Math 412: Advanced Linear Algebra
Lecture Notes

Lior Silberman

Contents

Introduction

Lior Silberman, `lior@Math.UBC.CA`, `https://www.math.ubc.ca/~lior/`
Office: Math Annex 1112
Phone: 604-827-3031

For administrative details see the syllabus.

0.1. Goals and course plan (Lecture 1)

Math	Metamath	Skills
Reinforce basics • Vector space, subspace • Linear independence, basis • Linear map • Eigenvalue, eigenvector	Abstraction	Hard problems Multiple ideas
New ideas • Direct sum, product • $\mathrm{Hom}(U,V)$ and duality • Quotients • Tensor products	Constructions Universal properties	Abstract examples
Structure theory for linear maps • Matrix decompositions • LU, $LL^{\dagger}$ and Computation • Minimal poly, Cayley–Hamilton • Jordan canonical form	Canonical forms Algorithms	
Analysis • Norms • Holomorphic calculus (e^{tX})	Combining fields	

0.2. Review

0.2.1. Basic definitions. We want to give ourselves the *freedom* to have scalars other than real or complex.

DEFINITION 1 (Fields). A *field* is a quintuple $(F,0,1,+,\cdot)$ such that $(F,0,+)$ and $(F\setminus\{0\},1,\cdot)$ are abelian groups, and the *distributive law* $\forall x,y,z\in F : x(y+z)=xy+xz$ holds.

LEMMA 2. *In a field* $x\cdot 0=0$.

PROOF. $x\cdot 0=x\cdot(0+0)=x\cdot 0+x\cdot 0$. Adding $-(x\cdot 0)$ and using the axioms gives $0=x\cdot 0$. □

COROLLARY 3. *The associative and commutative laws for multiplication also holds for products involving* 0.

EXERCISE 4. Generalized associative and commutative laws holds for both addition and multiplication.

EXAMPLE 5. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$. $\mathbb{F}_2$ (via addition and multiplication tables; ex: show this is a field), $\mathbb{F}_p$.

EXERCISE 6. Every finite field has p^r elements for some prime p and some integer $r\geq 1$. Fact: there is one such field for every prime power.

DEFINITION 7. A *vector space* over a field F is a quadruple $(V,\underline{0},+,\cdot)$ where $(V,\underline{0},+)$ is an abelian group, and $\cdot : F\times V\to V$ is a map such that:

(1) $1_F\underline{v}=\underline{v}$.
(2) $\alpha(\beta\underline{v})=(\alpha\beta)\underline{v}$.
(3) $(\alpha+\beta)(\underline{v}+\underline{w})=\alpha\underline{v}+\beta\underline{v}+\alpha\underline{w}+\beta\underline{w}$.

LEMMA 8. $0_F\cdot\underline{v}=\underline{0}$, $a\cdot\underline{0}=\underline{0}$.

PROOF. See Lemma 2. □

EXAMPLE 9. $\{\underline{0}\}$, $\mathbb{R}^n$, F^X.

EXERCISE 10 (PS1). Check vector space axioms for F^X

0.2.2. Linear maps and subspaces.

0.2.3. Bases and dimension. Fix a vector space V.

DEFINITION 11. Let $S\subset V$.

- $\underline{v}\in V$ *depends on* S if there are $\{\underline{v}_i\}_{i=1}^r\subset S$ and $\{a_i\}_{i=1}^r\subset F$ such that $\underline{v}=\sum_{i=1}^r a_i\underline{v}_i$ [empty sum is $\underline{0}$]
- Write $\mathrm{Span}_F(S)\subset V$ for the set of vectors that depend on S.
- Call S *linearly dependent* if some $\underline{v}\in S$ depends on $S\setminus\{\underline{v}\}$, equivalently if there are distinct $\{\underline{v}_i\}_{i=1}^r\subset S$ and $\{a_i\}_{i=1}^r\subset F$ not all zero such that $\sum_{i=1}^r a_i\underline{v}_i=\underline{0}$.
- Call S *linearly independent* if it is not linearly dependent.

AXIOM 12 (Axiom of choice). *Every vector space has a basis.*

0.3. Euler's Theorem

Let $G = (V, E)$ be a connected planar graph. A *face* of G is a finite connected component of $\mathbb{R}^2 \setminus G$.

THEOREM 13 (Euler). $v - e + f = 1$.

PROOF. Arbitrarily orient the edges. Let $\partial_E : \mathbb{R}^E \to \mathbb{R}^V$ be defined by $f((u,v)) = 1_v - 1_u$, $\partial_F : \mathbb{R}^F \to \mathbb{R}^E$ be given by the sum of edges around the face.

LEMMA 14. *∂_F is injective.*

PROOF. Faces containing boundary edges are independent. Remove them and repeat. □

LEMMA 15. $\operatorname{Ker}\partial_E = \operatorname{Im}\partial_F$.

