2n^2 - 2$. Other examples of intrinsically complex matrix Lie groups are $\mathrm{O}(n, \mathbb{C})$ and $\mathrm{SO}(n, \mathbb{C})$ which have $\dim_{\mathbb{C}} = \frac{1}{2}(n^2 - n)$.

2.6. Exercise. Prove that the matrix group

$$G = \mathrm{U}(2) = \{A \in \mathrm{M}(2, \mathbb{C}) : A^* A = I\}$$

is a smooth hypersurface of dimension $\dim_{\mathbb{R}}(G) = 4$ when we identify $\mathrm{M}(2, \mathbb{C}) \cong \mathbb{C}^4 \cong \mathbb{R}^8$ via the correspondence that sends $A \in \mathrm{U}(2)$,

$$A = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \qquad \text{with } z_{ij} = x_{ij} + \sqrt{-1}\, y_{ij}, \quad (x_{ij}, y_{ij} \in \mathbb{R}) \ ,$$

to the real 8-tuple $\mathbf{x} = \big(x_{11}, \ldots, x_{44}; y_{11}, \ldots, y_{44}\big)$ in $\mathbb{R}^8$ (see next exercise for further comments).
Hint: Write the identities implied by $A^*A = I_{2\times 2}$ as a system of several scalar equations in the variables $\mathbf{x} \in \mathbb{R}^8$, deleting any that are redundant. Define a $\mathcal{C}^\infty$ map $\phi : \mathbb{R}^8 \to \mathbb{R}^k$ has constant rank on and near U(2). (What is the appropriate value of k in this situation?)
Note: SU(2) cannot be a smooth complex manifold even though it is carved out of complex matrix space $\mathrm{M}(2,\mathbb{C})$, because $\dim_{\mathbb{R}} \mathrm{SU}(2) = 4 - 1 = 3$. $\square$

Observe that $\dim_{\mathbb{R}} \mathrm{SO}(3) = 3$ is also the real dimension of SU(2). This coincidence is no accident because there is a two-to-one homomorphism mapping SU(2) onto the Euclidean rotation group SO(3). This "double covering" of SO(3) has profound implications in quantum mechanics.

In the previous exercise you were asked to show that the polynomial map $\phi : \mathbb{R}^8 \to \mathbb{R}^k$ that identifies $M = \mathrm{U}(2)$ as a (real) hypersurface of dimension $d = 8 - r$ in $\mathbb{R}^8 \cong \mathrm{M}(2,\mathbb{C})$ if $\mathrm{rk}(d\phi)_A \equiv r$ (constant) at all points of M. In the next exercise you are asked to find valid choices of d coordinates $x_{i_1}, \ldots, x_{i_d}$ from $x_1, \ldots, x_8$ that parametrize M near p. As in the IFT, given a base point $g \in M$ you must find a partition of coordinate indices $[1,8] = J \cup J'$ such that the restriction $\big(\pi_{J'}|_M\big) : M \to \mathbb{R}^{J'}$ is an admissible coordinate chart near g.

2.7. Exercise. Consider the identity element $g = e = I_{2\times 2}$ in $G = \mathrm{U}(2)$.

1. Produce a partition of $[1,8] = J' \cup J$ such that the projection $\pi_{J'} : \mathbb{R}^8 \to \mathbb{R}^{J'}$ restricts to SU(2) to give a standard chart defined near the identity $I_{2\times 2}$.

This will show that $\dim_{\mathbb{R}}(M) = |J'| = 3$.

2. Determine *all* partitions $[1,8] = J' \cup J$ that produce admissible chart maps near the identity in SU(2).
3. Produce an explicit partition $[1,8] = J' \cup J$ with $|J'| = 3$ that *does not* yield valid coordinates describing SU(2) near the identity element I. $\square$

2.8. Exercise. Perform the calculations outlined in the previous exercise for the group $G = \mathrm{SO}(3)$, identifying the matrix $A = [a_{ij}]$ with the vector

$$\mathbf{x} = (x_1, \cdots, x_9) = \big(a_{11}, \cdots, a_{13}; \cdots; a_{31}, \cdots, a_{33}\big) \in \mathbb{R}^9 \ . \ \square$$

2.9. Exercise. We have already shown that $G = \mathrm{SL}(n,\mathbb{R})$ is a hypersurface in $\mathrm{M}(n,\mathbb{R})$ with $\dim_{\mathbb{R}}(G) = n^2 - 1$. Taking $n = 3$ and identifying matrices $A \in \mathrm{M}(3,\mathbb{R})$ with vectors $\mathbf{x} = (a_{11}, a_{12}, a_{13}; \cdots; a_{31}, a_{32}, a_{33}) \in \mathbb{R}^9$,

1. Exhibit a partition $[1,9] = J' \cup J$ such that the projection $\big(\pi_{J'}|_G\big) : G \to \mathbb{R}^{J'} \cong \mathbb{R}^8$ yields a standard chart on $\mathrm{SL}(3,\mathbb{R})$ near the identity element I.
2. Does *every* choice $J' = \{1 \le i_1 < \cdots < i_8 \le 9\}$ of eight matrix entries (out of 9) yield a standard coordinate chart about the identity element $I_{3\times 3}$ in G? If not, exhibit a choice for which this fails.
3. Repeat (1.) taking $A = \begin{pmatrix} 2 & 3 \\ 0 & \frac{1}{2} \end{pmatrix}$ as the base point in $\mathrm{SL}(3,\mathbb{R})$. $\square$

The need for detailed matrix calculations of this sort can sometimes be avoided by appeal to a theorem of Elie Cartan.

2.10. Theorem (E. Cartan). *Every closed subgroup $G \subseteq \mathrm{M}(n,\mathbb{R})$ can be covered with*

$\mathcal{C}^\infty$-related charts (x_α, U_α) with values in some $\mathbb{R}^d$ that make it into a real Lie group.

No hypothesis is made here that G is a level set for some $\mathcal{C}^\infty$ map $\phi : \mathbb{R}^m \to \mathbb{R}^n$; and even if G happens to be a level set of this kind, there is no need to investigate its rank on G. The subgroup G could even be a closed *discrete* subgroup (a zero-dimensional Lie group consisting of isolated points in matrix space) such as $G = \mathbb{Z} \subseteq \mathbb{R}$, or

$$\mathrm{SL}(n,\mathbb{Z}) = \{A \in \mathrm{M}(n,\mathbb{R}) : \det(A) = +1 \text{ and all entries } a_{ij} \text{ are integers}\}\ ,$$

for which differentiability considerations are irrelevant.

Translations and Automorphisms on Matrix Lie Groups.

There are some obvious remarks to be made about any matrix Lie group G. If $g \in G$ we can define left and right **translation operators** $\lambda_x, \rho_x : G \to G$, letting

$$\lambda_x(g) = x \cdot g \quad \text{and} \quad \rho_x(g) = g \cdot x\ .$$

These are continuous and invertible maps with the properties

$$\lambda_e = \mathrm{id}_G \qquad \lambda_{x\cdot y} = \lambda_x \circ \lambda_y \qquad \lambda_{x^{-1}} = \big(\lambda_x\big)^{-1}\ ,$$

and likewise for right translations, except that

$$\rho_{x\cdot y} = \rho_y \circ \rho_x$$

They are invertible bicontinuous maps when G is given the obvious metric it inherits from the surrounding space $\mathrm{M}(n,\mathbb{F})$. Furthermore, all these maps are **diffeomorphisms**: when λ_x or ρ_x are described in terms of chart coordinates they (and their inverses) become $\mathcal{C}^\infty$ maps of $\mathbb{R}^d \to \mathbb{R}^d$, where $d = \dim_{\mathbb{R}}(G)$.

The action of an element $g \in G$ by conjugation yields an **inner automorphism** of the group,

$$i_g(x) = gxg^{-1} \quad \text{for all } x \in G\ .$$

This is a diffeomorphisms of $G \to G$ because $i_g = \lambda_g \circ \rho_g^{-1}$ is a composite of $\mathcal{C}^\infty$ maps. (Note that λ_x, ρ_y commute for all $x, y \in G$.) Its inverse is the inner automorphism $\big(i_g\big)^{-1} = i_{g^{-1}}$, and the correspondence $g \mapsto i_g$ is a homomorphism from G into the full group $\mathrm{Aut}(G)$ of automorphisms of G, because

$$i_e = \mathrm{id}_G \qquad \text{and} \qquad i_{g_1 g_2} = i_{g_1} \circ i_{g_2}$$

If $G \neq \mathrm{GL}(n,\mathbb{R})$ is a matrix Lie group its C^∞ structure is obtained by identifying $G \subseteq \mathrm{M}(n,\mathbb{R}) \cong \mathbb{R}^{n^2}$ and then covering G with standard charts (x_α, U_α) constructed via the IFT, where U_α is an open subset in G and x_α is a bicontinuous map from U_α to the open set $V_\alpha = x_\alpha(U_\alpha)$ in coordinate space $\mathbb{R}^d$. As in the IFT, the chart maps x_α are just restrictions to G of various projection maps $\pi_{_{J'}} : \mathbb{R}^{n^2} \to \mathbb{R}^{J'}$ corresponding to some splitting $\mathbb{R}^{n^2} = \mathbb{R}^J \oplus \mathbb{R}^{J'}$ with $|J'| = d$. Because of the way standard charts are defined, it is not hard to verify that a $\mathcal{C}^\infty$ map $\psi : \mathbb{R}^k \to \mathrm{M}(n,\mathbb{R})$ whose range happens to lie within the set $G \subseteq \mathrm{M}(n,\mathbb{R})$, is automatically a $\mathcal{C}^\infty$ map $\psi : \mathbb{R}^k \to G$ when G is given its standard structure as a real $\mathcal{C}^\infty$ manifold.

2.11. Exercise. Left translations λ_x are obviously bijective maps on G. Prove that they are $\mathcal{C}^\infty$ maps for the standard $\mathcal{C}^\infty$ structure on G. Thus if if (x_α, U_α), (y_β, U_β) are standard charts about a point $p \in G$ and its image $\lambda_x(g) = x \cdot p$, explain why λ_x becomes a $\mathcal{C}^\infty$ map $\mathbb{F}^d \to \mathbb{F}^d$ when described in local chart coordinates – i.e. why is the coordinate map

$$\mathbf{y} = \Phi(\mathbf{x}) = y_\beta \circ \lambda_x \circ x_\alpha^{-1}(\mathbf{x})$$

a $\mathcal{C}^\infty$ map $\mathbb{F}^d \to \mathbb{F}^d$ wherever it is well-defined?
Note: Since $(\lambda_x)^{-1} = \lambda_{x^{-1}}$ is also a translation, λ_x is a $\mathcal{C}^\infty$ diffeomorphism of G. Likewise for right translations ρ_x. $\square$

2.12. Lemma. *Let G be a matrix Lie group in $\mathrm{M}(n,\mathbb{R})$ and $f : \mathbb{R}^k \to \mathrm{M}(n,\mathbb{R}) \cong \mathbb{R}^{n^2}$ a $\mathcal{C}^\infty$ map. If G is given the standard $\mathcal{C}^\infty$ structure as a smooth real hypersurface and if $\mathrm{range}(\psi) \subseteq G$, then $f : \mathbb{R}^k \to G$ is automatically a $\mathcal{C}^\infty$ map – i.e. if $x_0 \in \mathbb{R}^k$, $p_0 = \psi(x_0)$ and if (x_α, U_α) is a standard chart on G about p_0, then*

$$x_\alpha \circ \psi : \mathbb{R}^k \to \mathbb{R}^d$$

is a C^∞ map between Euclidean spaces.

X.3. The Exponential Map $\mathrm{Exp} : \mathrm{M}(n,\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$.

The matrix groups above were all described as smooth d-dimensional hypersurfaces M embedded in some Euclidean space $\mathbb{R}^m$ with $m \geq d$. At each base point $p \in M$ there is a well-defined d-dimensional **tangent space** TM_p, a translate $\mathrm{TM}_p = p + E_p$ of some d-dimensional vector subspace $E_p \subseteq \mathbb{R}^m$ (see Figure 10.7). It is an almost universal convention to indicate tangent vectors in TM_p by capital roman letters $X, Y, ...$, or by $X_p, Y_p, ...$ when it is necessary to indicate the base point.

The tangent hyperplane TM_p itself is not itself a vector subspace in matrix space (for one thing, it does not contain the zero element), but we may nevertheless define vector space operations on tangent vectors *attached to the same base point p*, and think of $\mathrm{TM}_p = p + E_p$ as a vector space, by referring everything back to E_p by translation.

Vector Operations in $\mathbf{TM}_p$. *Given tangent vectors X_1, X_2 in TM_p form their sum $X_1 + X_2$ in TM_p by*

(50)

(i) Subtracting p from each to get actual vectors $v_k = X_k - p$ in E_p, then add these to get $v_1 + v_2$ in E_p where $(+)$ makes sense.

(ii) Then move the result back into the tangent space to get $X_1 + X_2 = p + (v_1 + v_2)$ in TM_p.

Scalar multiples $\lambda{\cdot}X = p + (\lambda \cdot v)$ in TM_p are defined similarly. Note carefully that there is no natural way to define the sum of two tangent vectors attached to different base points $p \neq q$ in M.

Equipped with these operations the tangent hyperplane TM_p becomes a vector space isomorphic to E_p under the bijection $v \mapsto v + p$, $v \in E_p$.

To carry out calculations with tangent vectors we shall exploit the close connection between derivatives of smooth curves in $\mathbb{R}^m$ passing through p and tangent vectors in TM_p. A *parametric curve* in $E = \mathbb{R}^m$ is a continuous map $\gamma : [a,b] \to \mathbb{R}^m$, and if $\{\mathbf{e}_j\}$ is the standard basis in $\mathbb{R}^m$ we may describe γ by writing

$$\gamma(t) = \sum_{j=1}^{m} x_j(t)\,\mathbf{e}_j \ .$$

This is a $\mathcal{C}^\infty$ curve if the scalar components $x_j(t)$ are smooth functions.
Note: When $E = \mathbb{C}^m$ a parametric curve takes the form

$$\gamma(t) = \sum_{j=1}^{m} z_j(t)\,\mathbf{e}_j \quad \text{with complex coefficients } z_j(t) = x_j(t) + \sqrt{-1}\,y_j(t) \ ,$$

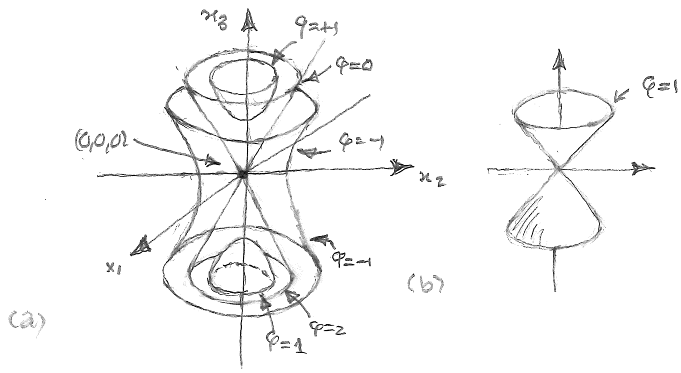

Figure 10.7. The tangent hyperplane TM_p to a smooth d-dimensional hypersurface in Euclidean space $\mathbb{R}^m$ is made up of vectors $p + \mathbf{v}$ attached to the base point p, where $\mathbf{v}$ lies in a certain vector subspace E_p passing through the origin in $\mathbb{R}^m$. Vector space operations are introduced into the tangent space by parallel transport between vectors $X = p + \mathbf{v} \in \mathrm{TM}_p = p + E_p$ and $\mathbf{v} \in E_p$, which actually is a vector subspace of $\mathbb{R}^m$.

and $\gamma(t)$ is a $\mathcal{C}^\infty$ curve if the real and imaginary parts $x_j, y_j : [a, b] \rightarrow \mathbb{R}$ are real-valued $\mathcal{C}^\infty$ functions. In this situation the vector derivative takes the form

$$\gamma'(t) = \sum_{j=1}^{m} \frac{dz_j}{dt}\,\mathbf{e}_j = \sum_{j=1}^{m} \Big(\frac{dx_j}{dt} + \sqrt{-1}\frac{dy_j}{dt}\Big)\,\mathbf{e}_j \qquad \square$$

3.1. Definition. *Let M be any smooth d-dimensional manifold embedded in Euclidean space $\mathbb{R}^m$ and consider any $\mathcal{C}^\infty$ curve $\gamma(t) \sum_{j=1}^{m} x_j(t)\,\mathbf{e}_j$ in $\mathbb{R}^m$ that remains in M for all t. By Lemma 2.12, γ is a $\mathcal{C}^\infty$ map from $\mathbb{R}$ into M equipped with its standard $\mathcal{C}^\infty$ structure, so if (x_α, U_α) is a standard chart on M the map $x_\alpha \circ \gamma(t)$ is differentiable from $\mathbb{R} \rightarrow \mathbb{R}^d$.*

Any $\mathcal{C}^\infty$ curve $\gamma(t)$ with values in $\mathbb{R}^m$ is differentiable, with vector derivative

$$\gamma'(t_0) = \frac{d\gamma}{dt}(t_0) = \lim_{\Delta t \to 0} \frac{\gamma(t_0 + \Delta t) - \gamma(t_0)}{\Delta t} = \frac{dx_1}{dt}\mathbf{e}_1 + \ldots + \frac{dx_m}{dt}\mathbf{e}_m$$

(limit computed in $\mathbb{R}^m$) for all $a < t_0 < b$. Moreover, if $\gamma(t)$ remains in M for all t and passes through p when $t = t_0$, then the vector derivative $\gamma'(t_0)$ must lie in the hyperplane tangent to M at p. Therefore $\gamma'(t_0)$ is an element of the tangent space TM_p, and in fact every tangent vector at p arises this way, as the vector derivative of a smooth curve in M passing through p.

In discussing tangent vectors to smooth curves we may as well assume $t_0 = 0$ because we can always reparametrize γ via $\gamma(t) \rightarrow \tilde{\gamma}(t) = \gamma(t - t_0)$ to make $\gamma(0) = p$ without changing the tangent vector at p.

Lie Algebra of a Matrix Lie Group. If M is a smooth hypersurface in Euclidean space $\mathbb{R}^m$ the tangent space TM_p at $p \in M$ acquires a vector space structure as explained in (50). But if the hypersurface is a matrix Lie group G, the tangent space TG_e at the identity element acquires additional algebraic structure, induced by the algebraic operations in G itself, and becomes a *Lie algebra*.

3.2. Definition. *An abstract* **Lie algebra** *over $\mathbb{F}$ is a vector space V over $\mathbb{F}$ equipped with a bracket operation $B(X, Y) = [X, Y]$ that is:*

1. Bilinear. $B(X,Y) = [X,Y]$ *is bilinear map from* $V \times V \to V$ (*linear in each entry when the other is held fixed*).

2. Antisymmetric. $[X,Y] = -[Y,X]$ *for all* $X, Y \in V$, *so* $[X,X] = 0$ *for all* X.

3. Jacobi Identity. $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$ *for all* $X,Y,Z \in V$.

A concrete example is provided by the imposing the "commutator operation" $[X,Y] = XY - YX$ on the algebra of $n \times n$ matrices over any field $\mathbb{F}$; it measures the degree to which X and Y fail to commute.

3.3. Exercise. Verify that the Jacobi identity in Definition 3.2 holds if we take $[X,Y] = XY - YX$ for X, Y in the matrix algebra $\mathrm{M}(n,\mathbb{F})$ $\square$

Similarly, the space $\mathrm{Hom}_{\mathbb{F}}(V,V)$ of $\mathbb{F}$-linear operators on a vector space is an associative algebra over $\mathbb{F}$ if we take composition of operators $S \circ T$ as its multiplication operation. It becomes a Lie algebra when equipped with the commutator operation $[S,T] = S \circ T - T \circ S$. (No surprise here: If $\dim_{\mathbb{F}}(V) = n$ and $\mathfrak{X}$ is a basis in V, the correspondence $T \mapsto [T]_{\mathfrak{X}}$ is a natural isomorphism $\mathrm{Hom}_{\mathbb{F}}(V,V) \cong \mathrm{M}(n,\mathbb{F})$ of associative algebras, which sends commutators to commutators.)

If G is a matrix Lie group there is a natural bracket operation $[X,Y]$ that makes the tangent space at the identity TG_e a Lie algebra – the **Lie algebra of** G – which will hereafter be denoted by the German "fraktur" letter $\mathfrak{g}$ instead of TG_e. But it is not immediately apparent how the multiplication operation $(\cdot)$ in G induces such a bracket operation in $\mathfrak{g}$. The key is to get a handle on tangent vectors $X \in \mathfrak{g}$ by realizing them as vector derivatives $\gamma'(0)$ of various $\mathcal{C}^\infty$ curves $\gamma(t)$ that pass through the group identity $p = e$ when $t = 0$,

$$X_p = \gamma'(0) = \frac{d}{dt}\{\gamma(t)\}\Big|_{t=0}$$

To pursue this we need some basic facts about derivatives of algebraic combinations of smooth curves in matrix space.

3.4. Lemma. *If* $\gamma, \eta : \mathbb{R} \to \mathrm{M}(n,\mathbb{F})$ *are* $\mathcal{C}^\infty$ *curves defined for* $a < t < b$, *then*

(*i*) Sum Rule: $\dfrac{d}{dt}\{\gamma(t) + \eta(t)\} = \dfrac{d}{dt}\{\gamma(t)\} + \dfrac{d}{dt}\{\eta(t)\}$

(*ii*) Product Rule: *The matrix product* $\gamma(t) \cdot \eta(t)$ *has derivative*

$$\frac{d}{dt}\{\gamma(t) \cdot \eta(t)\} = \frac{d\gamma}{dt} \cdot \eta(t) + \gamma(t) \cdot \frac{d\eta}{dt}$$

for all t.

Proof: Item (1.) is trivial. For (2.) write the appropriate difference quotients as

$$\begin{aligned} &\frac{\gamma(t+\Delta t)\cdot\eta(t+\Delta t) - \gamma(t)\cdot\eta(t)}{\Delta t} \\ &\quad = \frac{\gamma(t+\Delta t)\cdot\eta(t+\Delta t) - \gamma(t)\cdot\eta(t+\Delta t)}{\Delta t} + \frac{\gamma(t)\cdot\eta(t+\Delta t) - \gamma(t)\cdot\eta(t)}{\Delta t} \\ &\quad \to \gamma'(t)\cdot\eta(t) + \gamma(t)\cdot\eta'(t) \qquad \text{as } t \to 0 \ , \end{aligned}$$

because the product operation is *jointly continuous*, with $\gamma(s) \cdot \eta(t) \to \gamma(0) \cdot \eta(0)$ as $s, t \to 0$ independently. $\square$

This yields an intrinsic description of the vector space operations in the tangent space at the identity element of a matrix Lie group.

3.5. Lemma. *The Lie algebra* $\mathfrak{g} = \mathrm{TG}_e$ *of a matrix Lie group is a vector space over* $\mathbb{R}$.

Proof: If $c \in \mathbb{R}$ and $\gamma(t)$ a smooth curve passing through g when $t = 0$, and if $\gamma'(0) = X \in \mathfrak{g}$, then $\eta(t) = \gamma(ct)$ is also smooth, with

$$\frac{d}{dt}\{\eta(t)\} = c \cdot \frac{d\gamma}{dt}$$

Hence $\eta'(0) = c \cdot X$ is in $\mathfrak{g} = \mathrm{TG}_e$ if $X \in \mathfrak{g}$.

If γ, η have $\gamma'(0) = X$, $\eta'(0) = Y$ in $\mathfrak{g}$ then $\Phi(t) = \gamma(t) \cdot g^{-1} \cdot \eta(t)$ is a $\mathcal{C}^\infty$ map with $\Phi(0) = gg^{-1}g = g$, and by Lemma 3.4 we get

$$\Phi'(0) = \gamma'(0) \cdot g^{-1}\eta(0) + \gamma(0)g^{-1} \cdot \eta'(0) = X + Y \in \mathfrak{g}$$

since $\gamma(0) = \eta(0) = g$. $\square$

3.6. Exercise. If $\gamma : \mathbb{R} \to \mathrm{GL}(n, \mathbb{F})$ is a $\mathcal{C}^\infty$ matrix-valued curve that passes through $I = I_{n\times n}$ when $t = 0$, prove that

(a) For sufficiently small values of t the curve

$$\eta(t) = \big(\gamma(t)\big)^{-1} \quad \big(\text{inverse of } \gamma(t) \text{ in } \mathrm{M}(n, \mathbb{F})\big)$$

is a $\mathcal{C}^\infty$ curve with $\eta(0) = I$ and $\eta'(0) = -\gamma'(0)$.

(b) For fixed $g \in \mathrm{GL}(n, \mathbb{F})$, $\eta(t) = \lambda_g(\gamma(t)) = g\cdot\gamma(t)$ is a $\mathcal{C}^\infty$ curve with $\eta(0) = g$ and vector derivative $\eta'(t) = g\cdot\gamma'(t)$ for all $t \in \mathbb{R}$.

Note: In (a), $\gamma(0) = I \Rightarrow \det(\gamma(0)) = 1$; by continuity, $\det(\gamma(t)) \neq 0$ for all t near zero, so $\gamma(t)^{-1}$ exists. Furthermore, since $\gamma : \mathbb{R} \to \mathrm{M}(n, \mathbb{F})$ is a C^∞ map, as are $t \mapsto \det(\gamma(t))$ and $t \mapsto \det(\eta(t)) = 1/\det(\gamma(t))$ for small values of t. It follows that $\gamma(t)^{-1}$ is a $\mathcal{C}^\infty$ matrix-valued map for small values of t. $\square$

3.7. Lemma. *The similarity transformation* $i_g : X \to gXg^{-1}$ *leaves* $\mathfrak{g} = \mathrm{TG}_e$ *invariant for every* $g \in G$.

Proof: If $X \in \mathfrak{g}$ and γ is a $\mathcal{C}^\infty$ curve such that $\gamma(0) = e$ and $\gamma'(0) = X$, then $\eta(t) = g\gamma(t)g^{-1}$ is a smooth curve with $\eta(0) = e$ because products and inverses in G are $\mathcal{C}^\infty$ operations. By Lemma 3.4 we have $\eta'(t) = g\gamma'(t)g^{-1} = gXg^{-1}$. $\square$

Next, we observe that the tangent space TG_g to a matrix Lie group at a base point $g \in G$ is a left translate (by λ_g) of the tangent space $\mathrm{TG}_e = \mathfrak{g}$.

3.8. Lemma. *If* $g \in G$ *the tangent space to* G *at* g *is*

$$\mathrm{TG}_g = \lambda_g(\mathfrak{g}) = \{g\cdot X : X \in \mathfrak{g}\} \quad (\textit{matrix product in } \mathrm{M}(n, \mathbb{F}))\ .$$

Proof: For $x \in G$ the left/right translation operations $\lambda_g(x) = gx$ and $\rho_g(x) = xg$ are both bijective $\mathcal{C}^\infty$ maps on G with $\mathcal{C}^\infty$ inverses, so they are $\mathcal{C}^\infty$ diffeomorphisms of the Lie group.

If $\gamma(t)$ is a $\mathcal{C}^\infty$ curve in G with $\gamma(0) = e$ and $\gamma'(0) = X \in \mathfrak{g}$, then $\eta(t) = \lambda_g(\gamma(t)) = g \cdot \gamma(t)$ is a $\mathcal{C}^\infty$ curve with values in G that passes through $\eta(0) = g$. Furthermore, $\eta'(0) = g \cdot \gamma'(0) = g \cdot X$, and as noted earlier $X = \gamma'(0)$ is a typical tangent vector to G at e. Conversely any $Y \in \mathrm{TG}_g$ is obtained in this way, for if $Y = \eta'(0)$ for some smooth curve η passing through g when $t = 0$, then $\gamma(t) = \lambda_g^{-1} \cdot (\eta(t))$ passes through $\gamma(0) = e$ and $X = \gamma'(0) = g^{-1}\cdot\eta(0) = g^{-1}\cdot Y$ is in $\mathfrak{g}$ with $\lambda_g(X) = Y$. $\square$

This transfer of tangent vectors between base points in G is illustrated in Figure 10.8.

In view of Lemma 3.8, any tangent vector $X \in \mathrm{TG}_e$ at the identity element extends uniquely by left translations to a "field of tangent vectors" $\tilde{X}$ defined everywhere on G,

$$\tilde{X}_g = \lambda_g(X) = g \cdot X \qquad \text{for } g \in G. \tag{51}$$

We could also shift vectors $X \in \mathfrak{g}$ by *right* translation $X \mapsto X \cdot g = \rho_g(X)$, but unless G is commutative this does not produce the same field of tangent vectors as left translation.
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

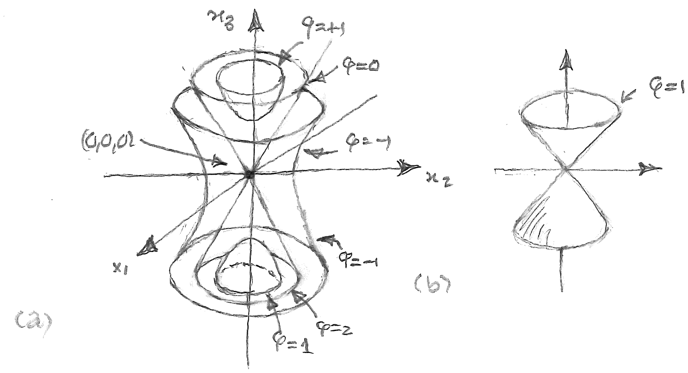

Figure 10.8. INSERT NEW CAPTION.

NOTE: A block of text about extending $X \in \mathfrak{g}$ to all of GL has been stored at EOF under heading "Fields on GL vs fields on G" Use it later on.

xxxxxxx ARCHIVED BLK: Fields on GL vs fields on G -¿ ARCHIVE xxxxxxxxxxxxxx

3.9. Definition (Smooth Vector Fields on G). *A map $\tilde{X} : G \to \mathrm{M}(n, \mathbb{F})$ that assigns a tangent vector $\tilde{X}_g \in \mathrm{TG}_g$ to each base point $g \in G$ is a* **vector field** *on G. It is a* **smooth vector field** *if the matrix elements in $\left[(\tilde{X}_g)_{ij}\right]$ are $\mathcal{C}^\infty$ scalar valued functions when described in local chart coordinates on G – i.e. if $\tilde{X}_{ij} \circ x_\alpha^{-1} : \mathbb{R}^d \to \mathbb{R}$ is a $\mathcal{C}^\infty$ function on coordinate space $\mathbb{R}^d$ ($d = \dim G$), for every standard chart (x_α, U_α) on G.*

A **left-invariant vector field** *on G is one such that*

$$\tilde{X}_{gh} = \lambda_g(\tilde{X}_h) = g \cdot X_h \qquad \textit{for all } g, h \in G\ .$$

There are many smooth vector fields on G, but a left-invariant field is completely determined by its value $\tilde{X}_e = X \in \mathfrak{g}$ at the identity element in G, because $\tilde{X}_g = \tilde{X}_{ge} = \lambda_g(\tilde{X}_e) = \lambda_g(X)$ in TG_g. Thus there is a bijective correspondence between elements $X \in \mathfrak{g}$ and the space of left-invariant vector fields $\tilde{X}_g$ on G.

The fields $\tilde{X}$ defined in (51) are typical left-invariant fields of vactors on G; they are also smooth vector fields as in Definition 3.10.

3.10. Exercise. If G is a matrix Lie group and $X \in \mathfrak{g} = \mathrm{TG}_e$, prove that the vector field $\tilde{X}_g$ defined in (51) is *smooth.* Thus if $\tilde{X}_g = [X_{ij}(g)]$ in $\mathrm{TG}_g \subseteq \mathrm{M}(n, \mathbb{F})$ and if (x_α, U_α) is one of the standard charts covering G, show that the matrix coefficients $X_{ij}(g)$ become $\mathcal{C}^\infty$ scalar-valued functions

$$X_{ij} \circ x_\alpha^{-1} = \big(\lambda_g(X)\big)_{ij} \circ x_\alpha^{-1} = (g \cdot X)_{ij} \circ x_\alpha^{-1}$$

from $\mathbb{R}^d \to \mathbb{R}$ for $1 \leq i, j \leq n$.
Hint: X and g are both matrices in $\mathrm{M}(n, \mathbb{F})$ and so is their product. $\square$

We now show how the tangent space space $\mathfrak{g} = \mathrm{TG}_e$ acquires a Lie algebra structure.

3.11. Proposition (Lie Algebra Property). *If G is a matrix Lie group in $\mathrm{M}(n, \mathbb{F})$ its tangent space at the identity $\mathfrak{g} = \mathrm{TG}_e$ is closed under formation of commutators in $\mathrm{M}(n, \mathbb{F})$,*

$$X, Y \in \mathfrak{g} \Rightarrow [X, Y] = XY - YX \text{ is also in } \mathfrak{g} \;, \tag{52}$$

so $\mathfrak{g}$ is a Lie subalgebra in $\mathrm{M}(n.\mathbb{F})$.