PROOF. Suppose a combo of edges is in the kernel. Following a sequence with non-zero coefficients gives a closed loop, which can be expressed as a sum of faces. Now subtract a multiple to reduce the number of edges with non-zero coefficients. □

LEMMA 16. $\operatorname{Im}(\partial_E)$ *is the the set of functions with total weight zero.*

PROOF. Clearly the image is contained there. Conversely, given f of total weight zero move the weight to a single vertex using elements of the image. [remark: quotient vector spaces] □

Now $\dim \mathbb{R}^E = \dim \operatorname{Ker}\partial_E + \dim \operatorname{Im}\partial_E = \dim \operatorname{Im}\partial_F + \dim \operatorname{Im}\partial_E$ so

$$e = f + (v - 1) \, .$$

□

REMARK 17. Using $\mathbb{F}_2$ coefficients is even simpler.

CHAPTER 1

Constructions

1.1. Overall plan

(1) Direct sums
   (a) External direct sum of two spaces
   (b) Internal direct sum of two spaces
   (c) Abstract direct sum of two or more spaces (through internal direct sum, not universal property)
(2) Quotients
   (a) Definition
   (b) Examples
   (c) Universal property in PS
(3) Duality
   (a) Hom space
   (b) Dual space
   (c) Dual basis
   (d) Dual of infinite direct sum leads to notion of direct product
   (e) Pairings and bilinear forms (but $\left(F^{\oplus\mathbb{N}}\right)' \simeq F^{\mathbb{N}}$ in PS3)
   (f) dual map defined (properties in PS3)
(4) Tensor products
   (a) Bilinear forms and maps
   (b) The tensor product
   (c) Universal properties (for direct sum, product, quotient, tensor product)
   (d) Symmetric and antisymmetric tensors

1.2. Direct sum, direct product (Lectures 2-4)

Fix a field F.

1.2.1. Simplest case (Lecture 2).

CONSTRUCTION 18 (External direct sum). *Let U,V be vector spaces. Their* direct sum*, denoted $U \oplus V$, is the vector space whose underlying set is $U \times V$, with coordinate-wise addition and scalar multiplication.*

LEMMA 19. *This really is a vector space.*

REMARK 20. The Lemma serves to review the definition of vector space.

PROOF. Every property follows from the respective properties of U,V. □

REMARK 21. Direct products of groups are discussed in 322.

LEMMA 22. $\dim_F(U \oplus V) = \dim_F U + \dim_F V$.

REMARK 23. This Lemma serves to review the notion of basis.

PROOF. Let B_U, B_V be bases of U, V respectively. Then $\{(\underline{u}, \underline{0}_V)\}_{\underline{u} \in B_U} \sqcup \{(\underline{0}_U, \underline{v})\}_{\underline{v} \in B_V}$ is a basis of $U \oplus V$. □

EXAMPLE 24. $\mathbb{R}^n \oplus \mathbb{R}^m \simeq \mathbb{R}^{n+m}$.

1.2.2. Internal sum and direct sum (Lecture 3). A key situation is when U, V are subspaces of an "ambient" vector space W.

LEMMA 25. *Let W be a vector space, $U, V \subset W$. Then* $\mathrm{Span}_F(U \cup V) = \{\underline{u} + \underline{v} \mid \underline{u} \in U, \underline{v} \in V\}$.

PROOF. RHS contained in the span by definition. It is a subspace (non-empty, closed under addition and scalar multiplication) which contains U, V hence contains the span. □

DEFINITION 26. The space in the previous lemma is called the *sum* of U, V and denoted $U + V$.

LEMMA 27. *Let $U, V \subset W$. There is a unique homomorphism $U \oplus V \to U + V$ which is the identity on U, V.*

PROOF. Define $f((\underline{u}, \underline{v})) = \underline{u} + \underline{v}$. Check that this is a linear map. □

PROPOSITION 28 (Dimension of sums). $\dim_F(U + V) = \dim_F U + \dim_F V - \dim_F(U \cap V)$.

PROOF. Consider the map f of Lemma 27. It is surjective by Lemma 25. Moreover $\mathrm{Ker} f = \{(\underline{u}, \underline{v}) \in U \oplus V \mid \underline{u} + \underline{v} = \underline{0}_W\}$, that is

$$\mathrm{Ker} f = \{(\underline{w}, -\underline{w}) \mid \underline{w} \in U \cap V\} \simeq U \cap V .$$

Since $\dim_F \mathrm{Ker} f + \dim_F \mathrm{Im} f = \dim(U \oplus V)$ the claim now follows from Lemma 22. □

REMARK 29. This was a review of that formula. Alternative proof by starting from a basis of $U \cap V$ and extending to bases of U, V, which is basically revisiting the proof of the formula.

DEFINITION 30 (Internal direct sum). We say the sum is *direct* if f is an isomorphism.

THEOREM 31. *For subspaces $U, V \subset W$ TFAE*

(1) *The sum $U + V$ is direct and equals W;*
(2) *$U + V = W$ and $U \cap V = \{\underline{0}\}$*
(3) *Every vector $\underline{w} \in W$ can be uniquely written in the form $\underline{w} = \underline{u} + vv$.*

PROOF. (1) $\Rightarrow$ (2): $U + V = W$ by assumption, $U \cap V = \mathrm{Ker} f$.
(2) $\Rightarrow$ (3): the first assumption gives existence, the second uniqueness.
(3) $\Rightarrow$ (1): existence says f is surjective, uniqueness says f is injective. □

1.2.3. Finite direct sums (Lecture 4). Three possible notions: $(U \oplus V) \oplus W$, $U \oplus (V \oplus W)$, vector space structure on $U \times V \times W$. These are *all the same*. Not just isomorphic (that is, not just same dimension), but also isomorphic when considering the extra structure of the copies of U, V, W. How do we express this?

DEFINITION 32. W is the *internal direct sum* of its subspaces $\{V_i\}_{i \in I}$ if it spanned by them and each vector has a unique representation as a sum of elements of V_i (either as a finite sum of non-zero vectors or as a zero-extended sum).

REMARK 33. This generalizes the notion of "linear independence" from vectors to subspaces.

LEMMA 34. *Each of the three candidates contains an embedded copy of U, V, W and is the internal direct sum of the three images.*

PROOF. Easy. $\square$

PROPOSITION 35. *Let A, B each be the internal direct sum of embedded copies of U, V, W. Then there is a unique isomorphism $A \to B$ respecting this structure.*

PROOF. Build up the isomorphism from the pieces. $\square$

REMARK 36. (1) Proof only used result of Lemma, not specific structure; but (2) proof implicitly relies on isomorphism to $U \times V \times W$; (3) We used the fact that a map can be defined using values on copies of U, V, W (4) Exactly same proof as the facts that a function on 3d space can be defined on bases, and that that all 3d spaces are isomorphic.

- Dimension of $\oplus_i V_i$ by induction.

DEFINITION 37. Abstract arbitrary direct sum.

- Block diagonality (but it's also a practice problem in PS2)

REMARK 38. Infinite direct sums are a supplement to PS2.

1.3. Quotients (Lecture 5)

Recall that for a group G and a normal subgroup N, we can endow the quotient G/N with group structure $(gN)(hN) = (gh)N$.

- This is well-defined, gives group.
- Have quotient map $q \colon G \to G/N$ given by $g \mapsto gN$.
- Homomorphism theorem: any $f \colon G \to H$ factors as $G \to G/\operatorname{Ker}(f)$ follows by isomorphism.
- If $N < M < G$ with both N, M normal then $q(M) \simeq M/N$ is normal in G/N and $(G/N)/(M/N) \simeq (G/M)$.

Now do the same for vector spaces.

LEMMA 39. *Let V be a vector space, W a subspace. Let $\pi \colon V \to V/W$ be the quotient as abelian groups. Then there is a unique vector space structure on V/W making π a surjective linear map.*

PROOF. We must set $\alpha(\underline{v} + W) = \alpha\underline{v} + W$. This is well-defined and gives the isomorphism. Use quotient map to verify vector space axioms. $\square$

EXAMPLE 40. $V = C^1(0,1)$, space of continuously differentiable functions. Then $\frac{d}{dx} \colon V \to C(0,1)$ vanishes on $\mathbb{R}\mathbb{1}$ and hence induces a map $\frac{d}{dx} \colon \left(C^1(0,1)/\mathbb{R}\mathbb{1}\right) \to C(0,1)$. Note that the inverse of this map is what we call "indefinite integral" – whose images is exactly an equivalence class "function$+c$".

EXAMPLE 41. The definite integral is a linear function which vanishes on functions which are non-zero at countably many points. More generally, integration works on functions modulu functions which are zero a.e.

FACT 42. *The properties above persist for vector spaces.*

- How to use quotients: "kill off" part of the vector spaces that is irrelevant (linear maps vanish there) or already understood.

REMARK 43. Bijection of subspaces of quotietns and universal property are a problem in PS2

1.4. Hom spaces and duality (Lectures 6-8)

1.4.1. Hom spaces (Lecture 5 continued or start of lecture 6).

DEFINITION 44. $\mathrm{Hom}_F(U,V)$ will denote the space of F-linear maps $U \to V$.

LEMMA 45. $\mathrm{Hom}_F(U,V) \subset V^U$ *is a subspace, hence a vector space.*

DEFINITION 46. $V' = \mathrm{Hom}_F(V,F)$ is called the *dual space*.

Motivation 1: in PDE. Want solutions in some function space V. Use that V' is much bigger to find solutions in V', then show they are represented by functions.