Note: Vector fields on G have values $X_g \in \mathrm{TG}_g \subseteq \mathrm{M}(n, \mathbb{F})$. The products XY and YX of fields appearing in (52) are defined pointwise, as products of their matrix values – i.e. $(XY)_g = X_g \cdot Y_g$ (matrix product), etc. It is worth noting that the individual products $(XY)_g$ and $(YX)_g$ *need not* lie in the tangent space TM_g, but their commutator $[X, Y]_g$ is always an element of TM_g. $\square$

> SOMETHING THAT NEEDS TO BE RESOLVED: Is $[X, Y]_g = \lambda_g([X, Y])$ actually equal to $[X_g, Y_g] = X_g Y_g - Y_g X_g$ (commutator of of matrices in $\mathrm{TG}_g \subseteq \mathrm{M}(n, \mathbb{F})$??

Proof: Let $\gamma(t), \eta(t)$ be $\mathcal{C}^\infty$ curves in G that pass through e when $t = 0$ and have vector derivatives $\gamma'(0) = X, \eta'(0) = Y$ in $\mathfrak{g}$. For $s, t \in \mathbb{R}$ let $f(t) = \gamma(s)\eta(t)\gamma(s)^{-1}$ (product of elements in G). Then f is $\mathcal{C}^\infty$, $f(0) = e$, and $f(t)$ remains within G for all t (and s) because G is a group, so $f'(0)$ is in the Lie algebra $\mathfrak{g}$. But we also have

$$\begin{aligned} f'(0) &= \frac{d}{dt}\{\gamma(t)\}\Big|_{t=0} = \frac{\partial}{\partial t}\big\{\gamma(s)\eta(t)\gamma(s)^{-1}\big\}\Big|_{t=0} \\ &= i\gamma(s) Y \gamma(s)^{-1} \end{aligned}$$

for all s near 0. Hence $\gamma(s) \cdot \mathrm{TG}_e \cdot \gamma(s)^{-1} \subseteq \mathrm{TG}_e$ for all s, and then

$$\begin{aligned} \frac{\partial}{\partial s}\big\{\gamma(s) Y \gamma(s)^{-1}\big\}\Big|_{s=0} &= \gamma'(0) Y \gamma(0)^{-1} + \gamma(0) Y \frac{d}{ds}\{\gamma(s)^{-1}\}\Big|_{s=0} \\ &= \gamma'(0) Y \gamma(0)^{-1} + \gamma(0) Y \cdot (-\gamma'(0)) \\ &= XY - YX = [X, Y] \quad \big(\text{commutator in } \mathrm{M}(n, \mathbb{R})\big) \end{aligned}$$

as required to show $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g}$. $\square$

Note: Another way to say this is:

$$[X, Y] = \frac{\partial^2}{\partial s\, \partial t}\big\{\gamma(s)\eta(t)\gamma(s)^{-1}\big\}\Big|_{s=t=0} \,. \tag{53}$$

Lie Algebras of Classical Groups.

We now compute some examples by determining which matrices in $\mathrm{M}(n, \mathbb{R})$ appear in $\mathfrak{g}$. When $G = \mathrm{GL}(n, \mathbb{F})$ there is nothing to do: because $\mathrm{GL}(n, \mathbb{F})$ is an open set in matrix space, $\mathfrak{gl}(n, \mathbb{F})$ is all of $\mathrm{M}(n, \mathbb{F})$ for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$.

3.12. Example. If $G = \mathrm{SL}(n, \mathbb{F})$ then $\mathfrak{sl}(n, \mathbb{F}) = \{X \in \mathrm{M}(n, \mathbb{F}) : \mathrm{Tr}(X) = 0\}$, which has $\dim_{\mathbb{F}}(\mathfrak{sl}) = n^2 - 1$.

Discussion: Let $\phi(t)$ be a $\mathcal{C}^\infty$ curve in G with $\phi(0) = I_{n \times n}$ and $\phi'(0) = X$ in the Lie algebra $\mathfrak{g}$. Then by definition of SL

$$1 \equiv \det\big(\phi(t)\big) = \sum_{\pi \in S_n} sgn(\pi) \cdot \prod_{j=1}^{n} \phi(t)_{j, \pi(j)} \quad \text{for all } t,$$

which by the Product Rule implies

$$0 = \frac{d}{dt}\{\phi(t)\} = \sum_{\pi\in S_n} sgn(\pi)\cdot\Big(\sum_{k=1}^{n}\phi(t)_{1,\pi(1)}\cdot\ldots\cdot\phi'(t)_{k,\pi(k)}\cdot\ldots\cdot\phi_{n,\pi(n)}(t)\Big)\ .$$

Setting $t = 0$ we have $\phi(0) = I_{n\times n}$. Therefore, for $j \neq k$ the j^{th} term in $\big(\sum_{k=1}^{n}\cdots\big)$ is zero unless unless $\pi(j) = j$; but then when $j = k$ we must have $\pi(k) = k$ too, so $\pi = \text{id}$ in S_n. We conclude that the entire sum over k is zero, except for the term corresponding to $\pi = \text{id}$ in the outer sum over S_n. Then, since $\phi'(0) = X$ in $\mathrm{M}(n,\mathbb{F})$ we get

$$0 = \sum_{k=1}^{n} 1\cdot\ldots\cdot 1\cdot x_{kk}\cdot 1\cdot\ldots\cdot 1 \;=\; \sum_{k=1}^{n} x_{kk} \;=\; \mathrm{Tr}(X)$$

where $x_{kk} = [\phi'(0)]_{kk}$

Conversely, if $X \in \mathrm{M}(n,\mathbb{F})$ and $\mathrm{Tr}(X) = 0$ the curve $\phi(t) = e^{tX} = \sum_{n=0}^{\infty}(t^n/n!)X^n$ is $\mathcal{C}^\infty$ and passes through I when $t = 0$. Furthermore the Exponent Law for matrix exponentials $e^{(s+t)X} = e^{sX}\cdot e^{tX}$ implies that

$$\frac{d}{dt}\{e^{tX}\} = X\cdot e^{tX} \qquad \text{for all } X \in \mathrm{M}(n,\mathbb{F}), t\in\mathbb{R} \tag{54}$$

In fact,

$$\begin{aligned}
\frac{d}{dt}\{e^{tX}\} &= \lim_{\Delta t\to 0}\frac{e^{(t+\Delta t)X} - e^{tX}}{\Delta t} \;=\; \lim_{\Delta t\to 0} e^{tX}\cdot\Big(\frac{e^{\Delta t\cdot X} - 1}{\Delta t}\Big)\\
&= \lim_{\Delta t\to 0} e^{tX}\cdot\Big(X + \frac{\Delta t\cdot X^2}{2!} + \frac{(\Delta t)^2 X^3}{3!} + \cdots\Big)\\
&= \lim_{\Delta t\to 0} e^{tX}\big(X + o(\Delta t)\big) \;=\; X\cdot e^{tX}
\end{aligned}$$

Thus $\phi'(0) = X$ and $\phi(0) = I$.

Furthermore, $X \in \mathfrak{g} \Rightarrow \phi(t)$ remains within G for all t because

$$\det(\phi(t)) = \det\big(e^{tX}\big) = e^{t\cdot\mathrm{Tr}(X)} = 1 \quad \text{ for all } t\ .$$

Thus $\mathfrak{sl}(n,\mathbb{F}) = \{X \in \mathrm{M}(n,\mathbb{F}) : \mathrm{Tr}(X) = 0\}$. When $\mathbb{F} = \mathbb{C}$, $\mathrm{SL}(n,\mathbb{C})$ is a complex group and $\mathfrak{sl}(n,\mathbb{C})$ a complex Lie algebra. In either case $\dim_{\mathbb{F}}(G) = \dim_{\mathbb{F}}(\mathfrak{sl}(n,\mathbb{F}))$, and our identification of the Lie algebra immediately shows $\dim_{\mathbb{F}}(\mathrm{SL}(n,\mathbb{F}) = n^2 - 1$. □

3.13. Exercise. *When $G = \mathrm{U}(n)$ or $\mathrm{SU}(n)$ the Lie algebras are*

$$\begin{aligned}
\mathfrak{u}(n) &= \{X \in \mathrm{M}(n,\mathbb{C}) : X^* = -X\}\\
\mathfrak{su}(n) &= \{X \in \mathrm{M}(n,\mathbb{C}) : X^* = -X \text{ and } \mathrm{Tr}(X) = 0\}
\end{aligned}$$

Proof ($\subseteq$)**:** Let $\phi : \mathbb{R} \to G$ be a $\mathcal{C}^\infty$ curve in G with $\phi(0) = I$, $\phi'(0) = X$ in $\mathfrak{g}$. Since

$$\frac{d}{dt}\{\phi^*(t)\} = \left(\frac{d\phi}{dt}\right)^*$$

and

$$0 = \frac{d}{dt}\{\phi^*(t)\cdot\phi(t)\} = \frac{d}{dt}\{\phi^*\}\cdot\phi(t) + \phi(t)^*\frac{d}{dt}\{\phi\}$$

we get

$$0 = \phi'(t)^*\phi(t) + \phi(t)^*\phi'(t) \qquad \text{for all small } t$$

Setting $t = 0$, we get $X^* = -X$, which shows that $\mathfrak{u}(n) \subseteq \{X : X^* = -X\}$ and similarly for $\mathfrak{su}(n)$.

Proof ($\supseteq$)**:** Conversely, if $X^* = -X$ and $\phi(t) = e^{tX}$ we have

$$\phi(t)^* = e^{t(X^*)} = e^{-tX} = \phi(t)^{-1}$$

and hence $\phi(\mathbb{R}) \subseteq G = \mathrm{U}(n)$. Since ϕ is $\mathcal{C}^\infty$, $\phi(0) = I$, and $\phi'(0) = X$ we see that X is a tangent vector in $\mathfrak{u}(n)$, proving the inclusion ($\supseteq$). The same argument works for $\mathfrak{su}(n)$.

These groups are real, not complex, even though they are carved out of $\mathrm{M}(n, \mathbb{C})$, because $X^* = -X$ forces diagonal elements to be in $i\mathbb{R}$ (imaginary). $\square$

3.14. Exercise. Show that $\dim_{\mathbb{R}}(\mathfrak{u}(n)) = n^2$ and $\dim_{\mathbb{R}}(\mathfrak{su}(n)) = n^2 - 1$. $\square$

3.15. Example. *If* $G = \mathrm{O}(n, \mathbb{C})$ *then* $\mathfrak{o}(n, \mathbb{C}) = \{X \in \mathrm{M}(n, \mathbb{C}) : X^{\mathrm{t}} = -X\}$.

Discussion. This is also $\mathfrak{so}(n, \mathbb{C})$ even though $\mathrm{O}(n, \mathbb{C}) \supsetneqq \mathrm{SO}(n, \mathbb{C})$. [Recall that $\mathrm{O}(n, \mathbb{C})$ consists of two connected closed cosets (though we have not yet shown $\mathrm{SO}(n, \mathbb{C})$ is an arcwise connected group). These are intrinsically complex Lie groups, and it is easily seen that $\dim_{\mathbb{C}} \mathfrak{o}(n, \mathbb{C}) = \dim_{\mathbb{C}} \mathfrak{so}(n, \mathbb{C})$ for the tangent spaces at the identity, which have the same dimensions as the groups.] Here $\phi(s)^{\mathrm{t}}\phi(s) = I$ so $\phi'(s)^{\mathrm{t}}\phi(s) + \phi(s)^{\mathrm{t}}\phi'(s) = 0$ for all s, which implies that $X^{\mathrm{t}} = X$, Clearly $\phi(s)$ is a $\mathcal{C}^\infty$ curve in matrix space such that $\phi(0) = I$ and $\phi'(0) = X$, by (3.4(ii)). Furthermore $\phi(s) = e^{sX}$ remains confined within $\mathrm{SO}(n, \mathbb{C})$) for all s because

$$\phi(s)^{\mathrm{t}} = \left(e^{sX}\right)^{\mathrm{t}} = e^{s(X^{\mathrm{t}})} = e^{-sX} = \phi(s)^{-1} ,$$

which insures that $\phi(s)^{\mathrm{t}}\phi(s) \equiv I$.

The subgroup $\mathrm{SO}(n, \mathbb{C})$ is a complex group with $\dim_{\mathbb{C}} \mathrm{SO}(n, \mathbb{C}) = \dim_{\mathbb{C}} \mathfrak{so}(n, \mathbb{C}) = \frac{1}{2}(n^2 - n)$, but $\mathrm{SO}(n, \mathbb{C})$ can also be regarded as a Lie group over $\mathbb{R}$ of twice the dimension, with $\dim_{\mathbb{R}} \mathfrak{so}(n, \mathbb{C}) = 2 \cdot (n^2 - n)$. $\square$

3.16. Example. $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are real Lie groups having the same Lie algebra because $\mathrm{O}(n)$ consists of two connected components (connectedness not yet proved), with $\mathrm{O}(n) = \mathrm{SO}(n) \cup R \cdot \mathrm{SO}(n)$ where $R = \mathrm{diag}(1, \cdots, 1, -1)$ is a real orthogonal matrix with $\det(R) = -1$. The same argument used in 3.15 shows that $\mathfrak{so}(n) = \{X \in \mathrm{M}(n, \mathbb{R}) : X^{\mathrm{t}} = -X\}$; the trace $\mathrm{Tr}(X)$ is automatically zero. It is also clear that $\dim_{\mathbb{R}} \mathrm{SO}(n) = \dim_{\mathbb{R}} \mathfrak{so}(n) = \frac{1}{2}(n^2 - n)$. $\square$

3.17. Example. The Lie algebra of the symplectic group $\mathrm{Sp}(2n, \mathbb{F})$ is

$$\mathfrak{sp}(2n, \mathbb{F}) = \{X \in \mathrm{M}(2n, \mathbb{F}) : \ X^{\mathrm{t}}J + JX = 0\} .$$

where

$$J = \begin{pmatrix} 0 & I_{n\times n} \\ -I_{n\times n} & 0 \end{pmatrix}_{2n\times 2n}$$

Here $2n$ is needed because only even dimensional vector spaces can support non-degenerate skew-symmetric bilinear forms.

Discussion: The inclusion ($\subseteq$) is seen by differentiating the relation that defines $\mathrm{Sp}(n, \mathbb{F})$:

$$A \text{ is non singular and } A^{\mathrm{t}}JA = J$$

Notice that $J^{\mathrm{t}} = J^{-1}$, so $\big(\det(A)\big)^2 = 1$ and $\det(A)$ can only be ± 1 for $A \in \mathrm{Sp}(n, \mathbb{F})$, for $\mathbb{F} = \mathbb{R}$ and $\mathbb{F} = \mathbb{C}$. The condition defining $\mathfrak{g}$ can be rewritten as

$$JXJ^{-1} = JXJ^{\mathrm{t}} = -X^{\mathrm{t}} ,$$

hence $\mathrm{Tr}(X) = 0$ automatically for $X \in \mathfrak{sp}(n, \mathbb{F})$.

To prove the inclusion ($\supseteq$), we rewrite $A^{\mathrm{t}}JA = J$ as $JAJ^{-1} = (A^{\mathrm{t}})^{-1}$. If $X \in \mathrm{M}(n, \mathbb{F})$ satisfies

$$X^{\mathrm{t}}J + JX = 0 \quad (\text{or } JXJ^{-1} = -X^{\mathrm{t}}, \text{ same thing}),$$

then $\phi(s) = e^{sX}$ is a $\mathcal{C}^\infty$ matrix-valued curve such that $\phi(0) = I$, $\phi'(0) = X$, and $\phi(t)$ remains wihtin $\mathrm{Sp}(n, \mathbb{F})$ because

$$J\phi(s)J^{-1} = e^{s(JXJ^{-1})} = e^{-s(X^{\mathrm{t}})} = (e^{-sX})^{\mathrm{t}} = \big(\phi(s)^{\mathrm{t}}\big)^{-1} \quad \text{for all } s \in \mathbb{R}. \ \square$$

X.4. Integral Curves for Vectors Fields.

4.1. Definition. *Let $\tilde{X}$ be a smooth vector field defined on an open subset E in a $\mathcal{C}^\infty$ manifold M, so $\tilde{X}(p)$ is a vector in TM_p at each base point. A* **solution curve** *(or* **integral curve***) for $\tilde{X}$ is any $\mathcal{C}^\infty$ curve $\gamma : (a, b) \to M$ for which*

$$\gamma'(t) = \tilde{X}_{\gamma(t)} \qquad \text{for all } a < t < b,$$

Thus the velocity vector at time t of the moving point $\mathbf{x} = \gamma(t)$ is at all times equal to the vector specified by the vector field at the base point $\gamma(t)$.

When $M = \mathbb{R}^n$ we have the following classical result from ODE, which we cite without proof.

4.2. Theorem (Local Existence and Uniqueness of Solutions). *Let*

$$\tilde{X}_{\mathbf{x}} = \sum_{i=1}^{n} x_i(\mathbf{x})\cdot\mathbf{e}_i \qquad (\mathbf{e}_1, \ldots, \mathbf{e}_n \text{ the standard basis vectors in } \mathbb{R}^n)$$

be a smooth field of vectors on an open subset $E \subseteq \mathbb{R}^n$ and let $p \in E$. Then there is an $\epsilon > 0$ and a $\mathcal{C}^\infty$ map $\gamma : (-\epsilon, \epsilon) \to E$ such that $\gamma(0) = p$ (the **initial condition** *for the solution), and*

$$\frac{d}{dt}\{\gamma(t)\} = \sum_{i=1}^{n} \frac{dx_i}{dt}\cdot\mathbf{e}_i = \tilde{X}_{\gamma(t)}$$

for all $|t| < \epsilon$. The solution is locally unique: any two solutions γ_1, γ_2 satisfying the same initial condition at $t = 0$ must agree in some neighborhood of zero.

If γ_1, γ_2 are solution curves defined on intervals $I_k = (a_k, b_k)$ containing t_0, and if they satisfy the same initial condition at t_0, then $\gamma_1(t) = \gamma_2(t)$ for $t \in (t_0 - \epsilon, t_0 + \epsilon)$, for some $\epsilon > 0$. This leads to a fundamental *global* existence theorem.

4.3. Lemma. *If $\gamma_1(t)$ and $\gamma_2(t)$ are solutions to the initial value problem that agree on some small neighborhood of t_0, they must agree throughout the intersection $I_1 \cap I_2$ of their domains.*

We may assume $t_0 = 0$ because the curves $\eta_k(t) = \gamma_k(t - t_0)$ defined near $t = 0$ have

$$\eta_k'(t) = \gamma_k'(t - t_0) = \tilde{X}_{\gamma_k(t-t_0)} = \tilde{X}_{\eta_n(t)} \ ,$$

so the η_k are solution curves for $\tilde{X}$ defined near $t_0 = 0$ such that $\eta_k(0) = p$.

Now let (α, β) be the largest interval about 0 such that $(\alpha, \beta) \subseteq I_1 \cap I_2 = (a_0, b_0)$ and $\gamma_1 = \gamma_2$ on (α, β). If $\alpha = a_0$, $\beta = b_0$, we are done. If not, suppose $\beta < b_0$. For t near β, γ_1 and γ_2 are both solution curves for $\tilde{X}$ such that $\gamma_1(\beta) = \gamma_2(\beta)$. But then by Lemma 3.20 we have $\gamma_1(t) = \gamma_2(t)$ in some interval about β, which contradicts the definition of β. Likewise if $\alpha > a_0$. $\square$

This simple observation allows us to piece together a "maximal solution" $\gamma(t)$ of the initial value problem.

4.4. Corollary (Maximal Solutions). *Let $\tilde{X}$ be a $\mathcal{C}^\infty$ vector field on an open set $E \subseteq \mathbb{R}^n$. Given $p \in \mathbb{R}^n$ and $t_0 \in \mathbb{R}$, there is a* LARGEST *open interval $I_{\max}$ containing t_0, and a $\mathcal{C}^\infty$ curve $\gamma(t)$ defined on it, such that $\gamma(t) \in E$ for all t in $I_{\max}$ and*

$$\frac{d\gamma}{dt} = \tilde{X}_{\gamma(t)} \text{ for all } t \in I_{\max} \quad \text{and} \quad \gamma(t_0) = p \tag{55}$$

This **maximal solution** *to the initial value problem and its domain $I_{\max}$ are unique.*

Proof: Consider all solutions γ_k to this initial-value problem defined on open intervals I_k containing t_0 (the intervals can be unbounded or even all of $\mathbb{R}$), and let $I_{\max} = \bigcup_k I_k$. By Lemma 3.19, the solutions agree wherever their domains I_k overlap and can be pieced together to get a well defined solution curve $\gamma(t)$ on $I_{\max}$. Obviously the solution cannot be extended to any larger interval containing t_0. $\square$

It can happen that the maximal domain $I_{\max}$ for a solution to the initial value problem (55) is not all of $\mathbb{R} = (-\infty, \infty)$. Examples show that $\gamma(t)$ can run off to ∞ in finite time, or run out of the set E on which $\tilde{X}$ is defined, or display other singular behavior that forces $I_{\max} \neq \mathbb{R}$.

If we identify $\mathrm{M}(n, \mathbb{F}) \simeq \mathbb{R}^{n^2}$ then $\mathrm{GL}(n, \mathbb{F})$ is an open dense set in matrix space, and if we specify a tangent vector $X \in \mathrm{T}(\mathrm{GL})_I = \mathfrak{gl}(n, \mathbb{F}) = \mathrm{M}(n, \mathbb{F})$ at the identity element in GL, left translates of X by elements of GL determine a unique field of tangent vectors on GL

$$\tilde{X}_g = \lambda_g(X) = g \cdot X \qquad \text{for } g \in \mathrm{GL}\ .$$

In this setting we can find explicit solutions for initial value problems of the form (55) posed for left-invariant vector fields on the general linear group $\mathrm{GL}(n, \mathbb{F})$, and in fact we will show that this can also be done for left-invariant vector fields on any matrix Lie group $G \subseteq \mathrm{M}(n, \mathbb{F})$. Because the underlying manifold for $\tilde{X}$ is a *group* the maximal solutions $\gamma(t)$ for such fields are defined for all $-\infty < t < \infty$ (they never become singular) and are given by the exponential map for matrices $\mathrm{Exp} : \mathrm{M}(n, \mathbb{F}) \to \mathrm{GL}(n, \mathbb{F})$.

4.5. Lemma. *When $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ the* **matrix exponential series**

$$\mathrm{Exp}(tA) = e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k \qquad \text{for } t \in \mathbb{R}$$

is absolutely norm convergent for all $A \in \mathrm{M}(n, \mathbb{F})$, and is a $\mathcal{C}^\infty$ map from $t \in \mathbb{R}$ (or $t \in \mathbb{C}$) into matrix space equipped with any convenient norm. Furthermore, if $X \in \mathrm{M}(n, \mathbb{F})$ and if $\tilde{X}$ is the unique left-translation invariant $\mathcal{C}^\infty$ vector field on GL such that $\tilde{X}_I = X$, then

$$\gamma(t) = e^{tX} = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k \quad \text{for } t \in (-\infty, \infty)$$

is the unique $\mathcal{C}^\infty$ solution curve for $\tilde{X}$ such that $\gamma(0) = I$.

Note: Given any $A \in \mathrm{M}(n, \mathbb{F})$, the curve $\gamma(t) = e^{tA}$ is a **one-parameter group** in GL in the sense that

$$\text{EXPONENT LAW:} \quad \gamma(t_1 + t_2) = \gamma(t_1) \cdot \gamma(t_2) \quad \text{and} \quad \gamma(0) = I, \gamma(1) = A. \tag{56}$$

In particular $\gamma(0) = I, \gamma(-t) = \gamma(t)^{-1}$ exists, and $\det(e^{tA}) \neq 0$ for all A and t. $\square$

We defer the proof of 4.5 while we review basic properties of the matrix exponential map $\mathrm{Exp} : \mathrm{M}(n, \mathbb{F}) = \mathfrak{gl}(n, \mathbb{F}) \to \mathrm{GL}(n, \mathbb{F}) \subseteq \mathrm{M}(n, \mathbb{F})$. We will then show there is an analogous "exponential map" associated with any matrix Lie group G. This map $exp : \mathfrak{g} \to G$ is $\mathcal{C}^\infty$ on all of the vector space $\mathfrak{g}$, and is a $\mathcal{C}^\infty$ diffeomorphism between some open neighborhood of the zero element in the Lie algebra $\mathfrak{g}$ and an open neighborhood of the identity element $e \in G$. Existence of this correspondence allows us to transfer many questions posed on the (nonlinear) manifold G to problems on the linear space $\mathfrak{g}$, where the tools of linear algebra can be used to find a solution that can be transferred back to the original group G via the exp map.

Properties of Matrix Exponential Map.

When $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ there are various natural norms on the vector space $V = \mathrm{M}(n, \mathbb{F})$ which allow us to discuss the distance $\|A - B\|$ between two matrices, and convergence $A_n \to A$ of matrices as $n \to \infty$. By definition, all norms have the properties

(57)

1. $\|X\| \geq 0$ and $\|X\| = 0 \Leftrightarrow X = 0$;
2. $\|\lambda A\| = |\lambda| \cdot \|A\|$ for $\lambda \in \mathbb{F}$;
3. $\|A \pm B\| \leq \|A\| + \|B\|$.

When V is finite-dimensional the choice of norm is unimportant because all norms on such spaces are equivalent, in the sense that there are constants $C, D > 0$ such that

$$\|x\|_2 \leq C \cdot \|x\|_1, \quad \|x\|_1 \leq D \cdot \|x\|_2 \quad \text{for all } x \in V\ . \tag{58}$$

Thus we have convergence $A_n \to A$ in one norm if and only if $A_n \to A$ in the other.

4.6. Exercise. Prove that (58) holds for any two norms on a finite dimensional vector space.
Hint: It suffices to compare a given norm $\|\cdot\|$ with the "Euclidean " norm $\|\cdot\|_2$ determined by a basis $\mathfrak{X} = \{e_k\}$ in V:

$$\|v\|_2 = \big\|\sum_{k=1}^n c_k e_k\big\|_2 = \Big(\sum_{k=1}^n |c_i|^2\Big)^{1/2}$$

Use the Reverse Triangle Inequality to show that the map $v \to \|v\|$ is continuous on $(V, \|\cdot\|_2)$. Then $m = \min\{\|v\| : \|v\|_2 = 1\}$ and $M = \max\{\|v\| : \|v\|_2 = 1\}$ are achieved on the compact set $S = \{v : \|v\|_2 = 1\}$. Etc. $\square$

On an associative algebra like $\mathrm{M}(n, \mathbb{F})$, in which both sums $(+)$ and products $(\cdot)$ are defined, it is really convenient to work with a norm $\|A\|$ having the multiplicative property

$$\|AB\| \leq \|A\| \cdot \|B\|$$

(It would desirable, but not necessary, to have $\|I\| = 1$ too.) Such norms exist. One is the **Hilbert-Schmidt norm**

$$\|A\|_{\mathrm{HS}} = \Big(\sum_{i,j=1}^n |a_{ij}|^2\Big)^{1/2}$$

which is the norm associated with the inner product $(A, B)_{\mathrm{HS}} = \mathrm{Tr}(B^*A)$. Another choice is the **operator norm** with respect to a given norm on $\mathbb{F}^n$,

$$\|A\|_{\mathrm{op}} = \max\{\|Ax\| : x \in \mathbb{F}^n, \|x\| \leq 1\}$$

The norm properties (57) and (58) are easily verified and we obviously have $\|I\|_{\mathrm{op}} = 1$.

4.7. Exercise. Show that $\mathrm{Tr}(B^*A)$ is an inner product on matrix space $\mathrm{M}(n, \mathbb{F})$, and

that $\|A\|_{\text{HS}}$ is the associated norm. Then verify that $\|AB\|_{\text{HS}} \leq \|A\|_{\text{HS}} \cdot \|B\|_{\text{HS}}$ for all A, B.
Hint: Schwartz inequality for inner products over $\mathbb{R}$ or $\mathbb{C}$.
Note: $\|I\|_{\text{HS}} = \sqrt{n}$. $\square$

For some purposes the most popular choice is the operator norm $\|A\|$ obtained by regarding A as the left-multiplication operator $L_A : \mathbb{F}^n \to \mathbb{F}^n$ and taking the Euclidean norm $\|v\|_2 = \big(\sum_{i=1}^n |v_i|^2\big)^{1/2}$ on $\mathbb{F}^n$. The operator norm $\|A\| = \max\{\|Ax\|_2 : \|v\|_2 \leq 1\}$ is the maximal length $\|Av\|_2$ of the image vector as v runs though vectors in the unit ball $B_1 = \{v : \|v\|_2 \leq 1\}$ in $\mathbb{F}^n$ (or the unit sphere $S = \{v : \|v\|_2 = 1\}$). Obviously, $\|I\| = 1$ in this norm. $\square$

4.8. Exercise. Verify that operator norm has the properties (57) and that $\|AB\| \leq \|A\| \cdot \|B\|$. $\square$

Despite its convenient properties, the operator norm is less convenient if you need to compute its numerical value.

4.9. Exercise. If $A = [a_{ij}]$ and $\mathbb{C}^n$ is given the Euclidean norm $\|v\|_2$, find a formula for computing $\|A\|_{\text{op}}$ in terms of the matrix coefficients a_{ij}. $\square$

By (58), $\mathrm{M}(n, \mathbb{F})$ is a complete metric space in either of these norms, for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. If $A \in \mathrm{M}(n, \mathbb{F})$ the exponential series $\sum_{n=0}^{\infty} \frac{1}{n!} A^n$ is then **absolutely norm convergent**, with

$$\sum_{n=0}^{\infty} \frac{1}{n!} \|A\|^n < \infty \ ,$$

by the norm triangle inequality. This immediately forces norm convergence of the series because the partial sums $S_n = I + A + \cdots + \frac{1}{n!} A^n$ form a Cauchy sequence in $\mathrm{M}(n, \mathbb{F})$, with

$$\|S_m - S_n\| = \Big\|\frac{1}{m!} A^m + \ldots + \frac{1}{(n+1)!} A^{n+1}\Big\| \leq \sum_{k=n+1}^{m} \frac{1}{k!} \|A\|^k \ .$$

Then $\|S_m - S_n\| \to 0$ as $m, n \to 0$ because

$$\sum_{k=0}^{\infty} \frac{1}{k!} \|A\|^k = e^{\|A\|} < \infty \ .$$

Similarly, if a power series is absolutely convergent with $\sum_{n=0}^{\infty} |a_n||z|^n < \infty$ for $|z| < R$, we may conclude that $\sum_{n=0}^{\infty} a_n A^n$ is absolutely norm convergent for all A such that $\|A\| < R$. As an example we have

4.10. Example. Let $\|A\| < 1$. The Taylor series for $\mathrm{Log}(1+z)$ about $z = 0$ is

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1} z^n}{n} \quad \text{and similarly} \quad \mathrm{Log}(1-z) = (-1) \cdot \sum_{n=1}^{\infty} \frac{z^n}{n} \ .$$

The function $\mathrm{Log}(z)$ agrees with the series sum for all z inside the radius of convergence $R = 1$, so we can define $\mathrm{Log}(A)$ for all A such that $\|A - I\| < 1$ by writing

$$\mathrm{Log}(A) = \mathrm{Log}(I + (A - I)) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n} (A - I)^n$$

(absolutely norm convergent for $\|A - I\| < I$).

Chapter XI. Tensor Fields & Vector Calculus.

... or, what they never told you in Calculus III.

XI.1 Tangent Vectors, Cotangent Vectors and Tensors.

This chapter explores some aspects of linear algebra that lie at the heart of modern differential geometry, and ends with a reinterpretation of many of the the main results of multivariate Calculus. This is a vast subject, so the presentations of this chapter will not be as fully developed as those in preceding chapters.[1] Some of the missing details can be found in Appendices A - C at the end of this chapter.

Euclidean n-dimensional space E_n is a featureless space: a line, a plane, etc, perhaps equipped with a *metric*, a distance function $d(x,y)$ between points. By marking an origin and imposing coordinates we can model E_n as n-tuples of real numbers $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$. In addition to allowing us to describe locations of points p as coordinate n-tuples, $\mathbb{R}^n$ comes equipped with certain algebraic operations – *scaling* and *vector addition* – which make Euclidean space into a vector space over $\mathbb{R}$. This extra structure, unknown to Euclid, was inspired by the "parallelogram law" for vector addition that arose in physics as the correct law for adding forces. In the late 1800's additional operations on coordinate space $\mathbb{R}^n$ were introduced, such as the inner product $(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i$ (for arbitrary dimensions) and the cross product $\mathbf{x} \times \mathbf{y}$ (which makes sense only in $n = 3$ dimensions).

The tangent space TE_p to E_n at a base point p was thought of as a copy of the vector space $\mathbb{R}^n$ attached to E_n at p; elements in TE_p are sometimes described as pairs $(p, \mathbf{x})$ where p is a base point in E and $\mathbf{x}$ a vector in $\mathbb{R}^n$. There is a separate tangent space attached to each base point in E. In Calculus tangent vectors based at p are thought of as "arrows" attached to the base point. These can be scaled and added to other tangent vectors attached to the same base point via the rules

$$\lambda\cdot(p, \mathbf{x}) = (p, \lambda\mathbf{x}) \quad \text{and} \quad (p, \mathbf{x}_1) + (p, \mathbf{x}_2) = (p, \mathbf{x}_1 + \mathbf{x}_2)$$

However, *there is no meaningful way to add tangent vectors* $(p, \mathbf{x}) \in \mathrm{TE}_p$ *and* $(q, \mathbf{y}) \in \mathrm{TE}_q$ *attached to different base points* $p \neq q$. When elements of TE_p are viewed as arrows, those arrows might represent force vectors acting at p, or the velocity of a moving particle as it passes through p. They could also represent more general fields: an electric field, magnetic field, gravitational field, or a field of stress tensors that pervade some solid medium, etc.

Many mathematical models are concerned with *fields of vectors* $\mathbf{F}(\mathbf{x})$ defined on E_n (or some open subset). These assign a tangent vector $\mathbf{F}(p) \in \mathrm{TE}_p$ to each base point, and the coordinates imposed on E_n determine basis vectors $\{\mathbf{e}_1(p), \cdots, \mathbf{e}_n(p)\}$ in the tangent space TE_p that allow us to describe any tangent vector at p as a linear combination $\mathbf{a} = a_1\mathbf{e}_1 + \ldots + a_n\mathbf{e}_n$ ($a_i \in \mathbb{R}$). Similarly we can describe a field of vectors on E_n as

$$\mathbf{F}(p) = F_1(p)\,\mathbf{e}_1 + \cdots F_n(p)\,\mathbf{e}_n \ ,$$

where the scalar-valued coefficients $F_k(p)$ vary with the base point. If $\mathbf{a}, \mathbf{b} \in \mathrm{TE}_p$ and

[1] This chapter is a "compactified" version presenting highlights of Class Notes developed over several semesters for a full semester NYU Senior Honors course *Vector Analysis: Calculus on Manifolds*. A pdf version of those notes for personal use is available on request from Prof. Greenleaf at `fred.greenleaf@nyu.edu` who should be contacted regarding classroom use.

$n = 3$ their scalar multiples and vector sums are given by

$$\lambda \mathbf{a} = \sum_{i=1}^{n} \lambda a_i \, \mathbf{e}_i \qquad \text{and} \qquad \mathbf{a} + \mathbf{b} = \sum_{i=1}^{3} (a_i + b_i) \, \mathbf{e}_i \ ,$$

and their cross product is

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \\ &= (a_2 b_3 - a_3 b_2) \, \mathbf{e}_1 + (a_1 b_3 - a_3 b_1) \, \mathbf{e}_2 + (a_1 b_2 - a_2 b_1) \, \mathbf{e}_3 \end{aligned}$$

Scalar fields on E_n are just scalar-valued functions $F : E_n \to \mathbb{R}$.

This coordinate description of vector fields on $\mathbb{R}^3$ or $\mathbb{R}^n$ is used in Calculus to define the standard "vector operations" on vector fields $\mathbf{y} = \mathbf{F}(\mathbf{x}) = \sum_{i=1}^{n} F_i(\mathbf{x}) \, \mathbf{e}_i$ that have differentiable coefficients. Writing $D_{x_i}\phi(p)$ for the i^{th} partial derivative $\partial\phi/\partial x_i(p)$ of a scalar-valued function $\phi : \mathbb{R}^m \to \mathbb{R}$, we define the operations

1. Gradient:

   $$\mathbf{grad}\, F(p) = \nabla F(p) = \sum_{i=1}^{n} D_{x_i} F(p) \cdot \mathbf{e}_i$$

   for scalar fields $F : \mathbb{R}^n \to \mathbb{R}$.

2. Curl:

   $$\begin{aligned} \mathbf{curl}\, \mathbf{F}(p) &= \nabla \times \mathbf{F}(p) \ = \ \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial/\partial_{x_1} & \partial/\partial_{x_2} & \partial/\partial_{x_3} \\ F_1 & F_2 & F_3 \end{pmatrix} \\ &= \Big(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \Big) \mathbf{e}_1 \ - \ \Big(\frac{\partial F_3}{\partial x_1} - \frac{\partial F_1}{\partial x_3} \Big) \mathbf{e}_2 + \Big(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \Big) \mathbf{e}_3 \end{aligned}$$

   for smooth vector fields $\mathbf{F} = F_1 \, \mathbf{e}_1 + \ldots + F_3 \, \mathbf{e}_3$. This definition only works for fields on 3-dimensional space $\mathbb{R}^3$, though it can be adapted to deal with vector fields on the plane $\mathbb{R}^2$.

3. Divergence:

   $$\mathbf{div}\, \mathbf{F}(p) = \nabla \circ \mathbf{F}(p) = \frac{\partial F_1}{\partial x_1}(p) \, \mathbf{e}_1 + \ldots + \frac{\partial F_3}{\partial x_3}(p) \, \mathbf{e}_3 = \mathrm{Trace}\Big[\frac{\partial F_i}{\partial x_j} \Big](p)$$

   for smooth vector fields $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$. The result is a scalar field. This also works only for fields on 3-dimensional Euclidean space, but can be adapted to work for fields on $\mathbb{R}^2$.

Some question lurk in the background of these definitions.

> **Question:** *All these operators are described in terms of Cartesian coordinates imposed on the blank slate of Euclidean space. What happens in polar coordinates or other coordinates systems? What bases in the tangent spaces* TE_p *are induced by these alternative coordinates on* E_n*? How should the vector operations* ∇F*,* $\nabla \times \mathbf{F}$*,* $\nabla \circ \mathbf{F}$ *be described in non-Cartesian coordinates?*

One might also wonder what analogs of these vector operators that pervade physics might exist in dimensions $n \geq 4$, where the classical definitions of **curl** and **div** no longer make sense.

Calculus on Manifolds. Flat and featureless Euclidean space is a very limited setting for developing mathematical physics, or for understanding the possible meaning of "tangent vectors." It is simply too special to reveal the subtleties involved. For instance, suppose you wanted to model phenomena on the spherical surface of the Earth.

- What does a vector field on the 2-dimensional sphere
$$S^2 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$$
mean? How about vector fields on the 3-dimensional sphere
$$S^3 = \{\mathbf{x} \in \mathbb{R}^4 : \|\mathbf{x}\|^2 = 1\}$$
that lives in 4-dimensional space?
- What becomes of the classical operations on vector fields in this higher dimensional setting?

The same questions apply to any curved lower-dimensional hypersurface embedded in $\mathbb{R}^n$.

There are deeper and more fundamental issues to deal with in a consistent re-imagining of Calculus. Suppose we regard S^2 or S^3 as the entire universe of discourse. For instance, in general relativity the universe can be viewed as a 3-dimensional sphere S^3 with a metric structure that evolves with time. There is no outside space $\mathbb{R}^4$ in which this object lives, and any reference to it is pure metaphysics, divorced from any phenomena observable from within S^3. "Tangent vectors" or "tangent spaces" viewed as arrows or hyperplanes attached to a base point $p \in S^3$ and extending into the "surrounding space" simply have no meaning. There is no "outside" to the universe. So, let's see how Calculus might be developed using concepts intrinsic to S^2 or S^3, or more general spaces that might be used to model phenomena.

Local Coordinates. An m-dimensional locally Euclidean space M is a set of points that can be covered with **coordinate charts** (x_α, U_α), each consisting of an open set $U_\alpha \subseteq M$ and a continuous map $x_\alpha : U_\alpha \rightarrow \mathbb{R}^m$ that assigns coordinates $\mathbf{x} = x_\alpha(u) = (x_1, \ldots, x_m)$ in $\mathbb{R}^m$ to each point $u \in U_\alpha$. We require

(1)

1. The chart domains U_α cover all of M.
2. $V_\alpha = x_\alpha(U_\alpha)$ is an open set in $\mathbb{R}^m$ for each index α.
3. The chart map x_α is a bicontinuous bijection (a *homeomorphism*) from U_α to the open set $V_\alpha = x_\alpha(U_\alpha)$ in coordinate space $\mathbb{R}^m$.

Then $x_\alpha : U_\alpha \rightarrow V_\alpha \subseteq \mathbb{R}^m$ and its inverse $x_\alpha^{-1} : V_\alpha \rightarrow U_\alpha \subseteq M$ are both continuous maps. Thus a locally Euclidean space M looks exactly like $\mathbb{R}^m$ near any base point $p \in M$; in particular, Cartesian coordinates can be imposed on M near p. But more is needed to do Calculus on M – in order to discuss derivatives of functions on M the charts (x_α, U_α) should be "differentiably related" whenever two chart domains U_α and U_β overlap.

1.1. Definition. *For $k \in \mathbb{N}$ a* $\mathcal{C}^{(k)}$**-structure** *on an m-dimensional locally Euclidean space M is a covering by charts (x_α, U_α), α in some index set I, such that the* **coordinate transition maps** *between overlapping charts (x_α, U_α) and (y_β, U_β),*
$$x_\alpha \circ y_\beta^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^m \quad \text{and} \quad y_\beta \circ x_\alpha^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^m ,$$
are of class $\mathcal{C}^{(k)}$ on the open sets in coordinate space $\mathbb{R}^m$ where they are defined. Recall that a map $\mathbf{y} = f(\mathbf{x})$ from $\mathbb{R}^m \rightarrow \mathbb{R}^n$ is class $\mathcal{C}^{(k)}$ on an open set $U \subseteq \mathbb{R}^m$ if all partial derivatives of the scalar components $f_i : U \rightarrow \mathbb{R}$ in $\mathbf{y} = f(\mathbf{x}) = \big(f_1(\mathbf{x}), \cdots, f_m(\mathbf{x})\big)$,
$$D_x^\alpha f_i = D_{x_1}^{\alpha_1} \cdot \ldots \cdot D_{x_m}^{\alpha_m} f_i(\mathbf{x}) \quad \text{where } \alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{Z}_+^m ,$$

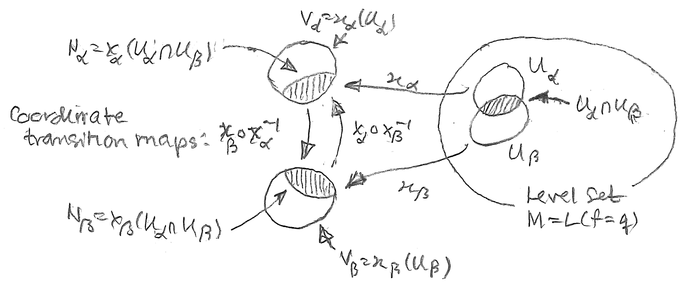

Figure 11.1. The coordinate transition maps $x_\alpha \circ y_\beta^{-1}$ and $y_\beta \circ x_\alpha^{-1}$ between two charts (x_α, U_α) an (y_β, U_β), and their (shaded) domains of definition N_α, N_β in $\mathbb{R}^m$ are shown. The domains N_α, N_β live in different copies of coordinate space $\mathbb{R}^m$; both correspond to the intersection $U_\alpha \cap U_\beta$ of the chart domains, which is an open set in the locally Euclidean space M.

exist and are continuous for $1 \leq i \leq m$ *and all* m*-tuples* $\alpha = (\alpha_1, \cdots, \alpha_m)$ *of total degree* $|\alpha| = \alpha_1 + \ldots + \alpha_m \leq k$. *The "exponents"* $\alpha = (\alpha_1, \cdots, \alpha_m)$ *are called multi-indices and* D_x^α *is a partial derivative of total degree* $|\alpha| = \alpha_1 + \ldots + \alpha_m$. *By convention,* $D_{x_i}^{\alpha_i} = I$ *(the identity operator) if* $\alpha_i = 0$*, and the multi-index* $\alpha = (0, \cdots, 0)$ *yields* $D_x^\alpha = I$ *(the identity operator).*

The meaning of the coordinate transition maps is illustrated in Figure 11.1.

Once we have a "founding set" of $\mathcal{C}^\infty$-related charts $\mathcal{U} = \{(x_\alpha, U_\alpha) : \alpha \in I\}$ that cover M, these determine a **maximal atlas** of consistent charts on M.

> *The* **maximal atlas** $\overline{\mathcal{U}}$ *determined by* $\mathcal{U}$ *consists of all charts* (ϕ, W) *with* $\phi : W \to \mathbb{R}^m$ *as in* (1) *such that* (ϕ, W) *is differentiably related to all the founding charts* $(x_\alpha, U_\alpha) \in \mathcal{U}$.

1.2. Lemma. *All charts* (ϕ, W), (ψ, W') *in the atlas* $\overline{\mathcal{U}}$ *are* $\mathcal{C}^{(k)}$*-related to each other, so coordinate transition maps* $\phi \circ \psi^{-1}$ *and* $\psi \circ \phi^{-1} : \mathbb{R}^m \to \mathbb{R}^m$ *are* $C^{(k)}$ *where defined if the chart domains* W, W' *overlap.*

Proof: This follows from the classical chain rule: If $p \in W \cap W'$ and (x_α, U_α) is a founding chart in $\mathcal{U}$ that contains p, we can split

$$\phi \circ \psi^{-1} = (\phi \circ x_\alpha^{-1}) \circ (x_\alpha \circ \psi^{-1})$$

The factors $(\phi \circ x_\alpha^{-1})$, $(x_\alpha \circ \psi^{-1})$ are both $\mathcal{C}^{(k)}$ maps between copies of $\mathbb{R}^m$ because ϕ, ψ are $\mathcal{C}^\infty$-related to the founding chart (U_α, x_α). $\square$

1.3. Example. Consider the 2-sphere $M = S^2$ in $\mathbb{R}^3$. A family of charts that cover M is given by $(H_i^\pm, P_i)$ where $H_i^\pm$ are the open hemispheres

$$H_i^+ = \{\mathbf{x} \in S^2 : x_i > 0\} \qquad H_i^- = \{\mathbf{x} \in S^2 : x_i < 0\} \qquad (1 \leq i \leq 3)$$

and P_i = the restriction to S^2 of the projection maps $P_i : \mathbb{R}^3 \to \mathbb{R}^2$

$$P_1(\mathbf{x}) = (x_2, x_3), \ \cdots \ , P_3(\mathbf{x}) = (x_1, x_2)$$

that project out the i^{th} coordinate of $\mathbf{x} = (x_1, x_2, x_3)$, see Figure 11.2 for an illustration.

In each case, *range*(P_i) is the open unit disc $D = \{(s_1, s_2) \in \mathbb{R}^2 : s_1^2 + s_2^2 < 1\}$. Obviously the restriction $P_i|_{H_i^\pm} : H_i^\pm \to D$ is one-to-one and continuous onto D and is the restriction of the $\mathcal{C}^\infty$ map $P_i : \mathbb{R}^3 \to \mathbb{R}^2$. Its inverse can be calculated explicitly, for instance

$$(P_3|_{H_3^+})^{-1}(s_1, s_2) = \left(s_1, s_2, \sqrt{1 - (s_1^2 + s_2^2)}\right) \quad \text{for all } \mathbf{s} \in D$$

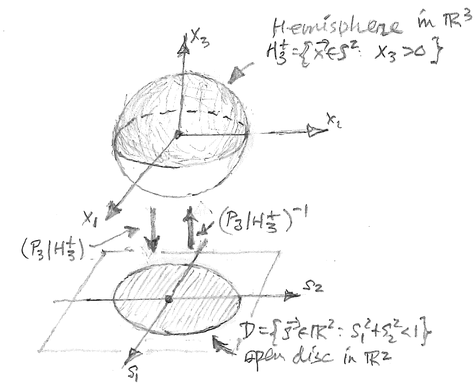

Figure 11.2. One of the charts (x_α, U_α) that determine the standard manifold structure for the smooth hypersurface $M = S^2$, a 2-dimensional sphere embedded in $\mathbb{R}^3$. The chart domain is the hemisphere

$$U_\alpha = H_3^+ = \{\mathbf{x} : \|\mathbf{x}\| = 1 \text{ and } x_3 > 0\} \ .$$

The chart map $x_\alpha = P_3|H_3^+$ sending $\mathbf{x} \to (x_1, x_2)$ projects the hemisphere U_α onto an open disc D of radius $r = 1$ in the x_1, x_2-plane.

is clearly a $\mathcal{C}^\infty$ map from $D \subseteq \mathbb{R}^2$ onto H_3^+. The coordinate transition maps between two such charts, say $x_\alpha = (P_3|_{H_3^+})$ and $y_\beta = \big(P_2|_{H_2^-}\big)$, are then

$$x_\alpha \circ y_\beta^{-1}(\mathbf{s}) = P_3\Big(s_1, -\sqrt{1 - (s_1^2 + s_2^2)}, s_2\Big) = \Big(s_1, -\sqrt{1 - (s_1^2 + s_2^2)}\Big)$$

and

$$y_\beta \circ x_\alpha^{-1}(\mathbf{s}) = P_2\Big(s_1, s_2, +\sqrt{1 - (s_1^2 + s_2^2)}\Big) = \Big(s_1, +\sqrt{1 - (s_1^2 + s_2^2)}\Big)$$

for $\mathbf{s} = (s_1, s_2) \in D$. Both are $\mathcal{C}^\infty$ maps from $\mathbb{R}^2 \to \mathbb{R}^2$.

In the maximal atlas we find other charts, for example, a chart imposing spherical coordinates (see Figure 11.3) on the open hemisphere H_1^+: if $\mathbf{x} = (x_1, x_2, x_3) \in S^2$ with $x_1 > 0$, we define a chart (z_γ, U_γ) on $U_\gamma = H_1^+$: for $\mathbf{x} \in S^2$ we let

$$(\theta, \phi) = z_\gamma(x_1, x_2, x_3) = \Big(\arctan\big(\frac{x_2}{x_1}\big), \arcsin(x_3)\Big)$$

so the transition map from $(x_\alpha, U_\alpha) = (P_3, H_3^+)$ to (z_γ, U_γ), $U_\gamma = H_1^+$ is

$$\begin{aligned}(\theta, \phi) &= z_\gamma \circ x_\alpha^{-1}(s_1, s_2) = z_\gamma\Big(s_1, s_2, \sqrt{1 - (s_1^2 + s_2^2)}\Big) \\ &= \Big(\arctan\big(\frac{s_2}{s_1}\big), \arcsin\big(\sqrt{1 - (s_1^2 + s_2^2)}\big)\Big)\end{aligned}$$

The inverse map can be handled similarly; both are $\mathcal{C}^\infty$ from $\mathbb{R}^2 \to \mathbb{R}^2$. $\square$

The maximal atlas $\overline{\mathcal{U}}$ determined by founding charts includes all possible determinations of spherical coordinates on open sets $U \subseteq S^2$ that avoid the north and south poles $N = (0, 0, 1)$ and $S = (0, 0, -1)$.

Smooth Functions and Mappings on M.

Once M is equipped with a differentiable structure (a maximal atlas of differentiably related covering charts), we can define

1. Smooth Functions $f : M \to \mathbb{R}$.

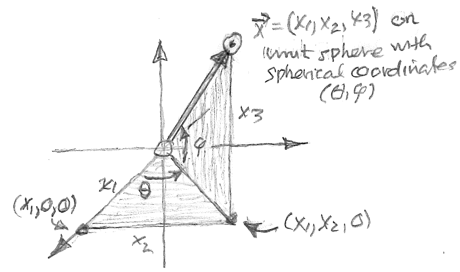

Figure 11.3. Spherical coordinates (θ, ϕ) on the two-dimensional sphere $S^2 = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$ are assigned as shown. We have chosen a somewhat non-standard definition of ϕ, one that agrees with the way latitude is assigned on the surface of the Earth. Both angles are ambiguous up to a multiple of 2π radians ($= 360°$); to get definite values, our convention restricts $-\pi < \theta < \pi$ and $-\pi < \phi < \pi$. The angles θ and ϕ cannot be defined at the north pole $N = (0, 0, 1)$ or south pole $S = (0.0. -1)$ on the sphere.

2. Smooth Parametrized Curves, $\gamma : \mathbb{R} \to M$.

3. Smooth Mappings $\phi : M \to N$ between differentiable manifolds M and N.

1.4. Definition. *A function $f : M \to \mathbb{R}$ is of class $\mathcal{C}^\infty$, indicated by writing $f \in \mathcal{C}^\infty(M)$, if $f \circ x_\alpha^{-1} : V_\alpha \to \mathbb{R}$ is $\mathcal{C}^\infty$ for every chart – i.e. f is smooth (of class $\mathcal{C}^\infty$) if it is smooth in the classical sense when described in local chart coordinates. The same definition, applied to any open set $U \subseteq M$, determines the space $\mathcal{C}^\infty(U)$ of smooth scalar valued functions on U. These are not only infinite-dimensional vector spaces, they are also associative algebras since they are closed under forming pointwise products $(f \cdot h)(u) = f(u)h(u)$, as well as scalar multiples $\lambda \cdot f$, sums $f_1 + f_2$, and linear combinations $\sum_{i=1}^r c_i f_i$. By the classical chain rule, smoothness of $f : M \to N$ does not depend on a particular choice of local coordinates.*

We now call attention to a special class of functions associated with a base point $p \in M$.

1.5. Definition (The Local Algebra $\mathcal{C}^\infty(p)$). *The* **local algebra** *$\mathcal{C}^\infty(p)$ at $p \in M$ is the associative algebra of $\mathcal{C}^\infty$ functions defined at and near p. A typical element in $\mathcal{C}^\infty(p)$ is a pair (f, U) involving an open neighborhood U of p and a $\mathcal{C}^\infty$ function $f : U \to \mathbb{R}$. We then define $(+)$, $(\cdot)$, and scaling operations*

$$
\begin{aligned}
\lambda \cdot (f, U) &= (\lambda \cdot f, U), \qquad \textit{for all } \lambda \in \mathbb{R} \\
(f_1, U_1) + (f_2, U_2) &= (f_1 + f_2, U_1 \cap U_2) \\
(f_1, U_1) \cdot (f_2, U_2) &= (f_1 \cdot f_2, U_1 \cap U_2) \quad \textit{(pointwise product)}
\end{aligned}
$$

on $\mathcal{C}^\infty(p)$. The zero element in this vector space is the pair $(0, M)$ and $\mathbf{1} = (1, M)$ is the multiplicative identity element. Note that $\mathcal{C}^\infty(p)$ contains the algebra $\mathcal{C}^\infty(M)$ of globally $\mathcal{C}^\infty$ functions on M, as well as $\mathcal{C}^\infty(U)$ for any open set U that contains p.

1.6. Definition. *A map $\phi : M \to N$ between differentiable manifolds M, N is* **smooth near** *p if it becomes a $\mathcal{C}^\infty$ map from $\mathbb{R}^m \to \mathbb{R}^n$ when described in local coordinates, so that*

$$
y_\beta \circ \phi \circ x_\alpha^{-1} : \mathbb{R}^m \to \mathbb{R}^n \qquad \textit{is } \mathcal{C}^\infty \textit{ in the classical sense} \ ,
$$

for all charts (x_α, U_α) about p and (y_β, U_β) about $q = \phi(p)$. The map is $\mathcal{C}^\infty$ on an open subset (or on all of M), if it is $\mathcal{C}^\infty$ near each base point in U.

1.7. Definition. *A* **parametrized curve** $\gamma : \mathbb{R} \to M$ *is of class* $\mathcal{C}^{(k)}$ *if, for all charts* (x_α, U_α) *on* M, $x_\alpha \circ \gamma(t)$ *is a* $\mathcal{C}^{(k)}$*-map from* $\mathbb{R} \to \mathbb{R}^m$ *on some open interval* $(a, b) \subseteq \mathbb{R}$. *It is* $\mathcal{C}^{(k)}$ *on a closed interval* $[a, b]$ *if there is a slightly larger interval* $(a - \epsilon, b + \epsilon)$ *on which* γ *is defined and of class* $\mathcal{C}^{(k)}$. *In local coordinates* γ *becomes a classical* $\mathcal{C}^{(k)}$ *curve from* $\mathbb{R} \to \mathbb{R}^m$

$$\tilde{\gamma}(t) = x_\alpha \circ \gamma(t) = \big(x_1(t), \cdots, x_m(t)\big)$$

with scalar components $x_k(t)$ *that are differentiable of class* $\mathcal{C}^{(k)}$.

The classical derivative $d\gamma/dt$ of a smooth curve $\gamma : [a, b] \to \mathbb{R}^m$ is

$$\begin{aligned}
\frac{d\gamma}{dt}(t_0) &= \lim_{\Delta t \to 0} \frac{\gamma(t_0 + \Delta t) - \gamma(t_0)}{\Delta t} = \\
&= \lim_{\Delta t \to 0} \frac{f \circ \gamma(t_0 + \Delta t) - f \circ \gamma(t_0)}{\Delta t} = \Big(\frac{dx_1}{dt}(t_0), \cdots, \frac{dx_m}{dt}(t_0)\Big)\ , \qquad (2)
\end{aligned}$$

But this makes no sense at all for smooth curves $\gamma : [a, b] \to M$ with values in a manifold; M is not a vector space and differences of points in M are undefined. (Try interpreting the difference quotients $\Delta\gamma/\Delta t$ in (2) for a curve that lives within in the 2-sphere $M = S^2 \subseteq \mathbb{R}^3$, without making any reference to the surrounding space $\mathbb{R}^3$.) We can, however, make sense of $\gamma'(t_0)$ as an *operator* $\gamma'(t_0) : \mathcal{C}^\infty(p) \to \mathbb{R}$ that acts on the local algebra of $\mathcal{C}^\infty$ functions defined near $p = \gamma(t_0)$, by defining

$$\langle \gamma'(t_0), f \rangle = \frac{d}{dt}\{f \circ \gamma(t)\}\Big|_{t=t_0} \qquad (3)$$

The operator $\gamma'(t_0)$ sends f to the time derivative of the f-values $f(\gamma(t))$ seen as t increases past t_0.

Given a chart (x_α, U_α) about p the action of $\gamma'(t_0)$ on functions can be computed by writing $f \circ \gamma(t)$ as a composite $(f \circ x_\alpha^{-1}) \circ (x_\alpha \circ \gamma) : \mathbb{R} \to \mathbb{R}^m \to \mathbb{R}$ and applying the classical chain rule. If $x_\alpha \circ \gamma(t) = \big(x_1(t), \ldots, x_m(t)\big)$ we get

$$\begin{aligned}
\langle \gamma'(t), f \rangle &= \frac{d}{dt}\{f \circ \gamma(t)\}\Big|_{t=t_0} \\
&= \frac{d}{dt}\{(f \circ x_\alpha^{-1}) \circ (x_\alpha \circ \gamma)\}\Big|_{t=t_0} \\
&= \sum_{i=1}^{m} D_{x_i}(f \circ x_\alpha^{-1})(x_\alpha \circ \gamma(t_0)) \cdot \frac{dx_i}{dt}(t_0) \qquad (4) \\
&= \sum_{i=1}^{m} \Big(D_{x_i}(f \circ x_\alpha^{-1})\Big) \circ x_\alpha(p) \cdot \frac{dx_i}{dt}(t_0)
\end{aligned}$$

Notation. Here, as in earlier chapters, we use "bracket notation" $\langle \ell, v \rangle$ to indicate the result $\ell(v)$ when we combine a vector $v \in V$ with a dual vector $\ell \in V^*$. (This turns out to be a good idea.) $\square$

1.8. Definition. *Any linear map* $\ell : \mathcal{C}^\infty(p) \to \mathbb{R}$ *with the property*

$$\langle \ell, f \cdot h \rangle = \langle \ell, f \rangle \cdot h(p) + f(p) \cdot \langle \ell, h \rangle$$

is referred to as a **derivation** *on the local algebra. In particular, if* $\gamma : [a, b] \to M$ *is a smooth curve passing through* $p \in M$ *when* $t = t_0$, *the directional derivative* $\gamma'(t_0) : \mathcal{C}^\infty(p) \to \mathbb{R}$ *along* γ *is a derivation on* $\mathcal{C}^\infty(p)$.

It can be shown that *all* derivations on the local algebra $\mathcal{C}^\infty(p)$ arise in this manner (see Appendix XI-A for the technical details). Given this, we arrive at the modern definition of tangent vector and tangent space for a differentiable manifold M.

Tangent Vectors and the Tangent Space TM_p of a Manifold.

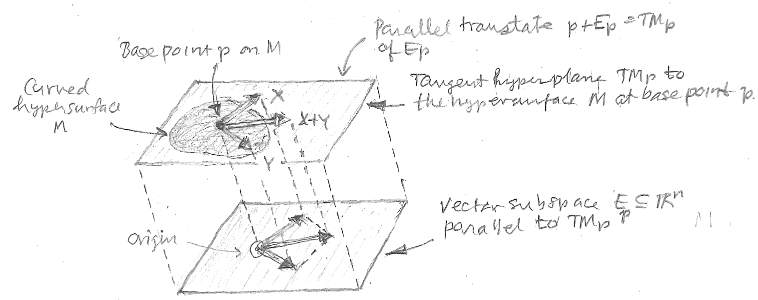

Figure 11.4. The tangent space TM_p to a smooth m-dimensional hypersurface $M \subseteq \mathbb{R}^n$ is a translate $p + E$ of some m-dimensional vector subspace $E \subseteq \mathbb{R}^n$. Elements $\mathbf{v} = p + \mathbf{a}$, $\mathbf{w} = p + \mathbf{b}$ in TM_p are indicated by arrows attached to the base point p. Scaling and vector addition in TM_p is accomplished by transferring arrows back to the corresponding vectors in E, which can be added via the usual parallelogram law for vector addition. This makes TM_p an m-dimensional vector space, see Figure 11.4.

Classically, manifolds were taken to be smooth hypersurfaces M embedded in a larger Euclidean space $\mathbb{R}^n$. The tangent space TM_p at a base point $p \in M$ consisted of "tangent vectors," arrows attached to p pointing out into the surrounding space. TM_p was then viewed as a translate $p + E \subseteq \mathrm{TM}_p$ of some m-dimensional vector subspace E; addition of vectors in TM_p was accomplished by translating arrows $p + \mathbf{a}$, $p + \mathbf{b}$ back to the corresponding vectors $\mathbf{a}, \mathbf{b}$ in E, where they could be added or scaled. The result was then translated back to an arrow $p + (\mathbf{a} + \mathbf{b})$ or $p + (\lambda\mathbf{a})$ attached to p.

Some Objections:

1. This makes no sense if we regard M as our entire universe of discourse and are forbidden to refer to objects lying in some "all encompassing Euclidean space."

2. A more intrinsic approach (still regarding M as an embedded hypersurface in $\mathbb{R}^n$) is to note that classical velocity vector for a moving point whose position is $\gamma(t)$ at time t, is given by
$$\frac{d\gamma}{dt}(t_0) = \lim_{\Delta t \to 0} \frac{\gamma(t_0 + \Delta t) - \gamma(t_0)}{\Delta t} \ .$$
(derivative taken in $\mathbb{R}^n$). When this velocity vector is viewed as an arrow attached to $p = \gamma(t_0)$ in M, it always lies in the classical tangent space to M at p because $\gamma(t)$ remains within M at all times. If we interpret tangent vectors at p as velocity vectors for various curves passing through p, this defines tangent vectors in terms of objects intrinsic to M – smooth parametrized curves in M – but there are still problems. The vector derivative $d\gamma/dt$ is a limit of difference quotients $\Delta\gamma/\Delta t$, which cannot be computed without referring to the surrounding space $\mathbb{R}^n$.

3. The tangent space to M at p is supposed to be a vector space. If smooth curves γ_1, γ_2 in M pass through p when $t = t_0$, it is not obvious how to find a curve $\eta(t)$ such that $\eta(t_0) = p$ and $\eta'(t_0)$ is the sum $\gamma_1'(t_0) + \gamma_2'(t_0)$ of the tangent vectors. There is also an ambiguous relationship between curves through p and tangent vectors at p since different curves through p can determine the same classical tangent vector $d\gamma/dt$.

Another approach is already implicit in the previous comments. We could view tangent vectors $\gamma'(t_0)$ to curves in terms of "directional derivatives" of functions $f : M \to \mathbb{R}$ defined near $p = \gamma(t_0)$, obtained by computing the time derivative of the f-values seen by the moving point $\gamma(t)$. From this point of view $\gamma'(t_0)$ becomes a *linear operator*

$$\langle \gamma'(t_0), f \rangle = \frac{d}{dt}\{f(\gamma(t))\}\Big|_{t=t_0}$$

acting on elements f in the local algebra $\mathcal{C}^\infty(p)$ defined in (3). The objects $f, \gamma(t)$ and $\gamma'(t_0)$ are all defined on M without reference to some mythical "surrounding Euclidean space."