1.4.2. The dual space, finite dimensions.

NOTE 47. In lecture ignore infinite dimensions (but make statements which are correct in general).

CONSTRUCTION 48 (Dual basis). *Let $B = \{\underline{b}_i\}_{i\in I} \subset V$ be a basis. Write $\underline{v} \in V$ uniquely as $\underline{v} = \sum_{i\in I} a_i \underline{b}_i$ (almost all $a_i = 0$) and set $\varphi_i(\underline{v}) = a_i$.*

LEMMA 49. *These are linear functionals.*

PROOF. Represent $\alpha\underline{v} + \underline{v}'$ in the basis. □

EXAMPLE 50. $V = F^n$ with standard basis, get $\varphi_i(\underline{x}) = x_i$. Note every functional has the form $\varphi(\underline{x}) = \sum_{i=1}^n \varphi(\underline{e}_i)\varphi_i(\underline{x})$.

REMARK 51. Alternative construction: φ_i is the unique linear map to F satisfying $\varphi_i(\underline{b}_j) = \delta_{i,j}$.

LEMMA 52. *The dual basis is linearly independent. It is spanning iff* $\dim_F V < \infty$.

PROOF. Evaluate a linear combination at $\underline{b}_j$.
If V is finite-dimensional, enumerate the basis as $\{\underline{b}_i\}_{i=1}^n$. Then for any $\varphi \in V'$ and any $\underline{v} \in V$ write $\underline{v} = \sum_i a_i \underline{b}_i$ and then

$$\varphi(\underline{v}) = \sum_i a_i \varphi(\underline{b}_i) = \sum_i (\varphi(\underline{b}_i))\, \varphi_i(\underline{v})$$

so

$$\varphi = \sum_i (\varphi(\underline{b}_i))\, \varphi_i \in \mathrm{Span}_F \{\varphi_i\} .$$

In the infinite-dimensional case let $\phi = \sum_{i\in I} \varphi_i$. Then ϕ is a well-defined linear functional which depends on every coordinate hence not in the span of the $\{\varphi_i\}$. □

REMARK 53. This isomorphism $V \to V'$ is not canonical: the functional φ_i depends on the *whole basis B and not only on $\underline{b}_i$, and the dual basis transforms differently from the original basis under change-of-basis.*

The argument above used evaluation – let's investigate that more.

PROPOSITION 54 (Double dual). *Given* $\underline{v} \in V$ *consider the* evaluation map $e_{\underline{v}} \colon V' \to F$ *given by* $e_{\underline{v}}(\varphi) = \varphi(\underline{v})$. *Then* $\underline{v} \mapsto e_{\underline{v}}$ *is a linear injection* $V \hookrightarrow V''$, *an isomorphism iff* V *is finite-dimensional.*

PROOF. The vector space structure on V' (and on F^V in general) is such that $e_{\underline{v}}$ is linear. That the map $\underline{v} \mapsto e_{\underline{v}}$ is linear follows from the linearity of the elements of V'. For injectivity let $\underline{v} \in V$ be non-zero. Extending $\underline{v}$ to a basis, let $\varphi_{\underline{v}}$ be the element of the dual such that $\varphi_{\underline{v}}(\underline{v}) = 1$. Then $e_{\underline{v}}(\varphi_{\underline{v}}) \neq 0$ so $e_{\underline{v}} \neq 0$. If $\dim_F V = n$ then $\dim_F V' = n$ and thus $\dim_F V'' = n$ and we have an isomorphism. □

The map $V \hookrightarrow V''$ is natural: the image $e_{\underline{v}}$ of $\underline{v}$ is intrinsic and does not depend on a choice of basis.

1.4.3. The dual space, infinite dimensions (Lecture 7).

LEMMA 55 (Interaction with past constructions). *We have*
(1) $(V/U)' \hookrightarrow V'$ *as* $\{\varphi \in V \mid \varphi(U) = \{0\}\}$.
(2) $(U \oplus V)' \simeq U' \oplus V'$.

PROOF. Universal property. □

COROLLARY 56. *Since* $(F)' \simeq F$, *it follows by induction that* $(F^n)' \simeq F^n$.

What about infinite sums?