1.9. Definition. *If M is a smooth manifold and $p \in M$ a base point, the* **tangent vectors** *at p are regarded as the derivations on the local algebra $\mathcal{C}^\infty(p)$. The* **tangent space** TM_p *at p is the set of all such derivations. It is a vectors space because a sum of derivations on $\mathcal{C}^\infty(p)$ is again a derivation.*

The following properties hold for directional derivatives along curves in M and their associated derivations on the local algebra $\mathcal{C}^\infty(p)$.

1. Directional derivatives $\gamma'(t_0)$ along smooth curves through p are linear maps from $\mathcal{C}^\infty(p)$ to $\mathbb{R}$ – i.e. they are *linear functionals* in the dual space V^* of the ∞-dimensional vector space $V = \mathcal{C}^\infty(p)$. The classical product formula of Calculus

$$\frac{d}{dt}\{(f \cdot h)(t)\} = \frac{df}{dt} \cdot h(t) + f(t) \cdot \frac{dh}{dt}$$

shows that they are also derivations on $\mathcal{C}^\infty(p)$ because we have $(f \cdot h)(\gamma(t)) = \big(f \circ \gamma(t)\big) \cdot \big(h \circ \gamma(t)\big)$ for all t, and

$$\begin{aligned} \langle \gamma'(t_0), f \cdot h \rangle &= \frac{d}{dt}\{\big(f \circ \gamma(t)\big) \cdot \big(h \circ \gamma(t)\big)\}\Big|_{t=t_0} \\ &= \frac{d}{dt}\{f \circ \gamma\}\Big|_{t=t_0} \cdot h(p) + f(p) \cdot \frac{d}{dt}\{h \circ \gamma(t)\}\Big|_{t=t_0} \\ &= \langle \gamma'(t_0), f \rangle \cdot h(p) + f(p) \cdot \langle \gamma'(t_0), h \rangle \quad \square \end{aligned}$$

for all $f, h \in \mathcal{C}^\infty(p)$.

2. Sums $\ell_1 + \ell_2$ and scalar multiples $\lambda \cdot \ell$ are again derivations on $\mathcal{C}^\infty(p)$, so the tangent space TM_p is a vector space over $\mathbb{R}$.

3. The directional derivative $\gamma'(t_0)$ is a *local operator* on $\mathcal{C}^\infty(p)$: the outcome $\langle \gamma'(t_0), f \rangle$ depends only on the behavior of f near $p = \gamma(t_0)$. In particular, if $f \equiv h$ on some open neighborhood of p then $\langle \gamma'(t_0), f \rangle = \langle \gamma'(t_0), h \rangle$, and if $f \equiv 0$ near p we get $\langle \gamma'(t_0), f \rangle = 0$.

Which curves passing through p determine the same operation on $\mathcal{C}^\infty(p)$? If we describe curves $\gamma(t), \eta(t)$ in local coordinates, say with

$$\mathbf{x}(t) = x_\alpha \circ \gamma(t) = \big(x_1(t), \cdots, x_m(t)\big) \quad \text{and} \quad \mathbf{y}(t) = x_\alpha \circ \eta(t) = \big(y_1(t), \cdots, y_m(t)\big)$$

then $\gamma'(t_0) = \eta'(t_0)$ as operations on $\mathcal{C}^\infty(p)$ if and only if the "first order terms" agree, so that

$$\mathbf{x}(t_0) = \mathbf{y}(t_0) \quad \text{and} \quad \frac{dx_1}{dt}(t_0) = \frac{dy_1}{dt}(t_0), \ldots, \frac{dx_m}{dt}(t_0) = \frac{dy_m}{dt}(t_0)\ .$$

This is clear from equation (4), from which it also follows that the higher order derivatives of $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are irrelevant.

These definitions are framed in a way that does not favor any single system of local coordinates near p over any other system in the maximal atlas. In this sense, all definitions so far are coordinate-independent.

Differential Operators $\partial/\partial x_i$ Determined by a Chart on M.
Given a chart (x_α, U_α) on M we can define derivations on $\mathcal{C}^\infty(p)$ that correspond to the familiar partial derivatives of Calculus. If p in M corresponds to a point $\mathbf{p} = x_\alpha(p) = (p_1, \dots, p_m) \in \mathbb{R}^m$ under the chart map x_α, we can define parametrized straight lines $\tilde{\gamma}_i(t)$ in $\mathbb{R}^m$ such that $\tilde{\gamma}_i(0) = \mathbf{p}$.

$$\tilde{\gamma}_i(t) = \mathbf{p} + t\mathbf{e}_i = (p_1, \cdots, p_i + t, \cdots, p_m) = \big(x_1(t), \cdots, x_m(t)\big) \qquad 1 \le i \le m \ ,$$

where $\mathbf{e}_1, \dots, \mathbf{e}_m$ are the standard basis vectors in $\mathbb{R}^m$. These curves in coordinate space can be transferred to $\mathcal{C}^\infty$ curves $\gamma_i(t) = x_\alpha^{-1} \circ \tilde{\gamma}_i(t)$ in the manifold that pass through p in M when $t = 0$. Directional derivatives along these curves are derivations $\gamma_i'(0)$ on $\mathcal{C}^\infty(p)$, which we denote by

$$\Big(\frac{\partial}{\partial x_i}\Big|_p\Big) : \mathcal{C}^\infty(p) \rightarrow \mathbb{R} \qquad \text{for } 1 \le i \le m \ .$$

These operator act on functions that live on the manifold M, while the classical partial derivatives D_{x_i} act on functions that live on coordinate space $\mathbb{R}^m$. The effect of $\big(\partial/\partial x_i|_p\big)$ on an element $f \in \mathcal{C}^\infty(p)$ is indicated using "bracket notation," by writing

$$\frac{\partial f}{\partial x_i}(p) = \Big\langle \frac{\partial}{\partial x_i}\Big|_p \,,\, f \Big\rangle \qquad \text{for } f \in \mathcal{C}^\infty(U_\alpha) \text{ and } 1 \le i \le m. \tag{5}$$

Note carefully: a chart (x_α, U_α) determines correlated tangent vectors $X_p = \big(\partial/\partial x_i|_p\big)$ simultaneously at *every* base point in U_α. The result is a "field of tangent vectors" on the chart domain U_α.

If we now let the base point p vary within M we obtain new functions $\partial f/\partial x_i(u) = \langle \gamma_i'(0), f\rangle$ defined throughout the chart domain U_α. A simplified version of (4) tells us how to compute these partial derivatives $\partial f/\partial x_i$ on the manifold.

1.10. Lemma. *If (x_α, U_α) is a chart on M and $f \in \mathcal{C}^\infty(U_\alpha)$, the partial derivatives $\partial f/\partial x_i$ on the manifold are given by*

$$\frac{\partial f}{\partial x_i}(p) = \big(D_{x_i}(f \circ x_\alpha^{-1})\big) \circ x_\alpha(p) \qquad 1 \le i \le n$$

for all $p \in U_\alpha$. Thus we get $\partial f/\partial x_i$ on a chart in M in three steps:

- *Transfer $f(u)$ in $\mathcal{C}^\infty(M)$ to a $\mathcal{C}^\infty$ function $(f \circ x_\alpha^{-1})(\mathbf{x})$ on $\mathbb{R}^m$.*
- *Take the classical partial derivative $D_{x_i}(f \circ x_\alpha^{-1})(\mathbf{x})$ on $\mathbb{R}^m$.*
- *Bring the result back to M to get*

$$\frac{\partial f}{\partial x_i}(u) = \big(D_{x_i}(f \circ x_\alpha^{-1})\big)(x_\alpha(u))$$

  for $u \in U_\alpha$.

Proof of 1.10: Write $\mathbf{p} = x_\alpha(p)$ and $\mathbf{x}(t) = x_\alpha \circ \gamma_i(t) = \big(x_1(t), \dots, x_m(t)\big)$. If we factor $f \circ \gamma_i = (f \circ x_\alpha^{-1}) \circ (x_\alpha \circ \gamma_i)$ we get maps $\mathbb{R} \rightarrow \mathbb{R}^m \rightarrow \mathbb{R}$, to which we may apply the classical chain rule to get

$$\begin{aligned}
\frac{\partial f}{\partial x_i}(p) &= \Big\langle \frac{\partial}{\partial x_i}\Big|_p \,,\, f\Big\rangle = \langle \gamma_i'(0)\,,\, f\rangle \\
&= \frac{d}{dt}\{f\circ\gamma_i(t)\}\Big|_{t=0} \;=\; \frac{d}{dt}\{(f\circ x_\alpha^{-1})\circ\big(x_\alpha(\gamma_i(t))\big)\}\Big|_{t=0} \\
&= \frac{d}{dt}\{(f\circ x_\alpha^{-1})(p_1,\ldots,p_i+t,\ldots,p_m)\}\Big|_{t=0} \\
&= \sum_{j=1}^{m}\Big[\Big(D_{x_j}(f\circ x_\alpha^{-1})\Big)(p_1,\ldots,p_i+t,\ldots,p_m)\cdot\frac{dx_j}{dt}(t)\Big]\Big|_{t=0} \\
&= 0+\ldots+1\cdot D_{x_i}(f\circ x_\alpha^{-1})(\mathbf{p})+\ldots+0 \;=\; D_{x_i}(f\circ x_\alpha^{-1})\circ x_\alpha(p)
\end{aligned}$$

because $x_\alpha\circ\gamma_i(t) = (p_1,\ldots,p_i+t,\ldots,p_m)$ if $\mathbf{p} = x_\alpha(p) = x_\alpha\circ\gamma(0)$. □

We will often invoke the following example.

1.11. Example. If (x_α, U_α) is a chart on a differentiable manifold the chart map can be written as

$$\mathbf{x} = x_\alpha(u) = \big(X_1(u),\ldots,X_m(u)\big)$$

whose components $X_k(u)$ are scalar-valued $\mathcal{C}^\infty$ functions on the chart domain U_α. It is amusing to compute $\partial X_k/\partial x_j$ on U_α as an exercise in understanding how the notation works. We claim that

$$\frac{\partial X_k}{\partial x_j}(u) \equiv \delta_{kj} \qquad \text{for all } u\in U_\alpha \text{ and all } j,k \tag{6}$$

where δ_{kj} is the **Kronecker delta symbol**, equal to 1 if $k=j$ and $=0$ when $j\neq k$.

Discussion: Since

$$\mathbf{x} = (x_1,\ldots,x_m) = x_\alpha(x_\alpha^{-1}(\mathbf{x})) = \big(X_1(x_\alpha^{-1}(\mathbf{x})),\ldots,X_m(x_\alpha^{-1}(\mathbf{x}))\big)$$

we have $X_k\circ x_\alpha^{-1}(\mathbf{x}) \equiv x_k$ for all $\mathbf{x}$ in $V_\alpha = x_\alpha(U_\alpha)\subseteq\mathbb{R}^m$. Applying Lemma 1.10 we get

$$\begin{aligned}
\frac{\partial X_k}{\partial x_j}(u) &= \Big\langle\frac{\partial}{\partial x_j}\Big|_u\,,\,X_k\Big\rangle \;=\; \big(D_{x_j}(X_k\circ x_\alpha^{-1})\big)(x_\alpha(u)) \\
&= D_{x_j}(x_k)\Big|_{\mathbf{x}=x_\alpha(u)} \;=\; \delta_{kj}
\end{aligned}$$

for all $u\in U_\alpha$. □

If (x_α, U_α) is a chart on M and $p\in M$, we can invoke (4) to evaluate the directional derivative of a function $f\in\mathcal{C}^\infty(p)$ along a smooth curve $\gamma(t)$ through p, by passing the problem over to coordinate space where the outcome can be determined using the familiar tools of multivariate Calculus. The final result is a pleasing and easily applied formula involving the functions $\partial f/\partial x_i$ on chart domains in M.

1.12. Corollary. *Let (x_α, U_α) be a chart on M, p a base point in U_α, and $f\in\mathcal{C}^\infty(p)$. If $\gamma(t)$ is a $\mathcal{C}^\infty$ curve such that $\gamma(0)=p$ whose description in local coordinates has the form*

$$x_\alpha(\gamma(t)) = \big(x_i(t),\ldots,x_m(t)\big) \qquad \textit{with } \mathcal{C}^\infty \textit{ coefficients } x_i(t)\ ,$$

then the tangent vector $X_p = \gamma'(0)$ determined by differentiating along $\gamma(t)$ is given by

$$\begin{aligned}
\langle X_p, f\rangle &= \langle\gamma'(0)\,,\,f\rangle \;=\; \sum_{j=1}^{m}\frac{dx_j}{dt}(0)\cdot\frac{\partial f}{\partial x_j}(p) \\
&= \Big\langle\sum_{j=1}^{m}\frac{dx_j}{dt}(0)\cdot\Big(\frac{\partial}{\partial x_j}\Big|_p\Big)\,,\,f\Big\rangle \qquad (7)
\end{aligned}$$

for all $f \in \mathcal{C}^\infty(p)$.

Proof: After factoring $f \circ \gamma = (f \circ x_\alpha^{-1}) \circ (x_\alpha \circ \gamma)$, apply the chain rule to get

$$\begin{aligned}\langle \gamma'(t), f\rangle &= \frac{d}{dt}\{f\circ\gamma(t)\} \;=\; \frac{d}{dt}\{f\circ x_\alpha^{-1}\big(x_1(t),\ldots,x_m(t)\big)\} \\ &= \sum_{j=1}^{m} \frac{dx_j}{dt}(t)\cdot\frac{\partial f}{\partial x_j}(\gamma(t))\end{aligned}$$

Now set $t = 0$. $\square$

1.13. Corollary. *If* (x_α, U_α) *is a chart on* M *the vectors*

$$\mathfrak{X} = \{\big(\frac{\partial}{\partial x_1}\Big|_p\big),\ldots,\big(\frac{\partial}{\partial x_m}\Big|_p\big)\}$$

are a basis for the tangent space TM_p *for every* $p \in U_\alpha$. *In particular* $\dim(\mathrm{TM}_p) = m$, *so the space of derivations on* $\mathcal{C}^\infty(p)$ *is finite dimensional even though the local algebra* $\mathcal{C}^\infty(p)$ *itself is infinite dimensional.*

Proof: As we will show in Appendix XI-A, every derivation $D : \mathcal{C}^\infty(p) \to \mathbb{R}$ is a directional derivative $\gamma'(0)$ along some (not necessarily unique) smooth curve such that $\gamma(0) = p$. The identity (7) shows that the vectors in $\mathfrak{X}$ span TM_p. They are also independent, for if there are scalars c_i such that $0 = \sum_{j=1}^m c_i(\partial/\partial x_i|_u)$ for all $u \in U_\alpha$, we may apply this functional on $\mathcal{C}^\infty(u)$ to each of the scalar components $X_k(u)$ of the chart map $x_\alpha(u) = \big(X_1(u),\ldots,X_m(u)\big)$ to get

$$0 = \sum_{i=1}^{m} c_i \Big\langle \frac{\partial}{\partial x_i}\Big|_u,\, X_k\Big\rangle \;=\; \sum_{i=1}^{m} c_i\cdot\delta_{ik} \;=\; c_k \quad \text{ for } k = 1, 2\ldots, m\ .$$

Thus $\mathfrak{X}$ is a basis for TM_p. $\square$

Change of Coordinates. A chart (x_α, U_α) determines basis vectors $\big(\partial/\partial x_i\big|_u\big)$ and partial derivatives $\partial f/\partial x_i(u)$ at each point u in the chart domain U_α. A different chart (y_β, U_β) containing p will assign other basis vectors $\big(\partial/\partial y_i\big|_u\big)$ and partial derivatives $\partial/\partial y_i(u)$ at points where the domains overlap. We often need to pass descriptions of these constructs from one coordinate system to another.

Given overlapping charts (x_α, U_α) and (y_β, U_β) on M, the scalar components in $x_\alpha = (X_1,\ldots,X_m)$ and $y_\beta = (Y_1,\ldots,Y_m)$ are $\mathcal{C}^\infty$ functions on the chart domains. For future reference we determine the **coordinate transition maps** $x_\alpha \circ y_\beta^{-1}$ and $y_\beta \circ x_\alpha^{-1}$, which are defined on the open sets in $\mathbb{R}^m$ that correspond to $U = U_\alpha \cap U_\beta$ in M under the chart maps x_α and y_β. A point $\mathbf{x} \in x_\alpha(U)$ maps to

$$\mathbf{y} \;=\; (y_1,\ldots,y_n) \;=\; y_\beta \circ x_\alpha^{-1}(\mathbf{x}) \;=\; \big(Y_1\circ x_\alpha^{-1}(\mathbf{x}),\ldots,Y_m\circ x_\alpha^{-1}(\mathbf{x})\,\big)$$

in $y_\beta(U)$, so $y_j = Y_j \circ x_\alpha(\mathbf{x})$. The vector-valued map $\mathbf{y} = F(\mathbf{x}) = y_\beta \circ x_\alpha^{-1}(\mathbf{x})$ from $\mathbb{R}^m$ to $\mathbb{R}^m$ is $\mathcal{C}^\infty$ and its classical Jacobian matrix is

$$(8) \qquad [DF(\mathbf{x})] = \Big[D_{x_j}(y_i)\Big] \;=\; \Big[D_{x_j}(Y_i\circ x_\alpha^{-1})(\mathbf{x})\Big] \quad \text{ for } \mathbf{x} \in x_\alpha(U)$$

In the reverse direction we have

$$\mathbf{x} = (x_1,\ldots,x_m) = G(\mathbf{y}) = x_\alpha \circ y_\beta^{-1}(\mathbf{y}) = (X_1\circ y_\beta^{-1},\ldots,X_m\circ y_\beta^{-1})$$

and this inverse map has Jacobian matrix

$$(9) \qquad [DG(\mathbf{y})] = \Big[D_{y_j}(x_i)\Big] \;=\; \Big[D_{y_j}(X_i\circ y_\beta^{-1})(|bfx)\Big] \quad \text{ for } \mathbf{y} \in y_\beta(U)$$

The transformation law below lets us write derivatives $\partial f/\partial x_i$ computed with respect to one chart on M in terms of the partial derivatives $\partial f/\partial y_j$ for another chart. We will make heavy use of these laws. Note that the resulting formulas are cast entirely in terms of functions that live *on the manifold* M.

1.14 Theorem (Change-of-Variable Formula). *Let (x_α, U_α) and (y_β, U_β) be overlapping charts on a manifold M. For any $\mathcal{C}^\infty$ function f on M we have*

$$\frac{\partial f}{\partial x_i}(u) = \sum_{j=1}^{m} \frac{\partial f}{\partial y_j}(u) \cdot \frac{\partial Y_j}{\partial x_i}(u) \qquad \text{for all } u \in U_\alpha \cap U_\beta \ , \tag{10}$$

where $Y_j \in \mathcal{C}^\infty(U_\beta)$ are the scalar components of the chart map $y_\beta = (Y_1, \ldots, Y_m)$.

Proof: The proof is a mildly strenuous exercise in applying the classical chain rule (and our definition of the operators $\partial/\partial x_i$ on the manifold). For $u \in U_\alpha \cap U_\beta$ we have

$$\frac{\partial f}{\partial x_i}(u) = \big(D_{x_i}(f \circ x_\alpha^{-1})\big) \circ x_\alpha(u) \ = \ \Big[D_{x_i}\big((f \circ y_\beta^{-1}) \circ (y_\beta \circ x_\alpha^{-1})\big)\Big] \circ x_\alpha(u)$$

This insertion of $y_\beta^{-1} \circ y_\beta$ is a crucial step. It makes the map $f \circ x_\alpha^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}$ a composite of smooth maps $\mathbb{R}^m \xrightarrow{\phi} \mathbb{R}^m \xrightarrow{F} \mathbb{R}$ between Euclidean spaces, with $F = f \circ y_\beta^{-1}$. The classical chain rule says:

$$D_{x_i}\big(F \circ \phi\big)(\mathbf{x}) \ = \ \sum_{j=1}^{m} D_{y_j} F\big(\phi(\mathbf{x})\big) \cdot D_{x_i} y_j(\mathbf{x})$$

where $\mathbf{y} = \phi(\mathbf{x}) = \big(y_1(\mathbf{x}), \ldots, y_m(\mathbf{x})\big)$ maps $\mathbb{R}^m \rightarrow \mathbb{R}^m$. Now set $F = f \circ y_\beta^{-1}$ and $\mathbf{y} = \phi(\mathbf{x}) = y_\beta \circ x_\alpha^{-1}(\mathbf{x}) = \big(y_1(\mathbf{x}), \ldots, y_m(\mathbf{x})\big)$ mapping $\mathbb{R}^m \rightarrow \mathbb{R}^m$ to get

$$\begin{aligned}
\frac{\partial f}{\partial x_i}(u) \ &= \ \sum_{j=1}^{m} \Big[\big(D_{y_j}(f \circ y_\beta^{-1})\big) \circ (y_\beta \circ x_\alpha^{-1})\Big]\big(x_\alpha(u)\big) \cdot \Big[D_{x_i}\big((y_\beta \circ x_\alpha^{-1})_j\big)\Big]\big(x_\alpha(u)\big) \\
&= \ \sum_{j=1}^{m} \Big[\Big(\big(D_{y_j}(f \circ y_\beta^{-1})\big) \circ y_\beta\Big)(u)\Big] \cdot \Big[D_{x_i}\big(Y_j \circ x_\alpha^{-1}\big)\Big]\big(x_\alpha(u)\big) \\
&= \ \sum_{j=1}^{m} \frac{\partial f}{\partial y_j}(u) \cdot \frac{\partial Y_j}{\partial x_i}(u) \quad \text{(definition of } \partial f/\partial y_j \text{ and } \partial Y_j/\partial x_i\text{)} \qquad \square
\end{aligned}$$

1.15 Corollary (Change of Variable Formula for Operators). *The operators $\partial/\partial x_i$ and $\partial/\partial y_j$ determined by overlapping charts (x_α, U_α) and (y_β, U_β) on a manifold M transform in the following way*

$$\frac{\partial}{\partial x_i}\Big|_u = \sum_{j=1}^{m} \frac{\partial Y_j}{\partial x_i}(u) \cdot \frac{\partial}{\partial y_j}\Big|_u \qquad \text{for all } 1 \leq i, j \leq m \text{ and } u \in U_\alpha \cap U_\beta \ . \tag{11}$$

as linear operators on functions in $\mathcal{C}^\infty(U_\alpha \cap U_\beta)$.

These "change of variable" formulas are easy to remember: the Y_j and y_j should "cancel" when the formula is written correctly, leaving only terms that involve $\partial/\partial x_i$.

Given overlapping charts $(x_\alpha, U_\alpha), (y_\beta, U_\beta)$ on M, entries in the Jacobian matrices (8) and (9) are $\mathcal{C}^\infty$ scalar-valued functions on open subsets of coordinate space $\mathbb{R}^m$,

$$\begin{aligned}
\Big[\frac{\partial y_i}{\partial x_j}(\mathbf{x})\Big] \ &= \ \Big[D_{x_j}(Y_i \circ x_\alpha^{-1})\Big] \quad \text{at points } \mathbf{x} \in x_\alpha(U_\alpha) \subseteq \mathbb{R}^m \\
\Big[\frac{\partial x_i}{\partial y_j}(\mathbf{y})\Big] \ &= \ \Big[D_{y_j}(X_i \circ y_\beta^{-1})\Big] \quad \text{at points } \mathbf{y} \in y_\beta(U_\beta) \subseteq \mathbb{R}^m
\end{aligned}$$

These matrices are inverses of each other when evaluated at base points $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^m$ that correspond under the transition maps. However, the relation between Jacobian matrices becomes clearer if we move these matrix-valued functions from coordinate space $\mathbb{R}^m$ to the manifold M itself. The resulting matrix-valued functions *defined on M*

$$\begin{aligned}
\Big[\frac{\partial Y_i}{\partial x_j}\Big] \;&=\; \Big[\frac{\partial y_i}{\partial x_j} \circ x_\alpha\Big] = \Big[\frac{\partial (Y_i \circ x_\alpha^{-1})}{\partial x_j} \circ x_\alpha\Big] \quad \text{(by Lemma 1.10)} \\
\Big[\frac{\partial X_i}{\partial y_j}\Big] \;&=\; \Big[\frac{\partial x_i}{\partial y_j} \circ y_\beta\Big] = \Big[\frac{\partial (X_i \circ y_\beta^{-1})}{\partial y_j} \circ y_\beta\Big] \quad \text{(Lemma 1.10 again)}
\end{aligned}$$

are then inverses of each other at every base point in $U_\alpha \cap U_\beta$.

1.16 Exercise. Prove that the transferred Jacobian matrices on M are inverses of each other at every base point in $U_\alpha \cap U_\beta$, so

$$\text{(12)} \qquad \Big[\frac{\partial Y_i}{\partial x_j}(u)\Big] \cdot \Big[\frac{\partial X_i}{\partial y_j}(u)\Big] \;=\; I_{m\times m} \quad \text{for all } u \in U_\alpha \cap U_\beta$$

Hint: By 1.11, $\partial Y_i/\partial y_j = \delta_{ij}$; rewrite $\partial/\partial y_j$ in terms of the $\partial/\partial x_k$ as in (11). $\square$

Vector Fields as Differential Operators on M.
On M, or any open subset thereof, a **vector field** $\tilde{X}$ is a map that assigns a tangent vector $X_p \in \mathrm{TM}_p$ at each $p \in M$. By Corollary 1.13, on any chart domain U_α there are uniquely determined coefficients $c_i(u)$ such that

$$\tilde{X}_u = \sum_{i=1}^{m} c_i(u) \cdot \Big(\frac{\partial}{\partial x_i}\Big|_u\Big) \quad \text{for } u \in U_\alpha .$$

These coefficients will change if we pass to another coordinate chart, but smoothness of the coefficients is always preserved.

1.17 Lemma. *If $X_u = \sum c_i(u) \cdot \big(\frac{\partial}{\partial x_i}\big|_u\big)$ for u near base point $p \in U_\alpha$ when described in local coordinates (x_α, U_α), then in any other chart (y_β, U_β) containing p we have*

$$X_u = \sum_{j=1}^{m} d_j(u) \cdot \Big(\frac{\partial}{\partial y_j}\Big|_u\Big)$$

with coefficients

$$d_j(u) = \sum_{k=1}^{m} c_k(u) \cdot \frac{\partial Y_j}{\partial x_k}(u) \ ,$$

where the Y_j are the scalar components of $y_\beta = \big(Y_1(u), \ldots, Y_m(u)\big)$.

Proof: This is immediate from equation (11). $\square$

A **smooth vector field** $\tilde{X}$ is a field $p \mapsto X_p$ of tangent vectors on M whose description in local coordinates has $\mathcal{C}^\infty$ coefficients for every chart. It follows from Lemma 1.17 that smoothness of $\tilde{X}$ on M has a coordinate-independent meaning because the scalar components $Y_j(u)$ of y_β and their derivatives $\partial Y_j/\partial x_k$ are smooth functions on M.

1.18 Corollary. *If X is a vector field on M whose description in local coordinates is $X_u = \sum c_i(u) \cdot \big(\frac{\partial}{\partial x_i}\big|_u\big)$ with coefficients that are $\mathcal{C}^\infty$ near p, then the same will be true for any other chart (y_β, U_β) about p.*

The set of smooth vector fields on M is denoted $\mathcal{D}^{(1,0)}(M)$. It becomes a vector space over $\mathbb{R}$ if we define

$$(X+Y)_p = X_p + Y_p \quad \text{and} \quad (\lambda X)_p = \lambda \cdot X_p \quad \text{for } p \in M \ .$$

This ∞-dimensional space is also a $\mathcal{C}^\infty(M)$**-module** because there is a natural action $\mathcal{C}^\infty(M) \times \mathcal{D}^{(1,0)}(M) \rightarrow \mathcal{D}^{(1,0)}(M)$ defined by pointwise multiplication

$$(f \cdot \tilde{X})_p = f(p) \cdot X_p \quad \text{ for all } p$$

There is also an action $\mathcal{D}^{(1,0)} \times \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$, obtained by letting a vector field $\tilde{X}$ act on functions in the following way

$$\tilde{X}f(u) = \langle X_u, f \rangle \quad \text{ for all } u \in M. \tag{13}$$

When $\tilde{X}f$ is described in local chart coordinates $\tilde{X}$ becomes an operator that acts on smooth functions on coordinate space $\mathbb{R}^m$ like a first order partial differential operator with variable coefficients. In fact, if (x_α, U_α) is a chart on M and we describe $\tilde{X}$ in chart coordinates we get

$$X_u = \sum_i c_i(u) \cdot \Big(\frac{\partial}{\partial x_i}\Big|_u\Big) \quad \text{ with } \quad c_i(u) \in \mathcal{C}^\infty(U_\alpha) \ ,$$

so that

$$\begin{aligned} \tilde{X}f(u) &= \langle X_u, f \rangle = \sum_{i=1}^m c_i(u) \cdot \Big\langle \frac{\partial}{\partial x_i}\Big|_u, f \Big\rangle \\ &= \sum_{i=1}^m c_i(u) \cdot \frac{\partial f}{\partial x_i}(u) \end{aligned}$$

is in $\mathcal{C}^\infty(U_\alpha)$. When smooth vector fields on M are regarded as differential operators on M, we see that vector fields have their own "global derivation" property in addition to being linear operators $\tilde{X} : \mathcal{C}^\infty(M) \rightarrow \mathcal{C}^\infty(M)$; they are linear derivations on $\mathcal{C}^\infty(M)$, with

$$\tilde{X}(f \cdot h)(u) = \tilde{X}f(u) \cdot h(u) + f(u) \cdot \tilde{X}h(u) \quad \text{ for } u \in M \text{ and } f, h \in \mathcal{C}^\infty(M) \ .$$

This follows directly from definition (13).

Another important property of this action involves the *support sets* of functions f in $\mathcal{C}^\infty(M)$,

$$\begin{aligned} \operatorname{supp}(f) &= \text{Closure in } M \text{ of the set } \{u \in M : f(u) \neq 0\} \\ &= \text{Complement in } M \text{ of the open set } \{u \in M : f(u) = 0\} \ , \end{aligned}$$

1.19 Proposition (Reduction of Supports). *The action of a smooth vector field $\tilde{X}$ can only decrease the support set* $\operatorname{supp}(f)$ *of a function $f \in \mathcal{C}^\infty(M)$, or equivalently*

$$f \equiv 0 \textit{ on an open set } U \subseteq M \ \Rightarrow \ \tilde{X}(f) \equiv 0 \textit{ on } U \ .$$

Proof: If $f \equiv 0$ near p, write $\tilde{X}$ in local chart coordinates to see $\tilde{X}f \equiv 0$ near p. $\square$

By looking at differences $f - h$ we get

$$f \equiv h \text{ on an open set } U \subseteq M \text{ implies } \tilde{X}\, f \equiv \tilde{X}\, h \text{ on } U, \text{ for all } f, h \in \mathcal{C}^\infty(M).$$

This means that smooth vector fields $\tilde{X}$ are *local operators* on $\mathcal{C}^\infty(M)$: the value of $\tilde{X}f$ at a point p is determined only by the behavior of f in the immediate vicinity of p.

Action of Smooth Maps $\phi : M \rightarrow N$ on Tangent Vectors.
If $\phi : M \rightarrow N$ is a $\mathcal{C}^\infty$ map, $p \in M$, and $q = \phi(p)$ in N, then under the action of ϕ,

- Points $p \in M$ get moved forward to points $q = \phi(p)$ in N.
- Functions $f \in \mathcal{C}^\infty(N)$ get "pulled back" to functions $\phi^{\mathrm{t}} f = f \circ \phi$ in $\mathcal{C}^\infty(M)$.
- Tangent vectors X_p in TM_p get "pushed forward" to vectors $(d\phi)_p(X_p)$ in TN_q by the **differential of** ϕ at p, the linear map $(d\phi)_p : \mathrm{TM}_p \to \mathrm{TM}_q$ defined as follows.

For $X_p \in \mathrm{TM}_p$ and $f \in \mathcal{C}_N^\infty(\phi(p))$, we define the tangent vector $(d\phi)_p X_p$ in $\mathrm{TN}_{\phi(p)}$ to be the unique tangent vector at $q = \phi(p)$ such that

$$\langle\, (d\phi)_p X_p\, ,\, f\,\rangle = \langle X_p, (d\phi)_p^{\mathrm{t}} f\,\rangle = \langle X_p, f \circ \phi\rangle\ , \tag{14}$$

for all $\mathcal{C}^\infty$ functions f defined on N near $\phi(p)$.

It is obvious that (14) determines a derivation $Y_q = (d\phi)_p X_p$ on $\mathcal{C}_N^\infty(\phi(p))$, so it is in TN_q; it is also obvious that $(d\phi)_p$ is a linear map from $\mathrm{TM}_p \to \mathrm{TN}_{\phi(p)}$. None of these definitions refer to coordinates on M or N.

To calculate the effect of $(d\phi)_p$ we impose local charts (x_α, U_α) about p and (y_β, U_β) about $q = \phi(p)$ which determine bases $\mathfrak{X} = \{\big(\partial/\partial x_i|_p\big)\}$ and $\mathfrak{Y} = \{\big(\partial/\partial y_j|_q\big)\}$ in TM_p and TN_q; the action of the linear operator $(d\phi)_p : \mathrm{TM}_p \to \mathrm{TN}_{\phi(p)}$ is completely determined once we know what it does to the basis vectors. This action is described by a law similar to the Change-of-Coordinates rule in Theorem 1.14.

1.20 Theorem. (Transformation Law for Tangent Vectors). *If $\phi : M \to N$ is a $\mathcal{C}^\infty$ map, $p \in M$, and $q = \phi(p)$, let (x_α, U_α) and (y_β, U_β) be charts about p and q respectively. If the scalar components of the chart map y_β on N are given by $y_\beta(v) = \big(Y_1(v) \ldots, Y_n(v)\big)$ for $v \in U_\beta$, then*

$$(d\phi)_p\Big(\frac{\partial}{\partial x_i}\Big|_p\Big) = \sum_{j=1}^n \frac{\partial(Y_j \circ \phi)}{\partial x_i}(p) \cdot \Big(\frac{\partial}{\partial y_j}\Big|_q\Big) \qquad \text{for } 1 \le i \le m. \tag{15}$$

Proof: On any chart (y_β, U_β) in N the scalar components of y_β have the property $\partial Y_i/\partial y_j \equiv \delta_{ij}$ (Kronecker delta) throughout the chart domain U_β. For any basis vector $\big(\partial/\partial x_i|_p\big)$ in TM_p, its image in $\mathrm{TN}_{\phi(p)}$ can be written

$$(d\phi)_p\Big(\frac{\partial}{\partial x_i}\Big|_p\Big) \;=\; \sum_{j=1}^n c_j(q) \cdot \Big(\frac{\partial}{\partial y_j}\Big|_q\Big) \qquad (q = \phi(p))$$

By (6) the coefficients $c_j(q)$ can be recovered by bracketing with the scalar components Y_k of the chart map y_β, so

$$\begin{aligned} c_j(q) \;&=\; \Big\langle\, (d\phi)_p\Big(\frac{\partial}{\partial x_i}\Big|_p\Big)\, ,\, Y_j \Big\rangle \\ &=\; \Big\langle\, \Big(\frac{\partial}{\partial x_i}\Big|_p\Big)\, ,\, Y_j \circ \phi \Big\rangle \qquad \text{(definition (14) of } (d\phi)_p) \\ &=\; \frac{\partial(Y_i \circ \phi)}{\partial x_i}(p) \end{aligned}$$

That proves (15). $\square$

Note the formal cancellation of "y_j" and "Y_j" when the formula is written correctly.

1.21 Exercise. (Composition of Maps). Let $M \overset{\psi}{\to} N \overset{\phi}{\to} Q$ be $\mathcal{C}^\infty$ maps between

manifolds. If $p \in M$ and $q = \psi(p)$ in N, explain why the composite $\phi \circ \psi : M \to Q$ is a $\mathcal{C}^\infty$ map, and prove that
$$d(\phi \circ \psi)_p = (d\phi)_{\psi(p)} \circ (d\psi)_p$$
as maps $\mathrm{TM}_p \to \mathrm{TN}_{\psi(p)} \to \mathrm{TQ}_{\phi(\psi(p))}$ $\square$.

The next example shows how to calculate the differential $(d\phi)_p : \mathrm{TM}_p \to \mathrm{TN}_{\phi(p)}$ of a smooth mapping between manifolds.

1.22 Example. The complex variable "squaring map" $w = \phi(z) = z^2$ takes the form
$$u + iv = \phi(x, y) = (x + iy)^2 = (x^2 - y^2) + i(2xy) = (x^2 - y^2, 2xy)$$
in Cartesian coordinates when we write $z = x + iy$ and $w = u + iv$. This is a map $\phi : M \to N$ between two copies of the complex plane $\mathbb{C}$, each of which can be regarded as a $\mathcal{C}^\infty$ manifold whose structure is determined by a single global chart:

- For M take (x_α, U_α) with $U_\alpha = M$ and $(x, y) = x_\alpha(z) = (X(z), Y(z)) \in \mathbb{R}^2$, where $X(z) = x$ and $Y(z) = y$ if $z = x + iy$.
- For N take (y_β, U_β) with $U_\beta = N$ and $(u, v) = y_\beta(w) = (U(w), V(w)) \in \mathbb{R}^2$, where $U(w) = u$ and $V(w) = v$ if $w = u + iv$.

At a typical point $z = x+iy$ in M the identity $(u, v) = y_\beta \circ \phi \circ x_\alpha^{-1}(x, y)$ forces the scalar components of the chart maps $x_\alpha = (X, Y)$ and $y_\beta = (U, V)$ to satisfy the relations
$$\begin{aligned} u &= (U \circ \phi)(z) = U(\phi(z)) = X^2(z) - Y^2(z) = x^2 - y^2 \\ v &= (V \circ \phi)(z) = V(\phi(z)) = 2X(z)Y(z) = 2xy \end{aligned}$$
where $w = u + iv = z^2$. The players involved are shown in the following diagram.

$$\begin{array}{ccccc} x + iy = z & \mathbb{C} & \xrightarrow{\quad\phi\quad} & \mathbb{C} & w = u + iv = z^2 \\ \downarrow & x_\alpha \downarrow & & \downarrow y_\beta & \downarrow \\ x_\alpha(z) = \big(X(z), Y(z)\big) = (x, y) & \mathbb{R}^2 & \xrightarrow{\Phi = y_\beta \circ \phi \circ x_\alpha^{-1}} & \mathbb{R}^2 & (u, v) = \big(U(w), V(w)\big) = \big(x^2 - y^2, 2xy\big) \end{array}$$

If $p \in M$ and $q = \phi(p) \in N$ the charts x_α, y_β determine basis vectors $(\partial/\partial x|_p), (\partial/\partial y|_p)$ and $(\partial/\partial u|_q), (\partial/\partial v|_q)$ in the tangent spaces $\mathrm{TM}_p, \mathrm{TN}_q$, which have dimension $= 2$ over $\mathbb{R}$. To compute $(d\phi)_p$ we describe ϕ in these local coordinates. The $\mathcal{C}^\infty$ map $\Phi = y_\beta \circ \phi \circ x_\alpha^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ between coordinate spaces is
$$(u, v) = \Phi(x, y) = y_\beta \circ \phi \circ x_\alpha^{-1}(x, y) = y_\beta(\phi(x + iy)) = \big(x^2 - y^2\ ,\ 2xy\big)$$
and its Jacobian matrix is
$$[D\Phi(x, y)] = \frac{\partial(u, v)}{\partial(x, y)} = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix}$$
If $p = x+iy$ in M, so $x_\alpha(p) = (X(p), Y(p)) = (x, y)$, basis vectors in TM_p are transformed to vectors in $\mathrm{TN}_{\phi(p)}$ via
$$(d\phi)_p\Big(\frac{\partial}{\partial x}\Big|_p\Big) \;=\; \frac{\partial(U \circ \phi)}{\partial x}(p) \cdot \Big(\frac{\partial}{\partial u}\Big|_{\phi(p)}\Big) + \frac{\partial(V \circ \phi)}{\partial x}(p) \cdot \Big(\frac{\partial}{\partial v}\Big|_{\phi(p)}\Big)$$
as in (15). But on $U_\alpha \subseteq M$ we have $\partial X/\partial x \equiv 1, \partial X/\partial y \equiv 0$, etc (courtesy of Example 1.11), hence
$$\begin{aligned} \frac{\partial(U \circ \phi)}{\partial x} &= \frac{\partial}{\partial x}\big(X^2 - Y^2\big) \;=\; 2X \cdot \frac{\partial X}{\partial x} - 2Y \cdot \frac{\partial Y}{\partial x} \;=\; 2X \\ \frac{\partial(V \circ \phi)}{\partial x} &= \frac{\partial}{\partial x}\big(2XY\big) \;=\; 2Y \cdot \frac{\partial X}{\partial x} + 2X \cdot \frac{\partial Y}{\partial x} \;=\; 2Y\ , \end{aligned}$$

so at $p = x + iy$ in M we have

$$
\begin{aligned}
(d\phi)_p\Big(\frac{\partial}{\partial x}\Big|_p\Big) &= 2X(p)\cdot\Big(\frac{\partial}{\partial u}\Big|_{\phi(p)}\Big) + 2Y(p)\cdot\Big(\frac{\partial}{\partial v}\Big|_{\phi(p)}\Big) \\
&= 2x\cdot\Big(\frac{\partial}{\partial u}\Big|_{\phi(p)}\Big) + 2y\cdot\Big(\frac{\partial}{\partial v}\Big|_{\phi(p)}\Big) \qquad (16)
\end{aligned}
$$

where $\phi(p) = p^2$. As a numerical example, suppose $p = 2 + i$ in M, so $\phi(p) = (2 + i)^2 = 3 + 4i$. By (16) we get

$$
(d\phi)_p\Big(\frac{\partial}{\partial x}\Big|_p\Big) = 4\cdot\Big(\frac{\partial}{\partial u}\Big|_{\phi(p)}\Big) + 2\cdot\Big(\frac{\partial}{\partial v}\Big|_{\phi(p)}\Big)
$$

at the base point $p = 2 + 1$ in M, and a similar calculation yields

$$
\begin{aligned}
(d\phi)_p\Big(\frac{\partial}{\partial y}\Big|_p\Big) &= -2Y(p)\cdot\Big(\frac{\partial}{\partial u}\Big|_{\phi(p)}\Big) + 2X(p)\cdot\Big(\frac{\partial}{\partial v}\Big|_{\phi(p)}\Big) \\
&= -2\cdot\Big(\frac{\partial}{\partial u}\Big|_{\phi(p)}\Big) + 4\cdot\Big(\frac{\partial}{\partial v}\Big|_{\phi(p)}\Big)
\end{aligned}
$$

Given the action on basis vectors, the action of $(d\phi)_p : \mathrm{TM}_p \to \mathrm{TN}_{\phi(p)}$ on arbitrary tangent vectors is determined by linearity. $\Box$

XI.2 Cotangent Vectors and Differential Forms.

As in Chapter IX, the *dual space* of a vector space V over $\mathbb{R}$ is the set V^* of all linear functionals on V, the linear maps $\ell : V \to \mathbb{R}$. If $\dim(V) < \infty$ then $\dim(V^*) = \dim(V)$, and any basis $\mathfrak{X} = \{e_1, \ldots, e_n\}$ in V induces a *dual basis* $\mathfrak{X}^* = \{e_1^*, \ldots, e_n^*\}$ in V^*, determined by the property

$$
\langle\, e_i^*,\, e_j\rangle = \delta_{ij} \quad \text{(Kronecker delta symbol)}
$$

If $v = \sum_{i=1}^n c_i e_i$ in V the dual functional e_k^* reads the k^{th} coefficient c_k of v, so $\langle e_k^*, v\rangle = c_k$, and if $\ell = \sum_{j=1}^n d_j e_j^*$ in V^* its coefficients can be found by bracketing ℓ with the basis vectors e_j to get

$$
d_j = \langle\, \ell, e_j\rangle \quad \text{for } 1 \le j \le n\ .
$$

If M is a differentiable manifold the tangent spaces TM_p and their dual spaces TM_p^* both play crucial roles in differential geometry.

2.1 Definition. *The* **cotangent space** *at $p \in M$ is the dual space TM_p^* to the tangent space TM_p. For reasons that will gradually emerge, this space is often referred to as the space of* **1-forms**, *or* **rank-1 differential forms** *on TM_p, and denoted $\bigwedge^1(\mathrm{TM}_p^*)$. By definition, the field of scalars $\mathbb{R}$ is regarded as the space of* **rank-0 differential forms** *on TM_pM, and as such is denoted by $\bigwedge^0(TM_p^*)$.*

Given a chart (x_α, U_α) on M and a point $p \in M$, we get a basis

$$
\mathfrak{X}_p = \Big\{\Big(\frac{\partial}{\partial x_i}\Big|_p\Big)\Big\} \text{ for } \mathrm{TM}_p \quad \text{and a dual basis} \quad \mathfrak{X}_p^* = \Big\{\Big(\frac{\partial}{\partial x_i}\Big|_p\Big)^*\Big\} \text{ in } \mathrm{TM}_p^*\ .
$$

For various reasons these dual vectors have come to be denoted by other symbols that may seem peculiar at first, but the notation will grow on you as its advantages become apparent. Hereafter we shall write

$$
(dx_i)_p \equiv \Big(\frac{\partial}{\partial x_i}\Big|_p\Big)^* \qquad (17)
$$

so the dual basis determined by a chart is written

$$(dx_1)_p, \ldots, (dx_m)_p \quad \text{instead of the more cumbersome} \quad \Big(\frac{\partial}{\partial x_1}\Big|_p\Big)^*, \ldots, \Big(\frac{\partial}{\partial x_m}\Big|_p\Big)^*$$

By definition of "dual basis" we then have

$$\Big\langle (dx_i)_p \, , \Big(\frac{\partial}{\partial x_j}\Big|_p\Big) \Big\rangle = \delta_{ij} \quad \text{(Kronecker delta)} \tag{18}$$

for $1 \leq i, j \leq m$.

The Canonical d-Operator on $\mathcal{C}^\infty(M)$.
Each $f \in \mathcal{C}^\infty(p)$ determines an element $(df)_p \in \mathrm{TM}_p^*$, the **rank-0 exterior derivative** of f at p. This is given by a construction that makes no mention of local coordinates:

$$(df)_p : \mathrm{TM}_p \to \mathbb{R} \quad \text{is obtained by letting} \quad \langle (df)_p, X_p \rangle = \langle X_p, f \rangle \tag{19}$$

for all $f \in \mathcal{C}^\infty(p)$ and $X_p \in \mathrm{TM}_p$. On any chart domain U_α the outcome can also be described in local coordinate to write $(df)_p$ as a linear combination of the dual basis vectors $(dx_1)_p, \ldots, (dx_m)_p$ determined by the chart (x_α, U_α).

2.2 Proposition. *Given a chart (x_α, U_α) and a function $f \in \mathcal{C}^\infty(U_\alpha)$, the rank-0 exterior derivative $(df)_p \in \mathrm{TM}_p^*$ has the following description in local chart cooordinates at every base point $p \in U_\alpha$.*

$$(df)_p = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(p) \cdot (dx_i)_p \ . \tag{20}$$

Proof: At each $p \in U_\alpha$ there exist unique coefficients such that $(df)_p = \sum_{i=1}^m c_i(u)\,(dx_i)_p$ in TM_p^*. To determine the c_i simply apply $(df)_p$ to the basis vectors $\big(\partial/\partial x_i\big|_p\big)$ in TM_p; by (6) we get

$$\begin{aligned} c_i(p) &= \Big\langle (df)_p, \frac{\partial}{\partial x_i}\Big|_p \Big\rangle = \Big\langle \frac{\partial}{\partial x_i}\Big|_p, f \Big\rangle \quad \text{(definition (19) of } (df)_p) \\ &= \frac{\partial f}{\partial x_i}(p) \quad \text{for } 1 \leq i \leq m \quad \square \end{aligned}$$

Observe that $(df)_p$ has the same general form in all coordinate systems. For example, given Cartesian coordinates $\mathbf{x} = x_\alpha(u) = (x, y)$ and polar coordinates $\mathbf{y} = y_\beta(u) = (r, \theta)$ on an open subset $U \subseteq M = \mathbb{R}^2$, we have

$$(df)_u = \frac{\partial f}{\partial x}(u) \cdot (dx)_u + \frac{\partial f}{\partial y}(u) \cdot (dy)_u \ = \ \frac{\partial f}{\partial r}(u) \cdot (dr)_u + \frac{\partial f}{\partial \theta}(u) \cdot (d\theta)_u$$

for all u where the charts overlap. Note the resemblance between the description of $(df)_p$ in (20) and the classical gradient $\nabla f(p) = \sum_{i=1}^n D_{x_i} f(p)\,\mathbf{e}_i$ of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$. This is not an accident. In the theory of manifolds the classical gradient ∇f becomes the rank-0 exterior derivative $d : \mathcal{C}^\infty(M) \to \bigwedge^1(\mathrm{TM}_p^*)$ and ∇f becomes the 1-form df of (19).

Every cotangent vector $\omega_p \in \mathrm{TM}_p^*$ is equal to $(df)_p$ for some f, but the f is not unique.

2.3 Lemma. *The map $d : \mathcal{C}^\infty(p) \to \mathrm{TM}_p^*$ that sends $f \mapsto (df)_p$ is linear and surjective. but not one-to-one. The nontrivial kernel of this map is*

$$\ker(d) = \{f \in \mathcal{C}^\infty(M) : \frac{\partial f}{\partial x_1}(p) = \ldots = \frac{\partial f}{\partial x_m}(p) = 0\}$$

Thus $(df)_p = 0$ if and only if p is a critical point for $f : M \to \mathbb{R}$ when f is described in local coordinates.

Proof: If we have $\partial f/\partial x_i(p) = 0$ for $1 \leq i \leq m$, for one chart (x_α, U_α) about p then by (10) this must be true for every other chart (y_β, U_β) about p. Thus the property "*p is a critical point for f*" is independent of any choice of local coordinates.

Surjectivity follows for an interesting reason. The scalar component functions $X_i(u)$ of the chart map $x_\alpha(u) = \big(X_1(u), \ldots, X_m(u))\big)$ are in $\mathcal{C}^\infty(U_\alpha)$, and by (6) we get

$$(dX_i)_u = (dx_i)_p \quad \text{for } 1 \leq i \leq m \tag{21}$$

because when we apply $(dX_i)_p \in \mathrm{TM}_P^*$ to a typical basis vector in TM_p equation (20) yields

$$(dX_i)_u = \sum_{j=1}^m \frac{\partial X_i}{\partial x_j}(u) \cdot (dx_j)_u = \sum_{j=1}^m \delta_{ij} \cdot (dx_j)_u = (dx_i)_u \ .$$

Thus $(dX_i)_u$ and the dual vector $(dx_i)_u = \big(\partial/\partial x_i\big|_u\big)^*$ determined by the chart (x_α, U_α) are identical elements in TM_u^* for all $u \in U_\alpha$, despite their very different origins. It is now clear that $range(d)$ contains a basis for TM_p^*, and therefore is surjective. $\square$

Let $\omega : p \to \omega_p \in \mathrm{TM}_p^*$ be a field of 1-forms on M. Given a chart (x_α, U_α) there are unique coefficients $c_j(p)$ such that

$$\omega_p = \sum_{j=1}^m c_j(p) \cdot (dx_j)_p \quad \text{in } \mathrm{TM}_p^* \text{ for } p \in U_\alpha \ . \tag{22}$$

By (18) the coefficients can be recovered by bracketing ω_p with the basis vectors $\{\big(\partial/\partial x_i|_p\big)\}$ in TM_p determined by the chart.

2.4 Definition. *We say that ω is a* **smooth field of 1-forms**, *or a* **smooth rank-1 differential form** *on M if for every chart (x_α, U_α) the coefficient functions $c_k(u)$ are in $\mathcal{C}^\infty(U_\alpha)$. As we will show in Proposition 2.5 below, this definition is coordinate-independent. The set of such fields on M is denoted by $\mathcal{D}^{(0,1)}(M)$, or sometimes by $\bigwedge^1(M)$ depending on the context. This becomes an infinite-dimensional vector space if sums and scalar multiples in $\mathcal{D}^{(0,1)}(M)$ are given by*

$$(\lambda\omega)_p = \lambda \cdot \omega_p \qquad \text{and} \qquad (\omega + \mu)_p = \omega_p + \mu_p \qquad \text{in } \mathrm{TM}_p^* \ .$$

for all $p \in M$, and fields $\omega, \mu \in \mathcal{D}^{(0,1)}(M)$. This space of 1-forms is also a $\mathcal{C}^\infty(M)$-module under the action $\mathcal{C}^\infty(M) \times \mathcal{D}^{(0,1)}(M) \to \mathcal{D}^{(0,1)}(M)$ given by

$$(f \cdot \omega)_p = f(p) \cdot \omega_p \qquad \text{for } p \in M, f \in \mathcal{C}^\infty(M), \omega \in \mathcal{D}^{(0,1)}(M) \ .$$

The following Change-of-Variable law for 1-forms justifies our claim that smoothness is a coordinate-independent property of a field of cotangent vectors. The proof follows easily from the previously established transformation law (11) for tangent vectors.

2.5 Proposition (Change of Variable for Cotangent Vectors). *The cotangent vectors $(dx_i)_u$ and $(dy_j)_u$ determined by charts (x_α, U_α) and (y_β, U_β) on a manifold M transform in the following way*

$$(dx_i)_u = \sum_{j=1}^m \frac{\partial Y_j}{\partial x_i}(u) \cdot (dy_j)_u \qquad 1 \leq i, j \leq m$$

for all base points $u \in U_\alpha \cap U_\beta$. Here the Y_i are the scalar components of the chart map $y_\beta = (Y_1, \ldots, Y_m)$.

Proof: There are unique coefficients such that $(dy_i)_u = \sum_{j=1}^m c_j(u)\cdot(dx_j)_u$. If the scalar components of y_β are $y_\beta(u) = \big(Y_1(u), \ldots, Y_m(u)\big)$, then by (6) and Lemma 2.3, bracketing with the dual basis vectors $\big(\partial/\partial x_k\big|_u\big)$ yields

$$\begin{aligned} c_k(u) &= \Big\langle (dy_i)_u, \frac{\partial}{\partial x_k}\Big|_u \Big\rangle = \Big\langle (dY_i)_u, \frac{\partial}{\partial x_k}\Big|_u \Big\rangle \\ &= \Big\langle \frac{\partial}{\partial x_k}\Big|_u, Y_i \Big\rangle = \frac{\partial Y_i}{\partial x_k}(u) \qquad \text{for } 1 \le i \le m, \end{aligned}$$

proving the formula. $\square$

A law similar to Theorem 1.20 for tangent vectors describes the action by which a $\mathcal{C}^\infty$ map $\phi : M \to N$ "pulls back" cotangent vectors from base points in N to base points in M. We won't pursue this here since it won't be needed in later discussion, but it is a fundamental result in differential geometry – see Appendix XI-B for details.

The Exterior Derivative (an Overview). Equation (19) defines a linear operator $d = d_0 : \bigwedge^0(M) \to \bigwedge^1(M)$ called the **rank-0 exterior derivative**, where we define the space of rank-0 differential forms to be $\bigwedge^0(M) = \mathcal{C}^\infty(M)$. As in (19) we have

$$\langle (df)_p, X_p \rangle = \langle X_p, f \rangle \qquad \text{for all } f \in \mathcal{C}^\infty(p), p \in M \ . \tag{23}$$

The rank-0 d-operator has its own derivation property:

$$\big(d(fh)\big)_u = (df)_u h(u) + f(u)(dh)_u \qquad \text{for all } u \in M \text{ and } f, h \in \mathcal{C}^\infty(M) \ .$$

We will soon define a hierarchy of spaces of smooth tensor fields

$$\textstyle\bigwedge^k(M) = (\text{the rank } k\text{-differential forms on } M)$$

and associated rank-k exterior derivatives $d = d_k : \bigwedge^k(M) \to \bigwedge^{k+1}(M)$ such that

$$\textstyle\bigwedge^0(M) \xrightarrow{d=d_0} \bigwedge^1(M) \xrightarrow{d=d_1} \bigwedge^2(M) \xrightarrow{d=d_2} \cdots \xrightarrow{d=d_{m-1}} \bigwedge^m(M) \xrightarrow{d=d_m} 0 \tag{24}$$

where $m = \dim(M)$. In all dimensions these d_k play the roles occupied by the vector operators *div, grad, curl* in $m = 3$ dimensions, but a few more ideas must be developed before that connection can be explained.

Primitives of 1-forms. If ω is a smooth 1-form on an open set $U \subseteq M$, a **primitive** of ω is an $f \in \mathcal{C}^\infty(U)$ such that $df = \omega$ on U. If f is such a primitive on U, consider a chart (x_α, U_α) about p; replacing $U_\alpha \to U \cap U_\alpha$ we may assume $U_\alpha \subseteq U$. On U_α we can describe ω in local coordinates, $\omega_u = \sum_{i=1}^m w_i(u)\cdot(dx_i)_u$ with $w_i \in \mathcal{C}^\infty(U_\alpha)$. By Proposition 2.2, $df = \omega$ on U_α means that

$$(df)_u = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(u)\,(dx_i)_u \quad \text{is equal to} \quad \omega_u = \sum_{i=1}^m w_i(u)\,(dx_i)_u$$

for all $u \in U_\alpha$. The coefficients must agree, so the $w_i(u)$ are related to the partial derivatives of f via a system of partial differential equations

$$w_i(u) = \frac{\partial f}{\partial x_i}(u) \qquad \text{for all } u \in U_\alpha \text{ and } 1 \le i \le m \ .$$

This system does not always admit solutions, even locally. In fact the coefficients $w_i(u)$ of ω in local coordinates must satisfy the following "consistency condition" if the equation $df = w$ is to have any local solutions within the chart domain.

$$\frac{\partial w_i}{\partial x_j} - \frac{\partial w_j}{\partial x_i} \equiv 0 \qquad \text{on } U_\alpha \text{ for } 1 \le i < j \le m \ . \tag{25}$$

These conditions are *necessary* because if f is of class at least $C^{(2)}$ its mixed 2^{nd} order partial derivatives must agree, so the $w_i(u)$ satisfy a system of $\frac{1}{2}(m^2-m)$ equations in the m unknowns $u_1, \ldots, u_m$.

$$\frac{\partial w_i}{\partial x_j} = \frac{\partial}{\partial x_j}\Big(\frac{\partial f}{\partial x_i}\Big) = \frac{\partial}{\partial x_i}\Big(\frac{\partial f}{\partial x_j}\Big) = \frac{\partial w_j}{\partial x_i} \qquad \text{for } 1 \le i, j \le m$$

We list without proof the following facts about local and global primitives of 1-forms $\omega \in \bigwedge^1(M)$ – see Appendix XI-C for full details.

1. The consistency conditions for ω take the same form (25) in every system of local coordinates. (This is fairly obvious.)

2. Using line integrals $\int_\gamma w$ of 1-forms along $C^{(1)}$ curves one can prove the existence of local solutions near any base point $p \in U_\alpha$ if the conditions (25) are satisfied.

3. Conditions (25) do not by themselves imply the existence of a "global solution" f in $\mathcal{C}^\infty(U)$ to the identity $df = \omega$. The geometry of U can present obstructions to existence of solutions for certain forms ω even when (25) holds, as in Example 2.7 below.

4. If ω is a smooth 1-form on a manifold M and the equation $df = \omega$ has a $\mathcal{C}^\infty$ solution on an open subset $U \subseteq M$ that is *connected*, all other solutions on U are obtained by adding an arbitrary constant to f.

2.6 Exercise. If f is a $\mathcal{C}^\infty$ function defined near $p \in M$, prove that the following properties are equivalent

(a) $\dfrac{\partial f}{\partial x_i} \equiv 0$ near p for $1 \le i \le m$

(b) $f \equiv$ (constant) on some open set containing p.

Hint: When transferred to coordinate space via a chart about p, this can be proved by Calculus methods.
Note: It follows from the definition of "connectedness" that $f \equiv$ (constant) on any connected open set $U \subseteq M$ such that $df \equiv 0$ throughout U. $\square$

2.7 Example. Let $M = \mathbb{R}^2 \sim \{0\}$ (punctured plane). The angle-variable function

$$\theta(x,y) = \arcsin\left(\frac{y}{\sqrt{x^2+y_2}}\right)$$

is multiple-valued, but on any open half-plane H bounded by a line through the origin there is a single-valued $\mathcal{C}^\infty$ determination of θ, and these can only differ on H by a added constant of the form $2\pi n$ ($n \in \mathbb{Z}$). However, all $\mathcal{C}^\infty$ determinations of $\theta(x,y)$, on any open set in $U \subseteq M$, have the same single-valued exterior derivative $\omega = d\theta \in \bigwedge^1 U$ such that

$$\omega_{\mathbf{x}} = \left(\frac{-y}{x^2+y^2}\right) dx + \left(\frac{x}{x^2+y^2}\right) dy \qquad \text{for } \mathbf{x} = (x,y) \in M \tag{26}$$

when ω is described in Euclidean coordinates on M. Thus ω has local primitives near every $p \neq 0$ in M, but the multiple-valued nature of the primitive $\theta(x,y)$ on the punctured plane M may prevent us from pieceing together these local solutions to get a global $\mathcal{C}^\infty$ function $f : M \to \mathbb{R}$ such that $df = \omega$ throughout M. $\square$

2.8 Exercise. The punctured plane $M = \mathbb{R}^2 \sim \{0\}$ is covered by the four open half-planes

$$H_1 = \{(x,y) : y > 0\} \quad H_2 = \{(x,y) : x > 0\} \quad H_3 = \{(x,y) : y < 0\} \quad H_4 = \{(x,y) : x < 0\}$$

For each half-plane give an explicit single-valued $\mathcal{C}^\infty$ determination of the angle variable $\theta(x,y)$. In each case show that the exterior derivative $\omega = d\theta$ has the form (26).
Hint: You may have to express your answer in terms of the functions *arcsin, arctan, or arccos* from Calculus, depending on which H_i you examine. $\square$

2.9 Exercise. The following 1-forms on $M_2 = \mathbb{R}^2 \sim \{0\}$ are described with respect to the standard Euclidean coordinates on these spaces. Identify those that can have well-defined local primitives.

1. $\omega = \Big(\dfrac{y}{x^2+y^2}\Big) \cdot (dx) + \Big(\dfrac{x}{x^2+y^2}\Big) \cdot (dy)$
2. $\omega = \Big(\dfrac{x}{x^2+y^2}\Big) \cdot (dx) + \Big(\dfrac{y}{x^2+y^2}\Big) \cdot (dy)$
3. $\omega = \dfrac{x}{(x^2+y^2)^{3/2}} \cdot (dx) + \dfrac{y}{(x^2+y^2)^{3/2}} \cdot (dy)$
4. $\omega = \ln(x^2+y^2) \cdot (dx) + \ln(x^2+y^2) \cdot (dy)$
5. $\omega = (x^2-y^2) \cdot (dx) + 2xy \cdot (dy)$
6. $\omega = (2xy) \cdot (dx) - (x^2-y^2) \cdot (dy)$

2.10 Exercise. Verify that following 1-form on $M_3 = \mathbb{R}^3 \sim \{0\}$

$$\omega = \frac{x}{(x^2+y^2+z^2)^{3/2}} \cdot (dx) + \frac{y}{(x^2+y^2+z^2)^{3/2}} \cdot (dy) + \frac{z}{(x^2+y^2+z^2)^{3/2}} \cdot (dz)$$

has as a primitive the famed "$1/r$ potential"

$$\phi(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|} \quad \text{in which } \|\mathbf{x}\| = \sqrt{x^2+y^2+z^2}. \ \square$$

Interpretations of 1-Forms and their Primitives.
In Calculus, "vector fields" on $\mathbb{R}^n$ (or open subsets thereof) are represented as $\mathbf{F}(\mathbf{x}) = \sum_{i=1}^n F_i(\mathbf{x})\,\mathbf{e}_i$ where $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are the "standard unit vectors" in $\mathbb{R}^n$, but the nature of these basis vectors is seldom mentioned. This is a problem because fields appearing in applications are not all of the same type. For example:

1. **Velocity fields** $\tilde{X}$ on $M = \mathbb{R}^n$ must be regarded as fields of *tangent vectors*. In fact, if $\gamma(t)$ is a curve of class $\mathcal{C}^{(1)}$ its vector derivative $\gamma'(t)$ is the instantaneous velocity of the moving point $p = \gamma(t)$, and this is a tangent vector in TM_p. Therefore, at any $p \in M$ the basis vectors $\mathbf{e}_i$ in the identity $\gamma'(t) = \sum_{i=1}^n v_i\,\mathbf{e}_i$ ($v_i \in \mathbb{R}$) should also be interpreted as vectors in TM_p. The instantaneous velocities of particles in a fluid flow at a particular moment in time provide a physical example of a smooth vector field of tangent vectors on an open subset $M \subseteq \mathbb{R}^n$.