- The universal property gives a bijection $(\bigoplus_{i\in I} V_i)' \longleftrightarrow \bigtimes_{i\in I} V_i'$, more generally
$$\mathrm{Hom}_F\left(\bigoplus_{i\in I} V_i, Z\right) \overset{1:1}{\longleftrightarrow} \bigtimes_{i\in I} \mathrm{Hom}_i\left(V_i, Z\right).$$
  - LHS has a vector space structure – should get one on the right.
  - Which leads us to observe that:
- Any Cartesian product $\bigtimes_{i\in I} W_i$ has a natural vector space structure, coming from pointwise addition and scalar multiplication.
  - Note that the underlying set is
$$\begin{aligned}\bigtimes_{i\in I} W_i &= \{f \mid f \text{ is a function with domain } I \text{ and } \forall i \in I : f(i) \in W_i\} \\ &= \left\{ f \colon I \to \bigcup_{i\in I} W_i \mid f(i) \in W_i \right\}.\end{aligned}$$
    - ∗ RMK: AC means Cartesian products nonempty, but our sets have a distinguished element so this is not an issue.
  - Define $\alpha\,(w_i)_{i\in I} + (w_i')_{i\in I} \overset{\text{def}}{=} (\alpha w_i + w_i')_{i\in I}$. This gives a vector space structure.
  - Denote the resulting vector space $\prod_{i\in I} W_i$ and called it the *direct product* of the W_i.
- The bijection $(\bigoplus_{i\in I} V_i)' \longleftrightarrow \prod_{i\in I} V_i'$ is now a linear isomorphism [in fact, the vector space structure on the right is the one transported by the isomorphism].

We now investigate $\prod_i W_i$ in general.

- Note that it contains a copy of each W_i (map $\underline{w} \in W_i$ to the sequence which has $\underline{w}$ in the ith position, and $\underline{0}$ at every other position).

- And these copies are linearly independent: if a sum of such vectors from distinct W_i is zero, then every coordinate was zero.
- Thus $\prod_i W_i$ contains $\bigoplus_{i\in I} W_i$ as an internal direct sum.
  - This subspace is exactly the subset $\{\underline{w} \in \prod_i W_i \mid \mathrm{supp}(\underline{w}) \text{ is finite}\}$.
  - And in fact, that subspace proves that $\bigoplus_{i\in I} W_i$ exists.
  - But $\prod_i W_i$ contains many other vectors – it is much bigger.

EXAMPLE 57. $\mathbb{R}^{\oplus\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$ – and the latter is the dual!

COROLLARY 58. *The dual of an infinite-dimensional space is much bigger than the sum of the duals, and the double dual is bigger yet.*

1.4.4. Question: we only have *finite* sums in linear algebra. What about infinite sums? Answer: no infinite sums in algebra. Definition of $\sum_{n=1}^{\infty} a_n = A$ from real analysis relies on *analytic properties* of A (it's a number close to the partial sums), not algebraic properties.

But, calculating sums can be understood in terms of linear functionals.

LEMMA 59 (Results from Calc II, reinterpreted). *Let $S \subset \mathbb{R}^{\mathbb{N}}$ denote the set of sequences $\underline{a}$ such that $\sum_{n=1}^{\infty} a_n$ converges.*

(1) $\mathbb{R}^{\oplus\mathbb{N}} \subset S \subset \mathbb{R}^{\mathbb{N}}$ *is a linear subspace.*
(2) $\Sigma\colon S \to \mathbb{R}$ *given by* $\Sigma(\underline{a}) = \sum_{n=1}^{\infty} a_n$ *is a linear functional.*

Philosophy: Calc I,II made element-by-element statements, but using linear algebra we can express them as statements on the whole space.

Now questions about summing are questions about *intelligently* extending the linear functional Σ to a bigger subspace. BUT: if an extension is to satisfy every property of summing series, it is actually the trivial (no) extension.

For more information let's talk about limits of sequences instead (once we have notions of generalized limits of sequences we can apply them to the sequence of partial sums of a series).

DEFINITION 60. Let $c \subset \ell^\infty \subset \mathbb{R}^{\mathbb{N}}$ be the sets of convergent, respectively bounded sequences.

LEMMA 61. *$c \subset \ell^\infty$ are subspaces, and* $\lim_{n\to\infty}\colon c \to \mathbb{R}$ *is a linear functional.*

EXAMPLE 62. Let $C\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$be the Cesàro map $(C\underline{a})_N = \frac{1}{N}\sum_{n=1}^{N} a_n$. This is clearly linear. Let $CS = C^{-1}(c)$ be the set of sequences which are Cesàro-convergent, and set $L \in CS'$ by $L(\underline{a}) = \lim_{n\to\infty}(C\underline{a})$. This is clearly linear (composition of linear maps). For example, the sequence $(0,1,0,1,\cdots)$ now has the limit $\frac{1}{2}$.

LEMMA 63. *If $\underline{a} \in c$ then $C\underline{a} \in c$ and they have the same limit. Thus L above is an extension of* $\lim_{n\to\infty}$.

THEOREM 64. *There are two functionals* $\mathrm{LIM}, \lim_\omega \in (\ell^\infty)'$ *("Banach limit", "limit along ultrafilter", respectively) such that:*

(1) *They are positive (map non-negative sequences to non-negative sequences);*
(2) *Agree with* $\lim_{n\to\infty}$ *on c;*
(3) *And, in addition*
   (a) $\mathrm{LIM} \circ S = \mathrm{LIM}$ *where* $S\colon \ell^\infty \to \ell^\infty$*is the* shift.
   (b) $\lim_\omega (a_n b_n) = (\lim_\omega a_n)(\lim_\omega b_n)$.