2. **Gradient fields** ∇f determined by a function $f : M \to \mathbb{R}$, are classically presented in the form
$$\nabla f(\mathbf{x}) = D_{x_1} f(\mathbf{x})\,\mathbf{e}_1 + \ldots + D_{x_m} f(\mathbf{x})\,\mathbf{e}_m\ .$$
But there are many reasons to interpret every gradient field ∇f as a smoothly varying field of *cotangent vectors*, not tangent vectors, and then the basis vectors $\mathbf{e}_i$ must also be regarded as cotangent vectors. Formula (20) strongly suggests the following interpretation: taking global coordinates $x_\alpha(\mathbf{x}) = (x_1, \ldots, x_m)$ on Euclidean space M (with $U_\alpha = \mathbb{R}^m$) we have a well-defined set of basis vectors $\mathfrak{X}_p^* = \{(dx_1)_p, \ldots, (dx_m)_p\}$ in TM_p^* for every base point p. If we impose the standard

chart coordinates on $\mathbb{R}^m$ the exterior derivative $d : \bigwedge^0(M) \rightarrow \bigwedge^1(M)$ takes the form (20),

$$(df)_p = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(p) \cdot (dx_i)_p \qquad \text{for all } p \in \mathbb{R}^m \ .$$

If $f \in \mathcal{C}^\infty(\mathbb{R}^m)$ and we identify the "standard basis vectors" $\mathbf{e}_1, \ldots, \mathbf{e}_m$ attached to $p \in \mathbb{R}^m$ with the basis vectors $(dx_1)_p, \ldots, (dx_m)_p$ in the cotangent space TM_p^*, then the exterior derivative $(df)_p$ becomes the classical gradient $\nabla f(p)$ everywhere on $\mathbb{R}^m$.

3. **Electric fields** $\mathbf{E}(\mathbf{x}) = E_1(\mathbf{x})\, \mathbf{e}_1 + \ldots + E_m(\mathbf{x})$ are also fields of cotangent vectors. It has long been realized by physicists that, at least locally, $\mathbf{E}$ fields are gradients $\mathbf{E} = \nabla\phi$ of scalar "potential functions" $\phi : M \rightarrow \mathbb{R}$. If we identify the $\mathbf{e}_i$ at a base point p with the 1-forms $(dx_i)_p$ determined throughout $\mathbb{R}^m$ by the standard coordinate chart, then the classical statement $\mathbf{E} = \nabla\phi$ becomes a statement about the exterior derivative of ϕ,

$$\mathbf{E}(\mathbf{x}) = (d\phi)_\mathbf{x} = D_{x_1}\phi(\mathbf{x})\,(dx_1)_\mathbf{x} + \ldots + D_{x_m}\phi(\mathbf{x}) \cdot (dx_m)_\mathbf{x}$$

for all $\mathbf{x} \in \mathbb{R}^m$.

4. **Magnetic fields** are classically described as fields of vectors, but the "vectors" involved are something quite different, neither tangent vectors nor cotangent vectors. These fields are in fact represented by fields of antisymmetric tensors of rank-2, which assign at each base point p a nondegenerate antisymmetric bilinear form B_p on the tangent space TM_p. In the next section we will see how these arise in the scheme (24). Understanding this last statement requires an excursion into multilinear algebra.

XI.3 Tensor Fields and the Exterior Derivative.

Multilinear forms (tensors) on an arbitrary finite-dimensional vector space V have already been discussed in Chapter IX of these *Notes*. The **rank-k tensors** on V are the multilinear forms $\omega : V \times \ldots \times V \rightarrow \mathbb{R}$ (k factors), which become a vector space $V^{(0,k)}$ when equipped with the addition and scaling laws operations

$$\begin{array}{rcl} (\omega_1 + \omega_2)(v_1, \ldots, v_k) & = & \omega_1(v_1, \ldots, v_k) + \omega_2(v_1, \ldots, v_k) \\ (\lambda \cdot \omega)(v_1, \ldots v_k) & = & \lambda \cdot \omega(v_1, \ldots, v_k) \end{array}$$

Tensors of rank-1 are just vectors in the dual space $V^* = V^{(0,1)}$ and the rank-0 tensors on V are the scalars $V^{(0,0)} = \mathbb{R}$. Tensors of arbitrary mixed type $V^{(r,k)}$ on V can be defined but we won't need them in this account; in this notation the space V itself is denoted by $V = V^{(1,0)}$ and consists of tensors of "type (1,0)."

If $\mathfrak{X} = \{e_i\}$ is a basis for V and $\mathfrak{X}^* = \{e_i^*\}$ is the dual basis in V^*, then the space $V^{(0,k)}$ of k-linear forms $\omega : V \times \ldots \times V \rightarrow \mathbb{R}$ is a vector space of dimension $\dim V^{(0,k)} = m^k$ if $\dim(V) = m$. We will see that it is spanned by "monomials," which are "tensor products" of vectors in the dual basis $\mathfrak{X}^*$: if $I = (i_1, \ldots, i_k)$ is a multi-index with each $i_j \in [1, m]$ we define the corresponding **monomial** $e_I^* \in V^{(0,k)}$ to be the k-linear map

$$\text{(27)} \qquad e_I^* = e_{i_1}^* \otimes \ldots \otimes e_{i_k}^* : (v_1, \cdots, v_k) \mapsto \prod_{j=1}^{k} \langle e_{i_j}^*, v_j \rangle \quad \text{where the } v_j \in V$$

These turn out to be a basis for $V^{(0,k)}$, so every rank-k multi-linear form can be written uniquely as $\omega = \sum_{I \in [1,m]^k} c_I e_I^*$ with $c_I \in \mathbb{R}$.

Our main interest in these *Notes* will be tensors that act on the tangent spaces at base points on a differentiable manifold M. At a base point p on M we consider tensors that act on $V = \mathrm{TM}_p$. A **tensor field of rank-k** assigns a k-linear form $\omega_u \in \mathrm{TM}_u^{(0,k)}$ at each base point $u \in M$, but many results are purely algebraic and are true for arbitrary vector spaces V. Given a chart (x_α, U_α) on M and base point $u \in U_\alpha$, we have bases $\{\partial/\partial x_i|_u\}$ in TM_u and dual bases $\{(dx_i)_u\}$ in TM_u^*, and then there are uniquely determined coefficient functions $c_I(u)$ on U_α for $I \in [1, m]^k$ such that

$$\omega_u = \sum_I c_I(u)\,(dx_I)_u = \sum_I c_I(u) \cdot (dx_{i_1})_u \otimes \ldots \otimes (dx_{i_k})_u \quad \text{for all } u \in U_\alpha$$

A tensor field ω on M is **smooth** if the coefficient functions $c_I(u)$ are in $\mathcal{C}^\infty(U_\alpha)$ for all charts on M. It is routine (but messy) to verify that smoothness of ω when described in one chart implies smoothness with respect to any other chart in the maximal atlas.

The space $\mathcal{D}^{(0,k)}(M)$ of smooth rank-k tensor fields is an infinite-dimensional vector space, and also a $\mathcal{C}^\infty(M)$-module if we define

$$(f\omega)_u = f(u) \cdot \omega_u \quad \text{ for } u \in M \text{ and } f \in \mathcal{C}^\infty(M).$$

The space $\mathcal{D}^{(0,1)}(M)$ is precisely the space $\bigwedge^1(M)$ of smooth 1-forms on M, and $\mathcal{D}^{(0,0)}(M) = \bigwedge^0(M) = \mathcal{C}^\infty(M)$ by definition. The space $\mathcal{D}^{(0,2)}(M)$ of smooth rank-2 tensor fields consists of smoothly varying fields of bilinear forms on the tangent spaces TM_p, etc.

3.1 Example (Riemannian Structure on M). On any vector space V an **inner product** $g(v_1, v_2)$ is a particular type of rank-2 tensor in $V^{(0,2)}$; there are many possible inner products on V. A **Riemannian structure** on a manifold M is a smooth field of inner products $p \mapsto g_p$ with values $g_p \in \mathrm{TM}_p^{(0,2)}$. When M is equipped with this extra structure we can define

- **Length** $\|X_p\| = \sqrt{g_p(X_p, X_p)}$ of any vector $X_p \in \mathrm{TM}_p$. This determines a vector space norm on each tangent space that allows us to speak of the "length" of a tangent vector.

- **Orthogonality** of vectors in TM_p, which we interpret to mean $g_p(X_p, Y_p) = 0$, and the *angle* between two nonzero tangent vectors, which is determined by

$$\cos\big(\theta(X_p, Y_p)\big) = \frac{g_p(X_p, Y_p)}{\|X_p\|\,\|Y_p\|}$$

Lengths, angles, and orthogonality of tangent vectors cannot be defined in the absence of a Riemannian structure on M.

In particular it now becomes meaningful to speak of *orthonormal* bases in each tangent space TM_p, as well as more exotic constructs such as a smooth "field of orthogonal frames:" a family of smooth vector fields $\{\tilde{X}_1, \ldots, \tilde{X}_m\}$ on M such that

$$(\tilde{X}_1)_p, \ldots, (\tilde{X}_m)_p \text{ is an orthonormal basis in } \mathrm{TM}_p \text{ for each } p$$

(orthonormal with respect to the inner product $g_p : \mathrm{TM}_p \times \mathrm{TM}_p \to \mathbb{R}$).

Furthermore, if $\gamma : [a, b] \to \mathbb{R}$ is a $\mathcal{C}^\infty$ curve, or even one of class $\mathcal{C}^{(1)}$, its length is given by the Riemann integral

$$\textsc{Arc Length:} \quad L(\gamma) = \int_a^b \|\gamma'(t)\|\,dt = \int_a^b \sqrt{g_{\gamma(t)}\big(\gamma'(t), \gamma'(t)\big)}\,dt$$

where $\gamma'(t)$ is the tangent vector to the curve at $p = \gamma(t)$. The change of variable formula for Riemann integrals shows that the value of $L(\gamma)$ is unchanged, with $L(\eta) = L(\gamma)$, if η is any orientation-preserving reparametrization of γ

$$\eta = \gamma \circ \phi \quad \text{where} \quad \phi : [c, d] \to [a, b] \text{ is } \mathcal{C}^{(1)} \text{ with } \frac{d\phi}{ds} > 0 \text{ for all } s.$$

For orientation-reversing reparametrizations we get $L(\eta) = (-1)\cdot L(\gamma)$.

With considerably more effort one can show that if $p \in M$ there is an open neighborhood U such that any $q \in U$ can be connected to p by a *geodesic* – a $\mathcal{C}^{(1)}$ curve $\gamma_0 : [a, b] \to U$ such that $\gamma_0(a) = p, \gamma_0(b) = q$, $\gamma_0(t) \in U$ for all t, and

$$L(\gamma_0) \leq L(\gamma) \text{ for } \textit{any } \mathcal{C}^{(1)} \text{ curve } \eta \text{ in } U \text{ that connects } p \text{ to } q$$

These "minimal length" curves in M are the analogs of straight line segments when Euclidean space $\mathbb{R}^n$ is equipped with its "natural" Riemannian structure, obtained by taking the standard global chart $x_\alpha(\mathbf{x}) = (x_1, \ldots, x_n)$ on $\mathbb{R}^n$ and defining $g_p = \sum_{i=1}^n (dx_i)_p \otimes (dx_i)_p$, so that

$$g_p\Big(\frac{\partial}{\partial x_i}\Big|_p , \frac{\partial}{\partial x_j}\Big|_p\Big) = \delta_{ij} \quad \text{(Kronecker)}$$

The bases $\{(\partial/\partial x_i|_p)\}$ induced in TM_p by the chart coordinates are then *orthonormal bases* with respect to the inner product g_p in each tangent space to $\mathbb{R}^n$.

This notion of arc length also induces a natural metric on any Riemannian manifold

$$d_M(p, q) = \inf\{L_g(\gamma) : \gamma \text{ any } \mathcal{C}^\infty \text{ curve such that } \gamma(a) = p, \gamma(b) = q\,\}$$

(However, verifying that d_M satisfies the triangle inequality takes considerable effort.) While $\inf\{\ldots\}$ is actually achieved by a unique $\mathcal{C}^\infty$ curve (a geodesic) for all q sufficiently close to p, this need not be true for points far from p in the manifold. Even if the minimal length $d_M(p, q)$ is achieved, the geodesic connecting p to q might not be unique if p and q are widely separated. Think of the north and south poles on the unit sphere $S^2 \subseteq \mathbb{R}^3$. The sphere inherits a natural Riemannian structure from the surrounding Euclidean space and the length-minimizing geodesics in it are segments of great circles (intersections of the sphere with planes through the origin). Every great circle from N to S is a geodesic with the same (minimal) length $L(\gamma) = \pi$. $\square$

We won't have time to explore the geometry of Riemannian manifolds in these *Notes*, but we emphasize that there are many examples. An excellent account of this subject is given in the book *Riemannian Geometry*, by Manfredo do Carmo. All $\mathcal{C}^\infty$ manifolds embedded in a Euclidean space $\mathbb{R}^n$, such as the spheres S^n of various dimensions, or smooth level hypersurfaces determined via the Implicit Function Theorem, inherit a natural Riemannian structure induced by the standard Riemannian structure on $\mathbb{R}^n$ described above. This Riemannian structure will, however, depend on how the manifold is embedded in $\mathbb{R}^n$.

Action of the Permutation Group S_k on Tensors and Tensor Fields.
Permutations $\sigma \in S_k$ act as linear operators on the space of rank-k tensors $t \in V^{(0,k)}$ if we let

$$\sigma \cdot t(v_1, \cdots, v_k) = t(v_{\sigma(1)}, \cdots, v_{\sigma(k)}) \qquad \text{for } v_j \in V\ .$$

Thus in evaluating the action of $\sigma \cdot t$, the inputs $v_1, \cdots, v_k$ are permuted by $\sigma \in S_k$ before being fed into the tensor t. Then $\sigma : V^{(0,k)} \to V^{(0,k)}$ is a linear map and we obtain a "left action of S_k on tensors," which means that

(28) Left Action Law: $\quad (\sigma\tau) \cdot t = \sigma \cdot (\tau \cdot t) \qquad$ for all $\sigma, \tau \in S_k$ and $t \in V^{(0,k)}$.

Proof of (28): Many find the proof of this property confusing. For a straightforward argument that does the job, observe that

$$\begin{aligned}
\sigma\cdot(\tau\cdot t)(v_1,\ldots,v_k) &= (\tau\cdot t)\big(v_{\sigma(1)},\ldots,v_{\sigma(k)}\big)\\
&= (\tau\cdot t)(w_1,\ldots,w_k)\Big|_{w_1=v_{\sigma(1)},\ldots,w_k=v_{\sigma(k)}}\\
&= t\big(w_{\tau(1)},\ldots,w_{\tau(k)}\big)\Big|_{w_i=v_{\sigma(i)}}\\
&= t\big(v_{\sigma(\tau(1))},\ldots,v_{\sigma(\tau(k))}\big) \qquad (\text{since } w_{\tau(i)}=v_{\sigma(\tau(i))})\\
&= t\big(v_{(\sigma\tau)(1)},\ldots,v_{(\sigma\tau)(k)}\big) \qquad (\text{by the Group Law (28)})\\
&= (\sigma\tau)\cdot t(v_1,\ldots,v_k)
\end{aligned}$$

for all $v_i \in V$. $\square$

3.2 Exercise. If $\ell_1,\ldots,\ell_k \in V^*$ we have defined their tensor product to be the k-linear form $\ell_1\otimes\ldots\otimes\ell_k$ as in (27). Verify that the action of S_k on such monomials is given by

$$\sigma\cdot(\ell_1\otimes\ldots\otimes\ell_k) = \ell_{\sigma^{-1}(1)}\otimes\ldots\otimes\ell_{\sigma^{-1}(k)}$$

(Compare with the inital definition.) $\square$

All of this applies to manifolds M if we we take $V = \mathrm{TM}_u$ at base points $u \in M$. If ω is a tensor field in $\mathcal{D}^{(0,k)}(M)$, so $\omega_u \in \mathrm{TM}_u^{(0,k)}$ for all base points in M, we let σ act independently on each tangent space. Then $(\sigma\cdot\omega)_u = \sigma\cdot(\omega_u)$ in $\mathrm{TM}_u^{(0,k)}$ for each $u \in M$ and we get a group action of S_k by linear operators

$$\sigma : \mathcal{D}^{(0,k)}(M) \to \mathcal{D}^{(0,k)}(M)$$

on smooth rank-k tensor fields.

Antisymmetric tensors and tensor fields on manifolds are objects of special interest in Calculus on Manifolds.

3.3 Definition. *If V is a finite-dimensional vector space, a tensor $\omega \in V^{(0,k)}$ is* **symmetric** *or* **antisymmetric** *if*

(29) $$\sigma\cdot\omega = \omega \text{ for all } \sigma \in S_k \qquad or \qquad \sigma\cdot\omega = sgn(\sigma)\cdot\omega$$

where $sgn(\sigma)$ is the signature of the permutation $\sigma \in S_k$. (Anti-)symmetry of a tensor field $\omega \in \mathcal{D}^{(0,k)}(M)$ means ω_p is an (anti-)symmetric tensor at each base point.

*On a manifold M, smooth fields of antisymmetric rank-k tensors are called k-***forms***, or* **differential forms of rank-***k, and the space of all such tensor fields is denoted $\bigwedge^k(M)$. This is a vector subspace of $\mathcal{D}^{(0,k)}(M)$; both become $\mathcal{C}^\infty(M)$-modules if we define an action $\mathcal{C}^\infty(M)\times\bigwedge^k(M)\to\bigwedge^k(M)$, letting*

$$(f\cdot\omega)_u(v_1,\cdots,v_k) = f(u)\cdot w_u(v_1,\cdots,v_k) \quad \text{for } v_1,\cdots v_k \in \mathrm{TM}_k, u \in M\ .$$

The space of **symmetric tensor fields** *of rank-k is denoted $S^k(M)$ and the symmetric rank-k tensors on TM_p are denoted by $S^k\big(\mathrm{TM}_p\big)$.*

Antisymmetric tensors on a vector space V can be created by "antisymmetrizing" arbitrary rank-k tensors in $V^{(0,k)}$ via a surjective linear projection map

$$\mathrm{Alt} : V^{(0,k)} \to \bigwedge^k(V) \subseteq V^{(0,k)}$$

defined as follows.

(30) $$\mathrm{Alt}(\omega) = \frac{1}{k!}\sum_{\sigma\in S_k} sgn(\sigma)\,(\sigma\cdot\omega) \qquad \text{for all } \omega \in V^{(0,k)}\ .$$

The fudge factor “$1/k!$” is needed to make *Alt* a projection, with $(Alt)^2 = Alt$. A similar *symmetrization operator* projecting $S : V^{(0,k)} \to S^k(V)i$ is obtained by dropping the factor “$sgn(\sigma)$” in Definition (30), but symmetrzation will not play a role in the present narrative.

On a manifold M, doing this to $V = \mathrm{TM}_u^{(0,k)}$ at each base point yields a smooth field of rank-k antisymmetric tensors; the resulting tensor field $\mathrm{Alt}(\omega)$ is then given by

$$\big(\mathrm{Alt}(\omega)\big)_u = \mathrm{Alt}(\omega_u) \qquad \text{for all } u \in M, \omega \in \mathcal{D}^{(0,k)}(M). \tag{31}$$

We present the basic facts about antisymmetrization without proof – see Appendix XI-C for details.

3.4 Theorem. *If V is a vector space, $Alt : V^{(0,k)} \to V^{(0,k)}$ is a linear map such that*

1. *$(Alt)^2 = Alt \circ Alt$ is equal to Alt, so Alt is a projection onto its range in $V^{(0,k)}$.*
2. *$\mathrm{Alt}\big(V^{(0,k)}\big) \subseteq \bigwedge^k(V^{(0,k)})$, and if ω was an antisymmetric tensor on V to begin with then $\mathrm{Alt}(\omega) = \omega$, so $Alt = \mathrm{id}$ on $\bigwedge^k(V)$.*

As an operator on tensor fields, Alt maps $\mathcal{D}^{(0,k)}(M) \to \bigwedge^k(M)$ if we take $V = \mathrm{TM}_u^{(0,k)}$ in (31) *at every base point.*

Tensor Product of Tensors and Tensor Fields.
If $\omega \in V^{(0,k)}$ and $\mu \in V^{(0,\ell)}$ their **tensor product** $\omega \otimes \mu$ is a rank-$(k+\ell)$ tensor such that

$$\omega \otimes \mu(v_1, \cdots, v_k, v_{k+1}, \cdots, v_{k+\ell}) = w(v_1, \cdots, v_k) \cdot \mu(v_{k+1}, \cdots, v_{k+\ell})$$

The $(\otimes)$ operation is easily seen to be associative, with

$$\omega \otimes (\mu \otimes \tau) = (\omega \otimes \mu) \otimes \tau \quad \text{ in } V^{k+\ell+m} \ ,$$

but it is not commutative: $\omega \otimes \mu$ and $\mu \otimes \omega$ are generally different elments of $V^{(0,k+\ell)}$.

The **wedge product** $\omega \wedge \mu$ is a bilinear map of antisymmetric tensors

$$\textstyle\bigwedge^k(V) \times \bigwedge^\ell(V) \to \bigwedge^{k+\ell}(V)$$

that takes $\omega \in \bigwedge^k(V)$ and $\mu \in \bigwedge^\ell(V)$ to an antisymmetric tensor of rank $k+\ell$ on V,

$$\omega \wedge \mu = \frac{(k+\ell)!}{k!\ell!}\mathrm{Alt}(\omega \otimes \mu) \ . \tag{32}$$

Note that $\omega \otimes \mu$ need not be antisymmetric even if both factors were antisymmetric, so $\omega \otimes \mu$ has to be antisymmetrized to end up in $\bigwedge^{k+\ell}(V)$. The algebraic properties of the wedge product are listed below. The purely algebraic proofs (some difficult) are relegated to Appendix XI-C.

3.5 Theorem. *Let V be a finite dimensional vector space with basis $\mathfrak{X} = \{e_i\}$ and dual basis $\mathfrak{X}^* = \{e_i^*\}$ in V^*.*

1. *If $A : V \to V$ is a linear operator its transpose A^{t} acts on the dual space V^*, with*

   $$\langle A^{\mathrm{t}}\ell, v\rangle = \langle \ell, Av\rangle \qquad \text{for } v \in V, \ell \in V^* \ .$$

   This construction can be extended to define a transpose $A^{\mathrm{t}} : V^{(0,k)} \to V^{(0,k)}$ that acts on tensors of rank-k, by letting A act on each input vector v_i in $(v_1, \cdots, v_k)$,

   $$A^{\mathrm{t}}\omega(v_1, \cdots, v_k) = \omega\big(A(v_1), \ldots, A(v_k)\big)$$

   The operator A^{t} leaves the subspace $\bigwedge^k(V)$ invariant, so A^{t} acts on the space of antisymmetric tensors $\bigwedge^k(V)$ as well as the full space $V^{(0,k)}$ of rank-k tensors on V.

2. *A^{t}* **respects wedge products:** *$A^{\mathrm{t}}(\omega\otimes\mu) = A^{\mathrm{t}}w\otimes A^{\mathrm{t}}\mu$ and $A^{\mathrm{t}}(\omega\wedge\mu) = A^{\mathrm{t}}\omega\wedge A^{\mathrm{t}}\mu$ for $\omega \in \bigwedge^k(V)$ and $\mu \in \bigwedge^\ell(V)$.*

3. **Antisymmetry of wedge:** *If ϕ_1, ϕ_2 are in V^* (rank-1 tensors on V) their wedge product is an antisymmetric tensor in $V^{(0,2)}$,*

$$\phi_1 \wedge \phi_2 = \frac{1}{2}(\phi_1 \otimes \phi_2 - \phi_2 \otimes \phi_1) = (-1)\,\phi_2 \wedge \phi_1 \ . \tag{33}$$

   In particular $\phi \wedge \phi = 0$ in $V^{(0,2)}$ for every rank-1 tensor $\phi \in V^*$. If $\omega \in \bigwedge^k(V)$, $\mu \in \bigwedge^\ell(V)$ we have the more general commutation relations

$$\omega \wedge \mu = (-1)^{kl} \mu \wedge \omega \tag{34}$$

4. **Associativity of wedge:** *If $\omega \in \bigwedge^k(V)$, $\mu \in \bigwedge^\ell(V)$, $\tau \in \bigwedge^m(V)$ then $\omega\wedge(\mu\wedge\tau) = (\omega\wedge\mu)\wedge\tau$. Associativity means that when we form the wedge product $\omega_1 \wedge \ldots \wedge \omega_r$ of antisymmetric tensors of various ranks we don't have to worry about where to put the parentheses. The proof yields a fact that greatly simplifies many calculations involving iterated wedge products.*

$$\omega \wedge (\mu \wedge \tau) = (\omega \wedge \mu) \wedge \tau = \frac{(k+\ell+m)!}{k!\ell!m!}\,\mathrm{Alt}(\omega \otimes \mu \otimes \tau) \tag{35}$$

   Notice that evaluation of

$$\omega \wedge (\mu \wedge \tau) = \frac{(k+\ell+m)!}{k!\ell!m!}\,\mathrm{Alt}\big(\omega \otimes (\mathrm{Alt}(\mu \otimes \tau)\big)$$

   involves two applications of "Alt" while the last expression in (35) *requires only one.*

5. **Evaluating Monomial Wedge Products:** *$e_J = e_{j_1} \wedge \ldots \wedge e_{j_k}$ when J is not an ordered monomial. If $\mathfrak{X} = \{e_1, \ldots, e_m\}$ is any basis in V and $\mathfrak{X}^* = \{e_1^*, \ldots, e_m^*\} \subseteq V^*$ is the dual basis, we can define $e_J^* = e_{j_1}^* \wedge \ldots \wedge e_{j_k}^*$ for any multi-index $J = (j_1, \ldots, j_k)$. One important computational fact is:*

$$e_1^* \wedge \ldots \wedge e_m^*(e_1, \ldots, e_m) = 1 \ ,$$

   More generally we can evaluate

$$e_{j_1}^* \wedge \ldots \wedge e_{j_k}^*\big(e_{i_1}, \ldots, e_{i_k}\big) = e_J^*\big(e_{i_1}, \ldots, e_{i_k}\big)$$

   for arbitrary multi-indices I, J in $[1,m]^k$ (ordered or not). The outcome is zero if there are any repeated entries in the k-tuple $J = (j_1, \ldots, j_k)$, and if the entries are distinct there is a unique permutation $\sigma \in S_k$ such that $I = \sigma\cdot J = (j_{\sigma(1)}, \ldots, j_{\sigma(k)})$ is in $\mathcal{E}_k$, with $x_J = \pm x_I$.

3.6 Theorem (The Basis Theorem). *Given a basis $\mathfrak{X} = \{e_i\}$ in V and dual basis $\mathfrak{X}^* = \{e_i^*\}$ in V^* we define the set $\mathcal{E}_k$ of* **ordered k-tuples**

$$I = (i_1, \ldots, i_k) \text{ such that } 1 \le i_1 < \ldots < i_k \le m = \dim(V) \ .$$

The corresponding "ordered monomials" in $\bigwedge^k(V)$

$$e_I^* = e_{i_1}^* \wedge \ldots \wedge e_{i_k}^*, \qquad (I \in \mathcal{E}_k)$$

are a basis for the space of k-forms $\bigwedge^k(V)$. In particular,

$$\dim\big(\textstyle\bigwedge^k(V)\big) = \binom{m}{k} = \#(\text{multi-indices in } \mathcal{E}_k)$$

When $k = 0$ we have $\dim \bigwedge^0(V) = \dim_{\mathbb{R}}(\mathbb{R}) = 1$. *Likewise the dimension of the space* $\bigwedge^m(V)$ *of antisymmetric tensors of maximal rank* $m = \dim(V)$ *is also equal to* 1, *and* $\bigwedge^k(V) = \{0\}$ *for all* $k > m = \dim(V)$.

3.7 Exercise. If $\phi_1, \phi_2 \in V^*$ verify by example that $\phi_1 \otimes \phi_2$ need not equal $\phi_2 \otimes \phi_1$ as elements of $V^{(0,2)}$. (Equality of tensors means they have the same action on all inputs (v_1, v_2). Remember that the inputs are *ordered lists*, so $(v_1, v_2) \neq (v_2, v_1)$.) $\square$

3.8 Exercise. Let V be a finite-dimensional vector space, let $\mathfrak{X} = \{e_i\}$ be a basis and $\mathfrak{X}^* = \{e_i^*\}$ the dual basis. In Chapter IX we explained how every bilinear form $B : V \times V \to \mathbb{R}$ is described by an $m \times m$ matrix ($m = \dim V$)

$$[B_{ij}]_{\mathfrak{X}} \qquad \text{with entries} \qquad B_{ij} = B(e_i, e_j)$$

Given two vectors $\phi_1 = \sum_{i=1}^m a_i e_i^*$, $\phi_2 = \sum_{i=1}^m b_i e_i^*$ in V^* we get the bilinear forms $\phi_1 \otimes \phi_2$ and $\phi_2 \otimes \phi_1$ on V. Compute the associated matrices with respect to $\mathfrak{X}$. Are they equal? $\square$

Many computations involving antisymmetric tensors reduce to working with wedge products $\phi_1 \wedge \ldots \wedge \phi_k$ of rank-1 tensors $\phi \in V^*$. The following simple formula, easily derived by induction from the associative law (35), provides an easy way to directly evaluate such products.

$$\phi_1 \wedge \ldots \wedge \phi_k = \frac{(1 + \ldots + 1)!}{1! \cdots 1!} \cdot \operatorname{Alt}(\phi_1 \otimes \ldots \otimes \phi_k) = k! \operatorname{Alt}(\phi_1 \otimes \ldots \otimes \phi_k)$$

Note that the factor out front is now $k!$, not $\frac{1}{k!}$ as in (35).

If $\{e_i\}$ is a basis for V and $\{e_i^*\}$ is the dual basis in V^* one can evaluate a wedge product $e_J^* = e_{j_1}^* \wedge \ldots \wedge e_{j_k}^*$ for an arbitrary multi-index $J = (j_1, \ldots, j_k)$ of length k with $j_s \in [1, m]$ whose entries need not be increasing or even distinct.

3.9 Lemma. *If V is a vector space with basis $\{e_1, \ldots, e_m\}$ and dual basis $\{e_1^*, \ldots, e_m^*\}$ in V^*, and if $e_J^* = e_{j_1}^* \wedge \ldots \wedge e_{j_k}^*$ is an arbitrary monomial in the dual basis vectors, with $j_1, \ldots, j_k \in [1, m]$, but not necessarily with $j_1 < \ldots < j_k$, then*

(a) *If we interchange adjacent factors $e_{j_s}^* \leftrightarrow e_{j_{s+1}}^*$ in the wedge product e_J^* we get $-e_J^*$.*

(b) *The wedge product e_J is zero if there is a repeated vector, say $e_{j_r}^* = e_{j_s}^*$ with $r \neq s$. (This is the real reason $\bigwedge^r(V) = (0)$ when $r > m = \dim(V)$.)*

(c) *If the indices $(j_1, \ldots, j_k)$ are distinct there is a unique permutation $\sigma \in S_k$ such that $I = \sigma \cdot J = \big(j_{\sigma(1)}, \ldots, j_{\sigma(k)}\big)$ has the same indices, listed in increasing order. Then e_I^* is one of the standard basis vectors for $\bigwedge^k(V)$ and $e_J^* = sgn(\sigma) \cdot e_I^*$.*

In any case, e_J^* is either zero or is $\pm e_I^*$ for a unique standard basis vector $e_I^* \in \bigwedge^k(V)$,

The proof is outlined in the following Exercise.

3.10 Exercise. Use the rules set forth in Theorem 3.5 together with the following Hints to prove the claims (a) - (c) in Lemma 3.9.
Hints: Part (a) is immediate from antisymmetry of "$\wedge$." If there are identical factors in e_J^*, repeated swapping of adjacent factors will bring the identical factors together, where they annihilate each other because $\phi \wedge \phi = 0$. (Also, tensor products $\omega \otimes \ldots \otimes \mu$ (or $\omega \wedge \ldots \wedge \mu$ for antisymmetric tensors) are zero if any factor is zero.)

If the factors in e_J^* are distinct then for some permutation of the entries yields a multi-indent I in $\mathcal{E}_k$; then $J = \sigma \cdot I$ for some $\sigma \in S_k$ and then $e_J^* = sgn(\sigma) \cdot e_I^*$. $\square$

Calculating $\omega \wedge \mu$ on a Chart in M.

If $J = \{j_1, \ldots, j_k\}$ is an arbitrary multi-index of length k (entries not necessarily increasing or distinct), we can still form the wedge product $(dx)_J = (dx_{j_1}) \wedge \ldots \wedge (dx_{j_k})$ of the 1-forms (dx_i) on U_α determined by the chart. But in view of the anti-commutation relations $\omega \wedge \mu = (-1)^{kl} \mu \wedge \omega$ for wedge products, we have $e_J^* = 0$ unless the entries in J are distinct, and if they are distinct then $e_J^* = \pm e_I^*$ for a unique standard basis vector e_I^*, $I \in \mathcal{E}_k$.

One last thing should be noted. In discussing wedge products $\omega \wedge \mu$ of differential forms on M we often encounter "weighted" expressions like $f\,(dx_I) \wedge h\,(dx_J)$ involving $\mathcal{C}^\infty$ functions f, h. But by definition, $(\omega \wedge \mu)_u = \omega_u \wedge \mu_u$ at every base point, hence

$$(f\omega) \wedge (h\mu) = (fh) \cdot (\omega \wedge \mu) \tag{36}$$

as differential forms on M.

3.11 Example. Taking $M = \mathbb{R}^2$, let U_α be an open set on which Cartesian coordinates $x_\alpha(\mathbf{x}) = (x, y)$ and polar coordinates $y_\beta(\mathbf{x}) = (r, \theta)$ are defined. The smooth 1-forms $dx \wedge dy$ and $dr \wedge d\theta$ each provide a basis for the 1-dimensional space $\bigwedge^2(\mathrm{TM}_p)$ at each $p \in U_\alpha$. We shall rewrite the smooth 2-form $dx \wedge dy$ as $F(r, \theta) \cdot (dr \wedge d\theta)$ using the properties described in Theorem 3.5 and the change-of-variable formula for cotangent vectors, Proposition 2.5.

Discussion: The chart maps have scalar components $(x, y) = x_\alpha(\mathbf{x}) = \big(X(\mathbf{x}), Y(\mathbf{x})\big)$ and $(r, \theta) = y_\beta(\mathbf{x}) = \big(R(\mathbf{x}), \Theta(\mathbf{x})\big)$ for $\mathbf{x} \in M$, and the coordinate transition maps are given by

$$X = R\cos(\Theta),\ Y = R\sin(\Theta) \quad \text{and} \quad R = (X^2 + Y^2)^{1/2},\ \Theta = \arctan(Y/X)\ . \tag{37}$$

Apply the exterior derivative to the scalar component functions X, Y in (37) to get

$$\begin{aligned}
dx = dX &= \frac{\partial X}{\partial r} \cdot dr + \frac{\partial X}{\partial \theta} \cdot d\theta \\
&= \frac{\partial}{\partial r}\{R\cos\Theta\} \cdot dr + \frac{\partial}{\partial \theta}\{R\sin\Theta\} \cdot d\theta \\
&= \Big[\,\frac{\partial R}{\partial r}\cos(\Theta) - R\sin(\Theta) \cdot \frac{\partial \Theta}{\partial r}\,\Big] \cdot dr + \\
&\qquad + \Big[\,\frac{\partial R}{\partial \theta}\sin(\Theta) + R\cos(\Theta) \cdot \frac{\partial \Theta}{\partial \theta}\,\Big] \cdot d\theta\ ,
\end{aligned}$$

where $dX = (dx)$ is the exterior derivative of the scalar component $X : U_\alpha \to \mathbb{R}$ in the chart map x_α. (By (21), (dx) and dX are the same smooth 1-form on U_α.) By (6) we have

$$\frac{\partial R}{\partial r} \equiv 1,\ \frac{\partial \Theta}{\partial \theta} \equiv 1,\ \frac{\partial R}{\partial \theta} \equiv \frac{\partial \Theta}{\partial r} \equiv 0,$$

on U_α, so on the chart domain

$$dx = dX = \cos(\Theta) \cdot dr - R\sin(\Theta) \cdot d\theta\ .$$

The same sort of calculation yields

$$dy = dY = \sin(\Theta) \cdot dr + R\cos(\Theta)\ \cdot d\theta\ .$$

Adopting the time-honored abuse of notation that ignores the distinction between a function $R(\mathbf{x})$ and its values r, etc, we get

$$\begin{aligned}
dx \wedge dy &= \big[\cos(\theta) \cdot dr - r\sin(\theta) \cdot d\theta\big] \wedge \big[\sin(\theta) \cdot dr + r\cos(\theta) \cdot d\theta\,\big] \\
&= \sin(\theta)\cos(\theta) \cdot (dr \wedge dr) - r\sin^2(\theta) \cdot (d\theta \wedge dr) + \\
&\qquad + r\cos^2(\theta) \cdot (dr \wedge d\theta) - r^2\sin(\theta)\cos(\theta) \cdot (d\theta \wedge d\theta) \\
&= -r\sin^2(\theta) \cdot (d\theta \wedge dr) + r\cos^2(\theta) \cdot (dr \wedge d\theta) \\
&= r\,(dr \wedge d\theta)
\end{aligned}$$

Indeed, $dr \wedge dr = d\theta \wedge d\theta \equiv 0$, and $d\theta \wedge dr = -dr \wedge d\theta$ by anti-symmetry. $\square$

3.12 Exercise. On $\mathbb{R}^2$ compute $dr \wedge d\theta$ in terms of $dx \wedge dy$. On $\mathbb{R}^3$ compute $dx \wedge dy \wedge dz$ in terms of $d\rho \wedge d\phi \wedge d\theta$ (spherical coordinates). Spherical coordinates (ρ, θ, ϕ) and Cartesian coordinates (x, y, z) on $\mathbb{R}^3$ are related via

$$z = \rho \sin(\phi) \qquad y = \rho \sin(\theta) \cos(\phi) \qquad x = \rho \cos(\theta) \cos(\phi) \ ,$$

see Figure 11.3. $\square$

The General Exterior Derivative $d_k : \bigwedge^k(M) \rightarrow \bigwedge^{k+1}(M)$**.** Smooth rank-$k$ differential forms ω are tensor fields whose values ω_p lie in $\bigwedge^k(\mathrm{TM}_p)$ for each $p \in M$. Given any chart (x_α, U_α), such a field is uniquely described as a sum

$$\omega_u = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (dx_{i_1})_u \wedge \ldots \wedge (dx_{i_k})_u = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (dx_I)_u \quad \text{ for } u \in U_\alpha \ ,$$

and smoothness means the c_I are in $\mathcal{C}^\infty(U_\alpha)$ for all charts. The general exterior derivatives d_k

$$\textstyle\bigwedge^0(M) \xrightarrow{d=d_0} \bigwedge^1(M) \xrightarrow{d=d_1} \bigwedge^2(M) \xrightarrow{d=d_2} \cdots \xrightarrow{d=d_{m-1}} \bigwedge^m(M) \xrightarrow{d=d_m} 0$$

are operations on rank-k differential forms, for $k = 0, 1, 2, \ldots$ The rank-0 derivative d_0 has been defined in a coordinate-free way on $\bigwedge^0(M) = \mathcal{C}^\infty(M)$ in (19): for $f \in \mathcal{C}^\infty(M)$, $(df) \in \bigwedge^1(M)$ is the 1-form such that.

$$\langle (df)_p, X_p \rangle = \langle X_p, f \rangle \qquad \text{for all } X_p \in \mathrm{TM}_p \ . \tag{38}$$

As in (20), in local chart coordinates df takes the form

$$(df)_p = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p) \cdot (dx_i)_p \qquad \text{for all } p \in U_\alpha \ .$$

Coordinate-free formulas exist describing the action of exterior derivatives d_k of all ranks $k = 0, 1, 2, \ldots$ For d_0 the formula (given above) is simple, but such descriptions become quite complicated and unintuitive for $k \geq 1$. There is, however, a fairly straightforward way to describe $d\omega$ in local chart coordinates, for ω of any rank k, and this is the approach we shall pursue. The only problem with this lies in showing that we get the same element $d\omega_p$ in $\bigwedge^{k+1}(\mathrm{TM}_p)$ at all base points $p \in M$, no matter which local chart (x_α, U_α) we use to compute it. (The proof that the definition is chart-independent is fairly arduous, and we won't have time to prove it in this brief survey.) Here is the definition of $d\omega$ in local coordinates.

3.13 Definition. (Exterior Derivative d_k**).** *If* $\omega \in \bigwedge^k(M)$ *and* (x_α, U_α) *is any chart,* ω *has a unique description on* U_α *as*

$$\omega_u = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (dx_I)_u = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (dx_{i_1} \wedge \ldots \wedge dx_{i_k}) \qquad \textit{for all } u \in U_\alpha \ .$$

Then the **exterior derivative** $d\omega \in \bigwedge^{k+1}(M)$ *is given on* U_α *by*

$$\begin{aligned} d\omega &= \sum_{I \in \mathcal{E}_k} (dc_I) \wedge (dx_I) = \sum_{I \in \mathcal{E}_k} \Big(\sum_{j=1}^m \frac{\partial c_I}{\partial x_j} \, dx_j \Big) \wedge (dx_I) \\ &= \sum_I \sum_{j=1}^m \frac{\partial c_I}{\partial x_j} \cdot (dx_j) \wedge (dx_{i_1} \wedge \ldots \wedge dx_{i_k}) \qquad (39) \end{aligned}$$

where dc_I is the usual exterior derivative of the scalar function $c_I(u)$.

At any base point $p \in U_\alpha$ the chart determines a basis $\mathfrak{X}_p^* = \{(dx_1)_p, \ldots, (dx_m)_p\}$ in TM_p^* for all $p \in U_\alpha$, and "standard" basis vectors in $\bigwedge^k(\text{TM}_p^*)$,

$$e_I^* = dx_I = (dx_{i_1} \wedge \ldots \wedge dx_{i_k}) \quad \text{for } I = (i_1 < \ldots < i_k) \text{ in } \mathcal{E}_k \ .$$

A monomial $dx_j \wedge (dx_{i_1} \wedge \ldots \wedge dx_{i_k})$ appearing in (39) will be zero in $\bigwedge^{k+1}(\text{TM}_p^*)$ if j is one of the entries in I, and will not be one of the standard basis vectors in $\bigwedge^{k+1}(\text{TM}_u^*)$ unless we happen to have $j < i_1 < \ldots < i_k$. But it is easy to rewrite any monomial $dx_j \wedge (dx_{i_1} \wedge \ldots \wedge dx_{i_k})$ appearing (39) as a sum involving ordered monomials using the rules provided in Theorem 3.5.

Note: An annoying technical issue arises in applying formula (39). This definition presumes that ω has been presented as a sum of monomials $c_I\, e_I^*$ involving standard basis vectors, with $I \in \mathcal{E}_k$. If we wish to compute the exterior derivative of a k-form $\omega = \sum_{J \in [1,m]^k} c_J \cdot e_{j_1}^* \wedge \ldots \wedge e_{j_k}^*$ involving *unordered* monomials e_J^*, it seems we would have to rewrite each summand in terms of the standard (ordered) basis monomials e_I^* with $I \in \mathcal{E}_k$ before applying (39). This could be done by repeated use of the commutation rule $dx_i \wedge dx_j = -dx_j \wedge dx_i$, transposing adjacent factors to get $dx_J = \pm dx_I$ for some $I \in \mathcal{E}_k$. That would be a terrible nuisance; the following observation saves the day.

3.13A Lemma (Exterior Derivative d_k). *Formula (39) remains valid for k-forms on a chart domain U_α*

$$\omega = \sum_{J \in [1,m]^k} c_J \cdot e_J^* \qquad (c_J \in \mathcal{C}^\infty(U_\alpha))$$

even if they involve monomials e_J^* that are unordered or have repeated indices.

Proof: If an arbitrary multi-index $J \in [1,m]^k$ has a repeated entry, say $j_r = j_s$, then $d(f\, dx_J) = 0$ and does not contribute to the sum (39). Otherwise there is a unique permutation $\sigma \in S_k$ such that $I = \sigma \cdot J$ is an ordered index in $\mathcal{E}_k$, and then for any $f \in \mathcal{C}^\infty(U_\alpha)$ we have

$$\begin{aligned} d(f \cdot x_J) &= d\big(f \cdot sgn(\sigma)\, dx_I\big) \\ &= sgn(\sigma) \cdot (df \wedge dx_I) \qquad \text{(definition of } d\text{-operator)} \\ &= df \wedge sgn(\sigma) \cdot dx_I \ = \ df \wedge dx_J \quad \square \end{aligned}$$

We will take advantage of this very useful fact in proving Lemma 3.15 below.

We now list the basic properties of the operators $d_k : \bigwedge^k(M) \to \bigwedge^{k+1}(M)$. The proofs are complicated so we defer them to Appendix XI-C in order to press on toward reinterpreting Multivariate Calculus in terms of differential forms.

3.14 Theorem. (Basic Properties of the d-Operators). *Although we have defined $d_k\omega$ on M using its description on a typical coordinate chart (x_α, U_α), the outcome is the same for all charts, so formula (39) determines a well-defined differential form $d_k\omega \in \bigwedge^{k+1}(M)$ for every $\omega \in \bigwedge^k(M)$. Furthermore,*

1. *The rank-0 derivative $d_0 : \bigwedge^0(M) \to \bigwedge^1(M)$ is just the d-operator discussed earlier in (19) and (20), such that $\langle d_0 f, X_p\rangle = \langle X_p, f\rangle$ for $X_p \in \text{TM}_p, f \in \mathcal{C}^\infty(p)$.*

2. *Each d_k is a linear operator from $\bigwedge^k(M) \to \bigwedge^{k+1}(M)$.*

3. Generalized Derivation Property: *If $\omega_1 \in \bigwedge^k(M)$ and $\omega_2 \in \bigwedge^\ell(M)$ then*

$$d_{k+\ell}\big(\omega_1 \wedge \omega_2\big) = d_k(\omega_1) \wedge \omega_2 + (-1)^k \omega_1 \wedge d_\ell(\omega_2)$$

in $\bigwedge^{k+\ell+1}(M)$. *This property governs the interaction between exterior derivatives and wedge products of differential forms on* M.

4. *By far the most interesting property is* $d^2 = 0$, *which is a shorthand for*

(40) $$d_{k+1} \circ d_k(\omega) = 0 \text{ in } \bigwedge^{k+2}(M) \quad \text{for all } \omega \in \bigwedge^k(M) \ .$$

This follows from equality of mixed second-order partial derivatives of class $\mathcal{C}^{(2)}$ *functions on* $\mathbb{R}^m$.

Here is a proof that $d^2 = 0$.

3.15 Lemma. *For each* $k = 0, 1, 2, \dots$ *we have* $0 = d^2 = d_{k+1} \circ d_k$.

Proof: It suffices to show $d^2\omega = 0$ on a coordinate chart (x_α, U_α). In local coordinates $\omega \in \bigwedge^k(M)$ is a sum over $I \in \mathcal{E}_k$ of terms like $f \cdot dx_I$; since d_k is linear, $d_k\omega$ is a sum of terms

$$d(f dx_I) = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \cdot (dx_i \wedge dx_I)$$

and then $d^2(\omega) = d_{k+1}\big(d_k\omega\big)$ consists of terms

$$d(d\omega) = \sum_{j=1}^{m}\sum_{i=1}^{m} \frac{\partial^2 f}{\partial x_j \partial x_i} \cdot dx_j \wedge \big(dx_i \wedge dx_I\big) \qquad \text{with } I \in \mathcal{E}_k \ .$$

The monomials $(dx_i \wedge dx_I)$ appearing in $d(f dx_i)$ might not be in standard form, but formula (39) is still valid, as explained in 3.13A.

Since $\dfrac{\partial^2 f}{\partial x_j \partial x_i} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$, anticommutativity of wedge products makes the terms

$$\frac{\partial^2 f}{\partial x_j \partial x_i}\,(dx_j \wedge dx_i \wedge dx_I)$$

and

$$\frac{\partial^2 f}{\partial x_i \partial x_j}\,(dx_i \wedge dx_j \wedge dx_I) = (-1)\frac{\partial^2 f}{\partial x_i \partial x_j}\,(dx_j \wedge dx_i \wedge dx_I)$$

in the double sum cancel in pairs for $i \neq j$, and when $j = i$ we have $dx_j \wedge dx_i = dx_i \wedge dx_i \equiv 0$ on U_α. Thus $d(d\omega) = 0$ on any chart. $\square$

The following observation is also useful in calculations involving exterior derivatives.

3.16 Lemma. *If* $f \in \mathcal{C}^\infty(M)$ *and* $\omega \in \bigwedge^k(M)$ *has the form* $\omega = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (dx_I)$ *in local coordinates, then the exterior derivative of the combined* k*-form* $f\omega$ *is described on the chart domain* U_α *by*

(41) $$d(f\omega) = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (df) \wedge (dx_I) + f(u) \cdot (dc_I) \wedge (dx_I)$$

in $\bigwedge^{k+1}(U)$.

Proof: The exterior derivative takes the form

$$\begin{aligned} d(f\omega) &= d\Big(\sum_{I \in \mathcal{E}_k} (f \cdot c_I)(u) \cdot (dx_I)\Big) \\ &= \sum_{I \in \mathcal{E}_k} d\big((f c_I) \cdot (dx_I)\big) \\ &= \sum_{I \in \mathcal{E}_k} d(f c_I) \wedge (dx_I) \qquad \text{(by (39))} \end{aligned}$$

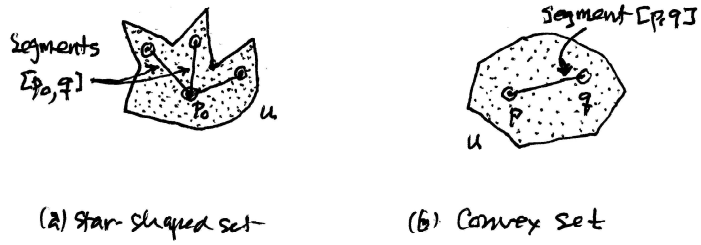

Figure 11.5. Definition of *convex* and *star-shaped* open sets in $\mathbb{R}^n$. Here $[p,q]$ is the straight line segment in $\mathbb{R}^n$ connecting p and q.

in local coordinates. Since d is a derivation on $\mathcal{C}^\infty(U_\alpha)$ we get (41). $\square$

Primitives of Rank-k Forms. We say that $\mu \in \bigwedge^k(M)$ is a **primitive** for ω in $\bigwedge^{k+1}(M)$ if $d_k\mu = \omega$ on M. The condition $d^2 \equiv 0$ imposes a necessary condition for ω to have a local primitive:

$$d\omega = d_{k+1}\omega \equiv 0 \quad \text{on } M\ .$$

In fact if $\omega = d\mu$ on some open set $U \subseteq M$ then $d\omega = d^2\mu = 0$. For rank-1 forms this identity is equivalent to the set of "consistency conditions" (25) mentioned earlier. It suffices to verify this on a typical chart, so suppose $d_1\omega = 0$ in $\bigwedge^2(U_\alpha)$ for $\omega = \sum_{i=1}^m c_i \cdot dx_i$ in $\bigwedge^1(U_\alpha)$. Then $dx_i \wedge dx_j = 0$ if $i = j$, so

$$\begin{aligned} d\omega &= \sum_{i=1}^m dc_i \wedge dx_i = \sum_i \sum_j \Big(\frac{\partial c_i}{\partial x_j}\Big) \cdot (dx_j \wedge dx_i) \\ &= \sum_{1 \le i < j \le m} \Big(\frac{\partial c_j}{\partial x_i} - \frac{\partial c_i}{\partial x_j}\Big) \cdot (dx_i \wedge dx_j) \end{aligned}$$

This sum over basis vectors in $\bigwedge^2(\mathrm{TM}_p)$ is zero at each $p \in U_\alpha$ if and only if (25) holds.

The higher rank analogs of these consistency conditions assert that if $\omega = d_{k-1}\mu$ for some $\mu \in \bigwedge^{k-1}(M)$, then on any chart the coefficients $c_I \in \mathcal{C}^\infty(U_\alpha)$ in $\omega = \sum_{I \in \mathcal{E}_k} c_I(u) \cdot (dx_I)$ must satisfy a system of partial differential equations embodied in the concise statement

$$d_k\omega \equiv 0 \text{ in } \textstyle\bigwedge^{k+1}(U_\alpha)\ .$$

In local coordinates, this system of differential equations becomes increasingly complicated for $k = 2, 3, \ldots$

3.17 Exercise. Write out the system of partial differential equations satisfied on a chart (x_α, U_α) by the coefficients of $\omega = \sum_{1 \le i<j \le m} c_{ij}(u) \cdot (dx_i \wedge dx_j) \in \bigwedge^2(U_\alpha)$ if $d_2\omega \equiv 0$ in $\bigwedge^3(U_\alpha)$. $\square$

The necessary condition $d\omega = 0$ actually insures that solutions of the equation $\omega = d\mu$ exist locally near any base point in U_α. There are two fundamental results about solutions of $d\mu = \omega$. The first asserts that *local* solutions exist if $d\omega \equiv 0$; it is a consequence of Poincare's Lemma, proved below, which also provides information about existence of *global* solutions.

3.18 Theorem. (Local Solutions). *Local primitives exist for $\omega \in \bigwedge^k(M)$ near a point $p \in M$ if and only if the necessary condition $d_k\omega \equiv 0$ in $\bigwedge^{k+1}(M)$ is satisfied.*

Once we know that $d\omega \equiv 0$ throughout M, the geometry of the manifold must be taken into account in seeking *global* solutions of $d\mu \equiv \omega$ on all of M. (There is also the problem of *constructing* the desired local and global solutions μ when they exist.)

A set $E \subseteq \mathbb{R}^n$ is **convex** if for any pair of points $\mathbf{p}, \mathbf{q} \in E$ the "line segment"

$$[\,\mathbf{p}, \mathbf{q}\,] = \{r\mathbf{p} + s\mathbf{q} : 0 \le r, s \le 1 \text{ and } r + s = 1\}$$

connecting $\mathbf{p}$ to $\mathbf{q}$ lies entirely within E; more generally, E is **star-shaped** if there is some $\mathbf{p} \in E$ from which every point $x \in E$ is "visible," with $[\,\mathbf{p}, \mathbf{x}] \subseteq E$ – see Figure 11.5.

3.19 Lemma (Poincare Lemma). *If E is a convex or star-shaped open set in $\mathbb{R}^n$ and $\omega \in \bigwedge^k(E)$ satisfies the condition $d_k\omega \equiv 0$ throughout E, then there is a global solution $\mu \in \bigwedge^{k-1}(E)$ to the equation $d_{k-1}\mu \equiv \omega$, so ω has a global primitive on E.*

Discussion: Given a $\mathcal{C}^\infty$ map $\phi : M \to N$ between manifolds there is a natural linear map $\delta\phi : \bigwedge^k(N) \to \bigwedge^k(M)$ that pulls back smooth k-forms on N to smooth k-forms on M; furthermore, $\delta\phi$ is a bijective linear isomorphism if ϕ is a diffeomorphism between M and N (invertible, with $\mathcal{C}^\infty$ inverse). Finally, $\delta\phi$ "intertwines" the exterior derivatives d^M and d^N, so that $\delta\phi\big(d^N(\omega)\big) = d^M\big(\delta\phi(\omega)\big)$ – i.e. the following diagram is commutative .

$$\begin{array}{ccc} \bigwedge^k(N) & \xrightarrow{(\delta\phi)} & \bigwedge^k(M) \\ d^N \downarrow & & \downarrow d^M \\ \bigwedge^{k+1}(N) & \xrightarrow{(\delta\phi)} & \bigwedge^{k+1}(M) \end{array}$$

We will not spell out the details here except to mention that this map $\delta\phi$ of smooth differential forms is a generalization of a natural linear map $(\delta\phi)_{\phi(p)} : \mathrm{TN}^*_{\phi(p)} \to \mathrm{TM}^*_p$ between cotangent spaces that transfers 1-forms in TN^*_q to 1-forms in TM^*_p. If $\phi(p) = q$, $(\delta\phi)_q$ is given by

$$\Big\langle (\delta\phi)_q\omega_q\,,\, X_p \Big\rangle = \Big\langle (\delta\phi)_{\phi(p)}\omega_{\phi(p)}\,,\, X_p \Big\rangle = \Big\langle \omega_{\phi(p)}\,,\, (d\phi)_p X_p \Big\rangle \tag{42}$$

for all $\omega_q \in \mathrm{TM}^*_q$ and $X_p \in \mathrm{TM}_p$. Thus $(\delta\phi)_{\phi(p)}$ is just the transpose

$$(\delta\phi)_{\phi(p)} = (d\phi)^{\mathrm{t}}_p : \mathrm{TN}^*_{\phi(p)} \to \mathrm{TM}^*_p$$

of the differential $(d\phi)_p : \mathrm{TM}_p \to \mathrm{TN}_{\phi(p)}$ discussed previously. Both $(\delta\phi)_{\phi(p)}$ and $(d\phi)_p$ are "natural maps," defined without reference to any particular system of local coordinates.

Transferring Problems from M to $\mathbb{R}^m$.
If (x_α, U_α) is a chart on a manifold M, the chart map $x_\alpha : U_\alpha \to V_\alpha = x_\alpha(U_\alpha)$ is a diffeomorphism between open sets in M and $N = \mathbb{R}^m$ (which is also a manifold). We may therefore use the chart map $\phi = x_\alpha$ to transfer questions about smooth differential forms on $U_\alpha \subseteq M$ to the corresponding problems on $V_\alpha \subseteq \mathbb{R}^m$. By commutativity of the preceding diagram, solutions of $d^M\mu = \omega$ on U_α correspond to solutions of $d^N\mu' = \omega'$ for the transferred forms $\mu' = \delta\phi(\mu)$ and $\omega' = \delta\phi(\omega)$. Likewise, the necessary conditions $d^M\omega = 0$ and $d^N\omega = 0$ match up under this correspondence.

Now observe that if $p \in M$ there are always (small) "star-shaped charts" (x_α, U_α) about p such that $V_\alpha = x_\alpha(U_\alpha)$ is a star-shaped neighborhood of $\mathbf{p} = x_\alpha(p)$ in $\mathbb{R}^m$ – in fact, any chart about p contains an open neighborhood U such that $p \in U \subseteq U_\alpha$ and $x_\alpha(U)$ is a (convex) open ball in $\mathbb{R}^n$. If we are lucky we can sometimes find quite large star-shaped charts on M, for instance if M is the punctured plane $\mathbb{R}^2 \sim (\mathbf{0})$, the "cut plane" $\mathbb{R}^2 \sim (-\infty, 0]$ is star-shaped with respect to any point p on the positive x-axis. Poincare's Lemma guarantees existence of global solutions for $d\mu = \omega$ on any star-shaped chart domain such that $d\omega \equiv 0$.

Elementary Proof of Poincare Lemma.[1] We may obviously assume that the "central point" $p \in E$ is the origin in $\mathbb{R}^n$. If $\omega = \sum_{I \in \mathcal{E}_k} w_I(u) \cdot (dx_I)_u$ on E we will define linear maps $I_k : \bigwedge^k(E) \to \bigwedge^{k-1}(E)$ for $k = 1, 2, \ldots$ such that

$$(43) \qquad d_{k-1} I_k(\omega) + I_{k+1} d_k(\omega) = \omega \qquad \text{on } E \text{ for } \omega \in \textstyle\bigwedge^k(E) \ .$$

If $d_k(\omega) = 0$ then $\mu = I_k(\omega) \in \bigwedge^{k-1}(E)$ is the desired solution of $d_{k-1}\mu = \omega$. Since I_k and d_k are linear operators we may restrict attention to weighted monomial k-forms

$$\omega = w_I(u) \cdot dx_I = w_I(u) \cdot (dx_{i_1} \wedge \ldots \wedge dx_{i_k}) \qquad \text{for } I \in \mathcal{E}_k,\ w_I \in \mathcal{C}^\infty(E) \ .$$

Since primitives of ω resemble antiderivatives of scalar-valued functions, it is not surprising to find that $I_k(\omega)$ can be constructed by taking integrals that involve the coefficients $w_I(u)$ and applying a basic result from Advanced Calculus.

Theorem. (Differentiating Under the Integral). *Let $B \subseteq \mathbb{R}^m$ be a closed rectangular set and $f : \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$ a function defined on some open set $U \supseteq B \times [a, b]$. If the partial derivatives $\partial f / \partial x_i(\mathbf{x}, y)$ exist and are continuous on U for $1 \leq i \leq m$, then $H(\mathbf{x}) = \int_a^b f(\mathbf{x}, t)\, dt$ has continuous partial derivatives*

$$\frac{\partial H}{\partial x_i}(\mathbf{x}) = \int_a^b \frac{\partial f}{\partial x_i}(\mathbf{x}, t)\, dt$$

on an open neighborhood of B for $1 \leq i \leq m$.

If the partial derivatives $D^\alpha_{\mathbf{x}} f(\mathbf{x}, t)$ in the $\mathbf{x}$-variables exist and are continuous on U for multi-indices of degree $|\alpha| = \alpha_1 + \ldots + \alpha_m \leq k$, then repeated differentiation under the integral shows that $H(\mathbf{x})$ is of class $\mathcal{C}^{(k)}$ on an open neighborhood of B, with

$$D^\alpha_{\mathbf{x}} H(\mathbf{x}) = \int_a^b D^\alpha_{\mathbf{x}} f(\mathbf{x}, t)\, dt$$

for $|\alpha| \leq k$.