1.4.5. Pairings and bilinear forms (Lecture 8). Goal: identify the dual of a vector space in concerete terms. For this we need an abstract notion of dual not tied to the particular realization V' (similar to how we have an abstract notion of direct sum as "space generated by independent copies of V_i" which is not tied to the concrete realization as the subspace of tuples of finite support in $\prod_i V_i$).

OBSERVATION 65. *The* evaluation map $V \times V' \to F$ given by

$$\langle \underline{v}, \varphi \rangle = \varphi(\underline{v})$$

is *bilinear* (=linear in each variable).

Note that linearity in the first variable is equivalent to the linearity of φ, while linearity in the second variable is equivalent to the definition of the vector space structure on V'.

DEFINITION 66 (Pairings / bilinear maps). For any two vector spaces U, V a (bilinear) *pairing* between U, V is a map

$$\langle \cdot, \cdot \rangle : U \times V \to F$$

which is linear in each variable separately. Similarly we define a *bilinear map* $U \times V \to Z$.

EXAMPLE 67. The standard inner product on F^n: $\langle \underline{u}, \underline{v} \rangle = \sum_i u_i v_i$. More generally, given $B \in M_{m\times n}(F)$ have a pairing on $F^m \times F^n$ given by

$$\langle \underline{u}, \underline{v} \rangle = \sum_i u_i B_{ij} v_j \,.$$

More generally, given any bilinear pairing of U, V choose bases $\{\underline{u}_i\}_{i\in I} \subset U$, $\left\{\underline{v}_j\right\}_{j\in J} \subset V$ and define the *Gram matrix* by

$$B_{ij} = \left\langle \underline{u}_i, \underline{v}_j \right\rangle \,.$$

We can then compute the pairing of any two vectors: by the distributive law ("FOIL")

$$\left\langle \sum_i a_i \underline{u}_i, \sum_i b_j \underline{v}_j \right\rangle = \sum_{i,j} a_j B_{ij} b_j \,.$$

Conversely, any matrix B defines a bilinear pairing (aside: this is a linear bijection if you give pairings the obvious vector space structure).

1.4.6. Pairings: duality and degeneracy. Fix a bilinear form$\langle \cdot, \cdot \rangle : U \times V \to F$. Then for any $\underline{u} \in U$ we get a map $\varphi_{\underline{u}} \colon V \to F$ by $\varphi_u(\underline{v}) = \langle \underline{u}, \underline{v} \rangle$.

(1) $\varphi_{\underline{u}}$ is linear ($\in V'$) iff the pairing is linear in the second variable.
(2) The map $U \to V'$ given by $\underline{u} \to \varphi_{\underline{u}}$ is linear iff the pairing is linear in the first variable.

We conclude that every pairing gives a map $U \to V'$, and equivalently also a map $V \to U'$.

LEMMA 68. *We have a linear bijection* $\{$*pairings on* $U \times V\} \longleftrightarrow \mathrm{Hom}_F(U, V')$

PROOF. The inverse map associates to each $f \in \mathrm{Hom}_F(U, V')$ the bilinear form

$$\langle \underline{u}, \underline{v} \rangle_f = (f(\underline{u}))(\underline{v}) \,.$$

□

DEFINITION 69. Call the bilinear map *non-degenerate* if both maps $U \to V'$, $V \to U'$ are embeddings.

LEMMA 70. *A pairing is non-degenerate iff for every non-zero $\underline{u} \in U$ there is $\underline{v} \in V$ such that $\langle \underline{u}, \underline{v} \rangle \neq 0$ and conversely.*

Key idea: if the map $V \to U'$ associated to a pairing is bijective, then we can use V as a *model* for U' via the pairing.

EXAMPLE 71. The dot product is a non-degenerate pairing $F^n \times F^n$ hence identifies $(F^n)'$ with F^n.

Two further examples from functional analysis:

First we fix a compact topological space X. Then for any finite Borel measure μ on X and any continuous $f \in C(X)$ we have the *integral*

$$\int f \, \mathrm{d}\mu \, .$$

This is a bilinear pairing $C(X) \times \{\text{finite signed measures on } X\}$ which is non-degenerate.

THEOREM 72 (Riesz representation theorem). *Let X be compact. Then every continuous linear functional on $C(X)$ is given by a finite measure (that is, the continuous dual $C(X)'$ can be represented by the space of measures).*

Second, the inner product on Hilbert space is a non-degenerate pairing!