For any weighted monomial $\omega = w_I(\mathbf{x}) \cdot dx_I$ with $I = (i_1 < \ldots < i_r) \in \mathcal{E}_r$, $r \geq 1$, we define the "rank-lowering" operator I_r, taking $I_r\omega(\mathbf{x}) \in \bigwedge^{r-1}(E)$ to be the sum of integrals

$$(44) \quad I_r\omega(\mathbf{x}) = \sum_{\ell=1}^{r} (-1)^{\ell-1} \Big(x_{i_\ell} \cdot \int_0^1 w_I\big(t\mathbf{x}\big)\, t^{r-1}\, dt \Big) \cdot (dx_{i_1} \wedge \ldots \wedge \widehat{dx_{i_\ell}} \wedge \ldots \wedge dx_{i_r}) \ ,$$

where the "hat" over dx_ℓ indicates an omitted factor. Hereafter we shall simplify notation by relabeling the variables $x_1, \ldots, x_m$ so $dx_I = dx_{i_1} \wedge \ldots \wedge dx_{i_r}$ becomes $dx_1 \wedge \ldots \wedge dx_r$

Taking $r = k$, differentiation under the integral shows that $I_k\omega(\mathbf{x})$ is a smooth $(k-1)$-form on E. To compute the term $d_{k-1} I_k \omega$ in (43) we find $D_{x_s}\{I_k\omega(\mathbf{x})\}$ for $1 \leq s \leq m$ by first differentiating under the integral to get

$$D_{x_s}\Big(x_\ell \cdot \int_0^1 w_I(t\mathbf{x})\, t^{k-1}\, dt \Big) = x_\ell \cdot \int_0^1 D_{x_s}\{w_I(t\mathbf{x})\}\, t^{k-1}\, dt + \delta_{s,\ell} \cdot \int_0^1 w_I(t\mathbf{x})\, t^{k-1}\, dt \ .$$

Here $\delta_{r,s}$ is the Kronecker delta symbol ($= 1$ if $r = s$ and zero otherwise). Noting that $dx_s \wedge dx_1 \wedge \ldots \wedge dx_k = 0$ if $s \in [1, k]$, the definition (39) of the exterior derivative d_{k-1}

[1] Adapted (with corrections) from M. Spivak, *Calculus on Manifolds*, Benjamin, 1965.

yields

$$
\begin{aligned}
d_{k-1}I_k\omega(\mathbf{x}) &= \sum_{\ell=1}^{k}\sum_{s=1}^{m} D_{x_s}\big(\cdots\big)\cdot(-1)^{\ell-1}dx_s\wedge(dx_1\wedge\ldots\wedge\widehat{dx_\ell}\wedge\ldots\wedge dx_k) \\
(45)\qquad &= \sum_{\ell=1}^{k}\sum_{s=1}^{m} D_{x_s}\big(\cdots\big)\cdot(-1)^{\ell-1}(-1)^{k-1}\,(dx_1\wedge\ldots\wedge\widehat{dx_\ell}\wedge\ldots\wedge dx_k)\wedge dx_s
\end{aligned}
$$

(shifting dx_s across $k-1$ vectors introduces a factor $(-1)^{k-1}$). Furthermore,

$$
\begin{aligned}
D_{x_s}\big(\cdots\big) &= D_{x_s}\Big(x_\ell\cdot\int_0^1 w_I(t\mathbf{x})\,t^{k-1}\,dt\Big) \\
(46)\qquad &= x_\ell\cdot\int_0^1 D_{x_s}w_I(t\mathbf{x})\,t^k\,dt+\delta_{\ell,s}\cdot\Big(\int_0^1 w_I(t\mathbf{x})\,t^{k-1}\,dt\Big)
\end{aligned}
$$

because $D_{x_s}\{w_I(t\mathbf{x})\}=t\cdot D_{x_s}w_I(t\mathbf{x})$. When (46) is substituted in (45) the first term in (46) yields the double sum

$$
\sum_{\ell=1}^{k}\sum_{s=1}^{m} x_\ell\cdot\Big(\int_0^1 D_{x_s}w_I(t\mathbf{x})\,t^k\,dt\Big)\cdot(-1)^{\ell-1}(-1)^{k-1}(dx_1\wedge\ldots\wedge\widehat{dx_\ell}\wedge\ldots\wedge dx_k\wedge dx_s)
$$

in which the nonzero terms are those with $s>k$ or $s=\ell$. Owing to the Kronecker delta in the second term of (46), substitution of this into (45) yields only k entries (those with $s=\ell$) and we get

$$
\begin{aligned}
&\sum_{\ell=1}^{k}\Big(\int_0^1 w_I(t\mathbf{x})\,t^{k-1}\,dt\Big)\cdot(-1)^{\ell-1}dx_\ell\wedge(dx_1\wedge\ldots\wedge\widehat{dx_\ell}\wedge\ldots\wedge dx_k) \\
&\qquad =\sum_{\ell=1}^{k}\Big(\int_0^1 w_I(t\mathbf{x})\,t^{k-1}\,dt\Big)\cdot dx_I \;=\; k\cdot\Big(\int_0^1 w_I(t\mathbf{x})\,t^{k-1}\,dt\Big)\,dx_I\ .
\end{aligned}
$$

As for the second term in (43), by our labeling convention we may write

$$
d_k\omega=d_k(w_I(\mathbf{x})\cdot dx_I)\;=\;\sum_{s=1}^{m}D_{x_s}w_I(\mathbf{x})\cdot dx_s\wedge dx_I\;=\;\sum_{s=1}^{m}D_{x_s}w_I(\mathbf{x})\cdot(-1)^k dx_I\wedge dx_s\ .
$$

In applying I_r when $r=k+1$ we must set the monomial "$dx_{i_1}\wedge\ldots\wedge dx_{i_{k+1}}$" appearing in (44) equal to "$dx_I\wedge dx_s$." Then by definition of I_{k+1} we get

$$
\begin{aligned}
I_{k+1}d_k(w_I\cdot dx_I)\;&=\;I_{k+1}\Big(\sum_{s=1}^{m}D_{x_s}w_I(\mathbf{x})\cdot(-1)^k(dx_1\wedge\ldots\wedge dx_k)\wedge dx_s\Big)= \\
&=\sum_{\ell=1}^{k+1}\sum_{s=1}^{m}\Big(x_\ell\cdot\int_0^1 D_{x_s}w_I(t\mathbf{x})\,t^{(k+1)-1}\,dt\Big)\cdot(-1)^k(-1)^{\ell-1}(dx_1\wedge\ldots\wedge\widehat{dx_\ell}\wedge\ldots\wedge dx_k\wedge dx_s) \\
&=\sum_{\ell=1}^{k}\sum_{s=1}^{m}\big(\cdots\big)+(\text{terms with }\ell=k+1)
\end{aligned}
$$

In the sum of terms with $\ell=k+1$ we have $dx_{k+1}=dx_s$; the monomial

$$
(dx_1\wedge\ldots\wedge\widehat{dx_\ell}\wedge\ldots\wedge dx_k\wedge dx_s)\;=\;(dx_1\wedge\ldots\wedge dx_k\wedge\widehat{dx_{k+1}})
$$

common to all these terms becomes $dx_1 \wedge \ldots \wedge dx_k = dx_I$, so this sum becomes

$$\sum_{s=1}^{m} \big(x_s \cdot \int_0^1 D_{x_s} w_I(t\mathbf{x})\, t^k\, dt\big) \cdot (-1)^k (-1)^k (dx_1 \wedge \ldots \wedge dx_k)$$
$$= \sum_{s=1}^{m} \big(\cdots\big) \cdot (dx_1 \wedge \ldots \wedge dx_k) \;=\; \sum_{s=1}^{m} \big(\cdots\big) \cdot dx_I$$

On the other hand, by the chain rule we also have

$$\begin{aligned} w_I(\mathbf{x}) &= \Big[w_I(t\mathbf{x})\, t^k \Big|_{t=0}^{1}\Big] \;=\; \int_0^1 \frac{d}{dt}\big\{w_I(t\mathbf{x})\, t^k\big\}\, dt \\ &= \Big(x_s \cdot \int_0^1 \sum_{s=1}^{m} D_{x_s} w_I(t\mathbf{x})\, t^k\, dt\Big) + k \cdot \Big(\int_0^1 w_I(t\mathbf{x})\, t^{k-1}\, dt\Big) \end{aligned}$$

Substituting this into the $\ell = k+1$ term we may rewrite

$$I_{k+1} d_k \omega = \sum_{\ell=1}^{k} \sum_{s=1}^{m} \big(\cdots\big) + w_I(\mathbf{x})\, dx_I - k \cdot \Big(\int_0^1 w_I(\mathbf{x})\, t^{k-1}\, dt\Big) \;.$$

When this is combined with $d_{k-1} I_k \omega$ the double sums cancel and so do the intergrals involving w_I, leaving only

$$I_{k+1} d_k \omega + d_{k-1} I_k \omega = \omega \;.$$

That completes the proof of the Poincare Lemma. $\square$

The identity (43) has a long history, which we can't spell out here. You might wonder how anyone deduced that an identity like (43) might hold, and how the definition (44) of the integral operators $I_k : \bigwedge^k(E) \to \bigwedge^{k-1}(E)$ was discovered.

XI.4 Div, Grad, Curl and All That.

We now consider what all this means on the Euclidean spaces $M = \mathbb{R}^2$ and $\mathbb{R}^3$, and its connection with the traditional vector operators $\nabla\circ =$ **div**, $\nabla\cdot =$ **grad**, and $\nabla\times =$ **curl** of Calculus. When $M = \mathbb{R}^3$ the sequence of exterior derivatives terminates

$$\textstyle\bigwedge^0(M) \xrightarrow{d_0} \bigwedge^1(M) \xrightarrow{d_1} \bigwedge^2(M) \xrightarrow{d_2} \bigwedge^3(M) \xrightarrow{d_3} (0)$$

because $\dim \bigwedge^k(\mathrm{TM}_p) = \binom{3}{k} = 0$ for $k > 3$.

Let $\{\mathbf{e}_i\}$ be the standard basis vectors in $V = \mathbb{R}^3$ and $\{\mathbf{e}_i^*\}$ the dual vectors in $V^* = \bigwedge^1(V)$. Fix a base point p in $M = \mathbb{R}^3$. If we interpret the $\mathbf{e}_i$ as tangent vectors at p, so $\mathbf{e}_i =$ the directional derivative $\big(\partial/\partial x_i|_p\big)$ in TM_p, then the dual basis vectors are $\mathbf{e}_i^* = (dx_i)_p$ in TM_p^*. We have already seen that

$$\text{Classical Gradient: } \nabla f = \sum_{i=1}^{3} D_{x_i} f(\mathbf{x}) \cdot \mathbf{e}_i$$

should be interpreted as a rank-1 differential form, the exterior derivative of $f \in \bigwedge^0(M) = \mathcal{C}^\infty(M)$,

$$\nabla f(p) = \sum_{i=1}^{3} \frac{\partial f}{\partial x_i}(p)\, \mathbf{e}_i^* = \sum_{i=1}^{3} \frac{\partial f}{\partial x_i}(p)\, (dx_i)_p = (df)_p \quad \text{ in } \mathrm{TM}_p^*$$

rather than as a tangent vector

$$\sum_{i=1}^{3} \frac{\partial f}{\partial x_i}(p)\, \mathbf{e}_i = \sum_{i=1}^{3} \frac{\partial f}{\partial x_i}(p) \Big(\frac{\partial}{\partial x_i}\Big|_p\Big) \quad \text{ in } \mathrm{TM}_p$$

As noted earlier, it is remarkable that the exterior derivative $df = \nabla f$ takes the same general form in all coordinate systems. For instance in spherical coordinates $y_\beta(\mathbf{x}) = (\rho, \theta, \phi)$ and Cartesian coordinates $x_\alpha(\mathbf{x}) = (x, y, z)$ we have

$$df = \frac{\partial f}{\partial x}\, dx + \ldots + \frac{\partial f}{\partial z}\, dz \;=\; \frac{\partial f}{\partial \rho}\, d\rho + \ldots + \frac{\partial f}{\partial \phi}\, d\phi$$

Now consider a classical "vector field" on $\mathbb{R}^3$ of the sort encountered in Calculus, which is usually written as

$$\mathbf{F}(\mathbf{x}) = F_1(\mathbf{x})\,\mathbf{i} + F_2(\mathbf{x})\,\mathbf{j} + F_3(\mathbf{x})\,\mathbf{k}$$

with smooth coefficients $F_i : \mathbb{R}^3 \to \mathbb{R}$. It is not clear how the basis vectors $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ in these classical narratives are to be interpreted – are they tangent vectors attached to each base point p? Cotangent vectors? Or as something else? The answer actually depends on the physical nature of the "field" being modeled. For instance we have seen that an electric field $\mathbf{E}(p)$, being $\nabla\phi$ for some scalar potential function, should be regarded as a field of cotangent vectors and not tangent vectors, etc.

Our interpretation of ∇f as a field of cotangent vectors suggests that, in this case at least, we might interpret the traditional basis vectors $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ as $\mathbf{i} = (dx)_p, \mathbf{j} = (dy)_p, \mathbf{k} = (dz)_p$ as we have for ∇f. On the other hand consider the operator $\mathbf{curl} = \nabla\times$ which sends a classical vector field $\mathbf{F}$ to a new vector field $\nabla \times \mathbf{F}$ given by

$$\begin{aligned} \nabla \times \mathbf{F} \;&=\; \mathrm{curl}(\mathbf{F}) \;=\; \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ F_1 & F_2 & F_3 \end{bmatrix} \\ &=\; \Big(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\Big)\,\mathbf{i} + \Big(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\Big)\,\mathbf{j} + \Big(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\Big)\,\mathbf{k} \qquad (47) \end{aligned}$$

Should we still interpret the symbols $\mathbf{i}, \mathbf{j}, \mathbf{k}$ the same way, as dual vectors in TM_p^*? Obviously not. The proper interpretation will depend on what type of tensor field $\mathbf{F}$ represents – a vector field in $\mathcal{D}^{(1,0)}(M)$, a field of cotangent vectors in $\mathcal{D}^{(0,1)}(M) = \bigwedge^1(M)$, or whatever. If $\mathbf{F}$ represents a smooth 1-form (say an electric field in space), it turns out that $\nabla \times \mathbf{F}$ should be interpreted as the exterior derivative $d\mathbf{F}$, making $\nabla \times \mathbf{F}$ a smooth 2-form in $\bigwedge^2(M)$, an antisymmetric rank-2 tensor field.

To see why this is the right interpretation write $\mathbf{F} = F_1\, dx_1 + F_2\, dx_2 + F_3\, dx_3$, regarding $\mathbf{F}$ as a 1-form in $\bigwedge^1(M)$. Its exterior derivative would then be

$$\begin{aligned} d\mathbf{F} \;&=\; \sum_{i=1}^{3} (dF_i) \wedge dx_i \;=\; \sum_{i=1}^{3} \Big(\sum_{j=1}^{3} \frac{\partial F_i}{\partial x_j}\, dx_j \Big) \wedge dx_i \\ &=\; \sum_{i \neq j} \frac{\partial F_i}{\partial x_j} \cdot dx_j \wedge dx_i \qquad (\text{since } dx_i \wedge dx_i = 0) \\ &=\; \sum_{1 \leq i < j \leq 3} \Big(\frac{\partial F_j}{\partial x_i} - \frac{\partial F_i}{\partial x_j} \Big) \cdot dx_i \wedge dx_j \qquad (\text{since } dx_j \wedge dx_i = -dx_i \wedge dx_j) \\ &=\; \Big(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \Big) \cdot dx_1 \wedge dx_2 + \Big(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3} \Big) \cdot dx_2 \wedge dx_3 + \Big(\frac{\partial F_3}{\partial x_1} - \frac{\partial F_1}{\partial x_3} \Big) \cdot dx_1 \wedge dx_3 \end{aligned}$$

We get a perfect match with the classical curl $\nabla \times \mathbf{F}$ if we interpret the "unit vectors" $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ in (44) as

$$\mathbf{i} = dx_2 \wedge dx_3 \qquad \mathbf{j} = dx_3 \wedge dx_1 = -dx_1 \wedge dx_3 \qquad \mathbf{k} = dx_1 \wedge dx_2$$

Note the cyclic order $dx_1 \to dx_2 \to dx_3 \to dx_1$ employed here.

Up to a $\pm$ sign these are the ordered basis monomials in $\bigwedge^2(M)$ when M is equipped with Cartesian coordinates. However, these "unit vectors" are no longer "vectors" of the same type as the unit vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ in the original field $F_1\,\mathbf{i} + \ldots + F_3\,\mathbf{k} \in \bigwedge^1(\mathrm{TM}_p)$ because they are rank-2 antisymmetric tensors in $\bigwedge^2(\mathrm{TM}_p)$. It is an accident that the tensor $\nabla \times \mathbf{F}(p)$ can be mistaken for the same kind of object as $\mathbf{F}(p)$. This happens because

$$\dim \bigwedge^1(\mathrm{TM}_p) = \binom{3}{1} = \binom{3}{2} = \dim \bigwedge^2(\mathrm{TM}_p) \;=\; 3$$

on any three-dimensional manifold.

Since vector spaces of the same dimension are "isomorphic" it is tempting to regard them as being "the same," but they are not, and there is no natural way to identify them. While it is often useful to introduce bases and identify both spaces $\bigwedge^1(\mathrm{TM}_p)$ and $\bigwedge^2(\mathrm{TM}_p)$ with coordinate space $\mathbb{R}^3$ in order to perform calculations, this completely obscures the very different nature of these tensors, which have different ranks and do not transform the same way when described in different systems of local coordinates. Moreover, this dodge cannot work in higher dimensions, for if physical space were four-dimensional we would have

$$\dim \bigwedge^1(\mathrm{TM}_p) = \binom{4}{1} \;=\; 4 \quad \text{while} \quad \dim \bigwedge^2(\mathrm{TM}_p) = \binom{4}{2} \;=\; 6$$

Furthermore, in vector Calculus on $\mathbb{R}^4$ a new "vector operator" would appear, the exterior derivative d_2 in

$$\bigwedge^0(\mathbb{R}^4) \xrightarrow{d_0} \bigwedge^1(\mathbb{R}^4) \xrightarrow{d_1} \bigwedge^2(\mathbb{R}^4) \xrightarrow{d_2} \bigwedge^3(\mathbb{R}^4) \xrightarrow{d_3} \bigwedge^4(\mathbb{R}^4) \xrightarrow{d_4} (0)$$

which is not encompassed in the three-dimensional theory of classical Calculus. In contrast, the theory of differential forms makes perfect sense for tensors of arbitrary ranks in all dimensions, and even makes sense on differentiable manifolds such as spheres $S^n \subseteq \mathbb{R}^{n+1}$, as well as various other manifolds that have no obvious realizations as smooth hypersurfaces embedded in Euclidean spaces.

Finally let us consider the divergence operator $\mathbf{div} = \nabla\circ$. In Calculus you were told that **div** acts on "vector fields" $\mathbf{F} = F_1\,\mathbf{i} + F_2\,\mathbf{j} + F_3\,\mathbf{k}$ to produce "scalar fields" via the formula

$$\text{(48)} \qquad \nabla \circ \mathbf{F} \;=\; \operatorname{div}\mathbf{F} \;=\; \mathrm{Trace}\begin{bmatrix} \partial F_1/\partial x_1 & \ldots & \partial F_1/\partial x_3 \\ \vdots & & \vdots \\ \partial F_3/\partial x_1 & \ldots & \partial F_3/\partial x_3 \end{bmatrix}$$

$$=\; \frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3} \qquad (sum of diagonal entries)$$

To make sense of this in terms of exterior derivatives let's interpret $\mathbf{F}$ as an antisymmetric rank-2 tensor field with $\mathbf{i}, \mathbf{j}, \mathbf{k}$ the basis vectors for the three-dimensional space $\bigwedge^2(\mathrm{TM}_p)$,

$$\mathbf{i} = dx_2 \wedge dx_3 \qquad \mathbf{j} = dx_3 \wedge dx_1 = -dx_1 \wedge dx_3 \qquad \mathbf{k} = dx_1 \wedge dx_2$$

Applying the exterior derivative $d_2 : \bigwedge^2(\mathbb{R}^3) \rightarrow \bigwedge^3(\mathbb{R}^3)$ to $\mathbf{F} = \sum_{I\in\mathcal{E}_2} F_I \, dx_I$ we get

$$\begin{aligned}
d\mathbf{F} &= \sum_{I\in\mathcal{E}_2} (dF_I) \wedge dx_I \\
&= \Big(\sum_{i=1}^{3} \frac{\partial F_1}{\partial x_i}\, dx_i\Big) \wedge (dx_2 \wedge dx_3) - \Big(\sum_{i=1}^{3} \frac{\partial F_2}{\partial x_i}\, dx_i\Big) \wedge (dx_1 \wedge dx_3) + \\
&\qquad + \Big(\sum_{i=1}^{3} \frac{\partial F_3}{\partial x_i}\, dx_i\Big) \wedge (dx_1 \wedge dx_2) \\
&= \Big(\frac{\partial F_1}{\partial x_1} + \frac{\partial F_2}{\partial x_2} + \frac{\partial F_3}{\partial x_3}\Big) \cdot dx_1 \wedge dx_2 \wedge dx_3 \qquad \text{in } \bigwedge^3(\mathbb{R}^3)
\end{aligned}$$

because $dx_i \wedge dx_j \wedge dx_k = 0$ if any two of the indices i, j, k are equal. As an example, magnetic fields $\mathbf{B}(\mathbf{x})$ are represented by rank-2 antisymmetric tensor fields. An important physical law asserts that $d\mathbf{B} = \nabla \circ \mathbf{B}$ is identically zero in any region of space ("*magnetic monopoles* do not exist").

Since $\dim \bigwedge^3(\mathrm{TM}_p) = 1$ on $M = \mathbb{R}^3$, all elements of $\bigwedge^3(\mathrm{TM}_p)$ are scalar multiples of the single basis vector $(dx_1 \wedge dx_2 \wedge dx_3)_p$, and therefore might easily be confused with actual scalars (which are elements of the one-dimensional space $\bigwedge^0(\mathrm{TM}_p) = \mathbb{R}$). But they are *not* scalars. In fact if we take a different Euclidean coordinate system on $M = \mathbb{R}^3$, say $(y_1, y_2, y_3) = y_\beta(\mathbf{x}) = (x_2, x_1, x_3)$ with *reversed orientation*, we find that

$$dy_I = dy_1 \wedge dy_2 \wedge dy_3 \;=\; dx_2 \wedge dx_1 \wedge dx_3 \;=\; -\, dx_1 \wedge dx_2 \wedge dx_3 \;=\; -dx_I$$

for all $I \in \mathcal{E}_3$, reversing the sign on the basis vector. That never happens for a true scalar field $f \in \bigwedge^0(\mathbb{R}^3) = \mathcal{C}^\infty(\mathbb{R}^3)$.

Summarizing these remarks, we have found that

- Gradient: $\mathbf{grad} = \nabla\cdot$ acting on scalar fields should be interpreted as the exterior derivative $d_0 : \bigwedge^0(\mathbb{R}^3) \rightarrow \bigwedge^1(\mathbb{R}^3)$. The result is a smooth 1-form.
- Curl: $\mathbf{curl} = \nabla\times$ acts on smooth 1-forms and should be interpreted as the exterior derivative $d_1 : \bigwedge^1(\mathbb{R}^3) \rightarrow \bigwedge^2(\mathbb{R}^3)$. The result is a smooth 2-form.
- Divergence: $\mathbf{div} = \nabla\circ$ acts on smooth 2-forms and should be interpreted as the exterior derivative $d_2 : \bigwedge^2(\mathbb{R}^3) \rightarrow \bigwedge^3(\mathbb{R}^3)$. The resulting smooth 3-form can be written as $\phi(\mathbf{x})\, dx_1 \wedge dx_2 \wedge dx_3$ for some $\mathcal{C}^\infty$ scalar function $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$.

At the next level we have $d_3 : \bigwedge^3(\mathbb{R}^3) \rightarrow \bigwedge^4(\mathbb{R}^3) = (0)$. Interpreting the classical vector operators in this manner we immediately obtain two classic vector identities

$$\nabla \times (\nabla f) = d_1 \circ d_0(f) = 0 \qquad \text{and} \qquad \nabla \circ (\nabla \times \mathbf{F}) = d_2 \circ d_1(\mathbf{F}) = 0 \tag{49}$$

from the fundamental property $d^2 = d_{k+1} \circ d_k = 0$ of exterior derivatives.

The two-dimensional case $M = \mathbb{R}^2$ is similar, but simpler.

4.1 Example. (The 2-Dimensional Case). In $M = \mathbb{R}^2$ there are just two nontrivial vector operators

$$\bigwedge^0(\mathbb{R}^2) \xrightarrow{d_0} \bigwedge^1(\mathbb{R}^2) \xrightarrow{d_1} \bigwedge^2(\mathbb{R}^2) \xrightarrow{d_2} (0)$$

because $\bigwedge^k(\mathbb{R}^2) = (0)$ for $k \geq 3$.

- $\mathbf{grad}$ maps $f \in \bigwedge^0(\mathbb{R}^2) = \mathcal{C}^\infty(\mathbb{R}^2)$ to

$$\nabla f = df = \frac{\partial f}{\partial x_1}\, dx_1 + \frac{\partial f}{\partial x_2}\, dx_2 \qquad \text{in } \bigwedge^1(\mathbb{R}^2)$$

- **curl** maps $\omega = F_1\,dx_1 + F_2\,dx_2$ in $\bigwedge^1(\mathbb{R}^2)$ to

$$\begin{aligned} d\omega &= \Big(\frac{\partial F_1}{\partial x_1}\,dx_1 + \frac{\partial F_1}{\partial x_2}\,dx_2\Big)\wedge dx_1 + \Big(\frac{\partial F_2}{\partial x_1}\,dx_1 + \frac{\partial F_2}{\partial x_2}\,dx_2\Big)\wedge dx_2 \\ &= \Big(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\Big)\cdot dx_1\wedge dx_2 \qquad \text{in } \bigwedge^2(\mathbb{R}^2) \end{aligned}$$

There is no two-dimensional analog of the divergence operator **div** since $d\omega = 0$ for every 2-form on $\mathbb{R}^2$. (Likewise on any 2-dimensional differentiable manifold.) $\square$

It is now clear that the d-operators on the manifold $M = \mathbb{R}^m$

$$\bigwedge^0(M) \xrightarrow{d_0} \bigwedge^1(M) \xrightarrow{d_1} \bigwedge^2(M) \xrightarrow{d_2} \ \dots \ \xrightarrow{d_{m-1}} \bigwedge^m(M) \xrightarrow{d_m} \bigwedge^{m+1}(M) = (0) \ \ ,$$

with $d\circ d = 0$ at every stage, are the natural higher-dimensional analogs of **div**, **grad**, **curl** in three dimensions.

In dealing with three dimensional space physicists and mathematicians have adopted notation devised (circa 1890) by J. Willard Gibbs, who exploited the coincidence

$$\dim \bigwedge^0 = \dim \bigwedge^3 = 1 \qquad \text{and} \qquad \dim \bigwedge^1 = \dim \bigwedge^2 = 3$$

to identify these spaces with the usual coordinate spaces $\mathbb{R}^1$ and $\mathbb{R}^3$. He also exploited the fact that the manifold structure of $M = \mathbb{R}^3$ is determined by a *single chart* $x_\alpha(\mathbf{x}) = (x_1, x_2, x_3)$. This provides us with globally defined smooth k-forms

For $k = 1$: $\quad dx_1 \qquad dx_2 \qquad dx_3$

For $k = 2$: $\quad dx_1\wedge\ dx_2 \qquad dx_2\wedge dx_3 \qquad dx_3\wedge dx_1 = -\,dx_1\wedge dx_3$

For $k = 3$: $\quad dx_1\wedge dx_2\wedge dx_3$

that determine correlated basis vectors at all base points. This allows us to compare the spaces $\bigwedge^k(\mathrm{TM}_p)$ and $\bigwedge^k(\mathrm{TM}_q)$ at different base points. Thus we can in a uniform way identify spaces having the same dimension: $\bigwedge^0(\mathrm{TM}_p) \cong \bigwedge^3(\mathrm{TM}_p) \cong \mathbb{R}$ and $\bigwedge^1(\mathrm{TM}_p) \cong \bigwedge^2(\mathrm{TM}_p) \cong \mathbb{R}^3$, for all base points. This is really convenient in doing calculations, but obscures the distinction between various kinds of "vectors" in $\mathbb{R}^2$ and "scalars" in $\mathbb{R}$. Most physicists and mathematicians voted for convenience, although most were well aware that these identifications are only a shorthand description of what is actually going on.

Differential Forms and the Cross-Product $\mathbf{a}\times\mathbf{b}$ in $\mathbb{R}^3$.
Gibbs also introduced the **cross product** $\mathbf{v}\times\mathbf{w}$ of vectors in $\mathbb{R}^3$, as a simplified proxy for the wedge product of cotangent vectors. If we write the basis vectors in $\bigwedge^1(\mathrm{TM}_p)$ as

$$\mathbf{i} = (dx_1)_p \qquad \mathbf{j} = (dx_2)_p \qquad \mathbf{k} = (dx_3)_p$$

and those in $\bigwedge^2(\mathrm{TM}_p)$ as

$$(50)\qquad \begin{aligned} \mathbf{i}' &= \ \mathbf{j}\wedge\mathbf{k} \ = \ (dx_2\wedge dx_3)_p \\ \mathbf{j}' &= \ \mathbf{k}\wedge\mathbf{i} \ = \ (dx_3\wedge dx_1)_p = -\,(dx_1\wedge dx_3)_p \\ \mathbf{k}' &= \ \mathbf{i}\wedge\mathbf{j} \ = \ (dx_1\wedge dx_2)_p \ , \end{aligned}$$

and if we now identify basis vectors $\mathbf{i}'\sim\mathbf{i}$, $\mathbf{j}'\sim\mathbf{j}$, $\mathbf{k}'\sim\mathbf{k}$ in these three-dimensional spaces, then the identities (47) take the form

$$(51)\qquad \mathbf{i}\wedge\mathbf{j} = \mathbf{k}'\sim\mathbf{k} \qquad \mathbf{j}\wedge\mathbf{k} = \mathbf{i}'\sim\mathbf{i} \qquad \mathbf{k}\wedge\mathbf{i} = \mathbf{j}'\sim\mathbf{j}$$

These should look familiar since the Calculus cross product of vectors $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $\mathbf{w} = w_1\mathbf{i} + w_2\mathbf{j} + w_3\mathbf{k}$ in $\mathbb{R}^3$ is defined to be

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3) \end{bmatrix} \\ &= (v_2w_3 - v_3w_2)\,\mathbf{i} + (v_3w_1 - v_1w_3)\,\mathbf{j} + (v_1w_2 - v_2w_1)\,\mathbf{k} \end{aligned} \tag{52}$$

In particular,

$$\mathbf{i} \times \mathbf{j} = \mathbf{k} \qquad \mathbf{j} \times \mathbf{k} = \mathbf{i} \qquad \mathbf{k} \times \mathbf{i} = \mathbf{j} \tag{53}$$

Under these identifications the classical cross product becomes the wedge product. If we regard $\mathbf{v}, \mathbf{w}$ as cotangent vectors in $\bigwedge^1(\mathrm{TM}_p)$, then $\mathbf{v} \times \mathbf{w}$ is the 2-form $\mathbf{v} \wedge \mathbf{w}$ in $\bigwedge^2(\mathrm{TM}_p)$, which we then identify with a vector back in $\bigwedge^1(\mathrm{TM}_p) \cong \bigwedge^2(\mathrm{TM}_p)$ to get the classical cross product $\mathbf{v} \times \mathbf{w}$ in (52). Antisymmetry of the wedge product yields the familiar Calculus identities

$$\mathbf{w} \times \mathbf{v} = -\,\mathbf{v} \times \mathbf{w} \qquad\qquad \mathbf{v} \times \mathbf{v} = \mathbf{0} \text{ for all } \mathbf{v}$$

which also follow from the Calculus-style definition (52). But beware: xitThe identifications that underlie this interpretation of $\mathbf{v} \times \mathbf{w}$ are only possible in three dimensional space, while the wedge product on $\bigwedge^k(\mathbb{R}^n)$ makes sense in *all* dimensions.

4.2 Exercise. For $k = 2, 3, 4$ compute components of the wedge product $\mathbf{v} \wedge \mathbf{w}$ of $\mathbf{v} = \sum_{i=1}^{l} v_i\mathbf{e}_i$ and $\mathbf{w} = \sum_{j=1}^{k} w_j\mathbf{e}_j$, interpreting $\mathbf{e}_i = (dx_i)$ in $\bigwedge^2(\mathbb{R}^k)$. $\square$

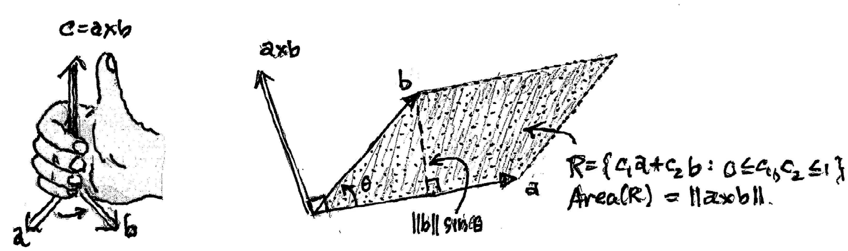

Figure 11.6. Geometric interpretation of the cross product $\mathbf{a} \times \mathbf{b}$ of two vectors in $\mathbb{R}^3$. The product is always perpendicular to the plane in $\mathbb{R}^3$ determined by the two vectors, and the length of the cross product is the area of the parallelogram R determined by the two vectors, $\|\mathbf{a} \times \mathbf{b}\| = \mathrm{Area}(R)$.

The geometric interpretation of the cross product $\mathbf{a} \times \mathbf{b}$ in $\mathbb{R}^3$ is shown in Figure 11.6. It is always perpendicular to the plane spanned by the two vectors and its length $\|\mathbf{a} \times \mathbf{b}\|$ is the area of the parallelogram R determined by $\mathbf{a}$ and $\mathbf{b}$. The meaning of the **vector triple product** $(\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c}$ of vectors in $\mathbb{R}^3$ is discussed in Exercises 4.3 - 4.4. Its numerical value (with sign) is the *oriented 3-dimensional volume* $\mathrm{Vol}(P)$ of the parallelopiped P whose edges are the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$. The signed volume of P is $(+)$ if the vectors form a "right-handed system" – i.e. if you wrap your fingers around the axis perpendicular to $\mathbf{a}$ and $\mathbf{b}$ with your index finger pointing in the same direction as $\mathbf{a}$ and middle finger aligned with $\mathbf{b}$, your thumb should be pointing in the same direction as $\mathbf{c}$; otherwise the signed volume is $(-)$.

4.3 Exercise. Prove that

(a) If $\mathbf{a}$ and $\mathbf{b}$ are nonzero then $\mathbf{a} \times \mathbf{b} = 0 \Leftrightarrow$ the vectors are collinear. Furthermore,

$$(\mathbf{a} \times \mathbf{b}) \bullet \mathbf{a} = 0 \quad \text{and} \quad (\mathbf{a} \times \mathbf{b}) \bullet \mathbf{b} = 0 \ ,$$

so $\mathbf{a} \times \mathbf{b}$ is orthogonal to a every vector in the plane $\mathbb{R}\mathbf{a} + \mathbb{R}\mathbf{b}$.

(b) $\mathbf{a} \bullet (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c}$

(c) $\|\mathbf{a} \times \mathbf{b}\|^2 = \|\mathbf{a}\|^2\|\mathbf{b}\|^2 \cdot (1 - \cos^2\theta) = \|\mathbf{a}\|^2\|\mathbf{b}\|^2 \cdot \sin^2(\theta)$ for the angle θ shown in Figure 11.6.

Hint: Show that the identity in (c) is equivalent to $\|\mathbf{a} \times \mathbf{b}\|^2 + |\mathbf{a} \bullet \mathbf{b}|^2$, which can be verified directly. $\square$

4.4 Example. If $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are in $\mathbb{R}^3$, verify that the **vector triple product** $(\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c}$, in which $(\bullet)$ is the usual inner product on $\mathbb{R}^3$, is just

$$(\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c} \;=\; \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \ ,$$

Explain why this is precisely the (oriented) volume of the parallelopiped P determined by the three vectors. (This will be zero (P is degenerate) $\Leftrightarrow$ the vectors are linearly dependent.) $\square$

4.5 Example. The triple product $\lambda = (\mathbf{a} \times \mathbf{b}) \bullet \mathbf{c}$ of vectors in $\mathbb{R}^3$ is a scalar. The space $\bigwedge^3(\mathbb{R}^3)$ is one-dimensional. If we interpret vectors as 1-forms in $\bigwedge^1(\mathbb{R}^3)$ as in Exercise 4.4, prove that

$$\mathbf{a} \wedge \mathbf{b} \wedge \mathbf{c} = \lambda \cdot (dx_1) \wedge (dx_2) \wedge (dx_3) \quad \square$$