THEOREM 73 (Riesz representation theorem). *Let $\mathcal{H}$ be a Hilbert space. Then every continuous linear functional on $\mathcal{H}$ is of the form $\langle \underline{u}, \cdot \rangle$.*

1.4.7. The dual of a linear map (Lecture 8, continued).

CONSTRUCTION 74. *Let $T \in \mathrm{Hom}(U,V)$. Set $T' \in \mathrm{Hom}(V',U')$ by $(T'\varphi)(\underline{v}) = \varphi(T\underline{v})$.*

LEMMA 75. *This is a linear map $\mathrm{Hom}(U,V) \to \mathrm{Hom}(V',U')$. An isomorphism if U,V finite-dimensional.*

LEMMA 76. $(TS)' = S'T'$

PROOF. PS3 $\square$

1.5. Multilinear algebra and tensor products (Lectures 9-14)

1.5.1. Bilinear forms (Lecture 9).

DEFINITION 77. Let $\{V_i\}_{i \in I}$ be vector spaces, W another vector space. A function $f \colon \times_{i \in I} V_i \to W$ is said to be *multilinear* if it is linear in each variable.

EXAMPLE 78 (Bilinear maps).
1. $f(x,y) = xy$ is bilinear $F^2 \to F$.
2. The map $(T, \underline{v}) \mapsto T\underline{v}$ is a multilinear map $\mathrm{Hom}(V,W) \times V \to W$.
3. For a matrix $A \in M_{n,m}(F)$ have $(\underline{x}, \underline{y}) \mapsto {}^{\mathrm{t}}\underline{x}A\underline{y}$ on $F^n \times F^m$.
4. For $\varphi \in U'$, $\psi \in V'$ have $(\underline{u}, \underline{v}) \mapsto \varphi(\underline{u})\psi(\underline{v})$, and finite combinations of those.

REMARK 79. A *bilinear function* on $U \times V$ is not the same as a *linear function on* $U \oplus V$. For example: is $f(a\underline{u}, a\underline{v})$ equal to $af(\underline{u}, \underline{v})$ or to $a^2 f(\underline{u}, \underline{v})$? That said, $\bigoplus V_i$ was *universal* for maps from V_i. It would be nice to have a space which is universal for multilinear maps. We only discuss the finite case.

EXAMPLE 80. A multilinear function $B\colon U\times\{\underline{0}\}\to F$ has $B(\underline{u},\underline{0})=B(\underline{u},0\cdot\underline{0})=0\cdot B(\underline{u},\underline{0})=0$. A multilinear function $B\colon U\times F\to F$ has $B(\underline{u},x)=B(\underline{u},x\cdot 1)=xB(\underline{u},1)=x\varphi(\underline{u})$ where $\varphi(\underline{u})=B(\underline{u},1)\in U'$.

We can reduce everything to Example 78(3): Fix bases $\{\underline{u}_i\},\{\underline{v}_j\}$. Then

$$B\left(\sum_i x_i\underline{u}_i,\sum_i y_j\underline{v}_j\right)=\sum_{i,j}x_iB(\underline{u}_i,\underline{v}_j)y_j={}^{\mathrm{t}}\underline{x}B\underline{y}$$

where $B_{ij}=B(\underline{u}_i,\underline{v}_j)$. Note: $x_i=\varphi_i(\underline{u})$ where $\{\varphi_i\}$ is the dual basis. Conclude that

$$B=\sum_{i,j}B\left(\underline{u}_i,\underline{v}_j\right)\varphi_i\psi_j\,. \tag{1.5.1}$$

Easy to check that this is an expansion in a basis (check against $(\underline{u}_i,\underline{v}_j)$). We have shown:

PROPOSITION 81. *The set $\{\varphi_i\psi_j\}_{i,j}$ is a basis of the space of bilinear forms $U\times V\to F$.*

COROLLARY 82. *The space of bilinear forms on $U\times V$ has dimension* $\dim_F U\cdot\dim_F V$.

REMARK 83. Also works in infinite dimensions, since can have the sum (1.5.1) be infinite – every pair of vectors only has finite support in the respective bases.

1.5.2. The tensor product (Lecture 10-11). Now let's fix U,V and try to construct a space that will classify bilinear maps on $U\times V$.

- Our space will be generated by terms $\underline{u}\otimes\underline{v}$ on which we can evaluate f to get $f(\underline{u},\underline{v})$.
- Since f is multilinear, $f(a\underline{u},b\underline{v})=abf(\underline{u},\underline{v})$ so need $(a\underline{u})\otimes(b\underline{v})=ab(\underline{u}\otimes\underline{v})$.
- Similarly, since $f(\underline{u}_1+\underline{u}_2,\underline{v})=f(\underline{u}_1,\underline{v})+f(\underline{u}_2,\underline{v})$ want $(\underline{u}_1+\underline{u}_2)\otimes(\underline{v}_1+\underline{v}_2)=\underline{u}_1\otimes\underline{v}_1+\underline{u}_2\otimes\underline{v}_1+\underline{u}_1\otimes\underline{v}_2+\underline{u}_2\otimes\underline{v}_2$.

CONSTRUCTION 84 (Tensor product). *Let U,V be spaces. Let $X=F^{\oplus(U\times V)}$ be the formal span of all expressions of the form $\{\underline{u}\otimes\underline{v}\}_{(\underline{u},\underline{v})\in U\times V}$. Let $Y\subset X$ be the subspace spanned by*

$$\{(a\underline{u})\otimes(b\underline{v})-ab(\underline{u}\otimes\underline{v})\mid a,b\in F,\,(\underline{u},\underline{v})\in U\times V\}$$

and

$$\{(\underline{u}_1+\underline{u}_2)\otimes(\underline{v}_1+\underline{v}_2)-(\underline{u}_1\otimes\underline{v}_1+\underline{u}_2\otimes\underline{v}_1+\underline{u}_1\otimes\underline{v}_2+\underline{u}_2\otimes\underline{v}_2)\mid **\}\,.$$

Then set $U\otimes V=X/Y$ and let $\iota\colon U\times V\to U\otimes V$ be the map $\iota(\underline{u},\underline{v})=(\underline{u}\otimes\underline{v})+Y$.

THEOREM 85. *ι is a bilinear map. For any space W any any bilinear map $f\colon U\times V\to W$, there is a unique linear map $\tilde{f}\colon U\otimes V\to W$ such that $f=\tilde{f}\circ\iota$.*

PROOF. Uniqueness is clear, since $\tilde{f}(\underline{u}\otimes\underline{v})=f(\underline{u},\underline{v})$ fixes $\tilde{f}$ on a generating set. For existence we need to show that if $\tilde{\tilde{f}}\colon X\to W$ is defined by $\tilde{\tilde{f}}(\underline{u}\otimes\underline{v})=f(\underline{u},\underline{v})$ then $\tilde{\tilde{f}}$ vanishes on Y and hence descends to $U\otimes V$. □

PROPOSITION 86. *Let B_U,B_V be bases for U,V respectively. Then $\{\underline{u}\otimes\underline{v}\mid\underline{u}\in B_U,\underline{v}\in B_V\}$ is a basis for $U\otimes V$.*

PROOF. Spanning: use bilinearity of ι. Independence: let $\{\varphi_{\underline{u}}\}_{\underline{u}\in B_U}\subset U'$, $\{\psi_{\underline{v}}\}_{\underline{v}\in B_V}\subset V'$ be the dual bases. Then $\varphi_{\underline{u}}\psi_{\underline{v}}$is a bilinear map $U\times V\to F$, and the sets $\{\underline{u}\otimes\underline{v}\}_{(\underline{u},\underline{v})\in B_U\times B_V}$ and $\{\widetilde{\varphi_{\underline{u}}\psi_{\underline{v}}}\}_{(\underline{u},\underline{v})\in B_U\times B_V}$ are dual bases. □

COROLLARY 87. $\dim_F(U\otimes V)=\dim_F U\cdot\dim_F V$.

EXAMPLE 88. Examples of tensor products

(1) $\mathbb{R}^n\otimes\mathbb{R}^m$ encoded as matrices.
(2) Polynomial algebra: $F[x]\otimes F[y]\simeq F[x,y]$
(3) Functions on product spaces: let X,Y be compact then $C(X)\otimes C(Y)$ dense in $C(X\times Y)$.
   - Note Fubini's Theorem can be obtained from this.
(4) Quantum mechanics: state space for two particles is the (completion of the) tensor product of the state spaces for the individual particles.
   - Note that many states are not "pure tensors".

1.5.3. The Universal Property (Lecture 12). Abstract view of tensor products (note the indefinite article!)

DEFINITION 89 (Abstract tensor product). A *tensor product* of the spaces U,V is a pair (W,ι) where W is a vector space, $\iota\colon U\times V\to W$ is bilinear, and for every bilinear map $f\colon U\times V\to Z$ there is a unique $\bar{f}\in\mathrm{Hom}_F(W,Z)$ such that $f=\bar{f}\circ\iota$.

REMARK 90. Informally, ι is the "most general bilinear map on $U\times V$.

EXAMPLE 91. We show that the image of ι spans W. Indeed if not there would be a non-zero functional $\bar{f}\in W'$ vanishing on the span of that image, and then $\bar{f}\circ\iota=0\circ\iota$ would both be the zero bilinear form, violating uniqeness.

Before we explain how to use this property, we return to the example of direct sum.

DEFINITION 92 (Abstract direct sum). A *direct sum* of $\{V_i\}_{i\in I}$ is a space W and maps $e_i\in\mathrm{Hom}_F(V_i,W)$ such that for every system of maps $f_i\in\mathrm{Hom}_F(V_i,Z)$ there is a unique $\bar{f}\colon W\to Z$ such that $f_i=\bar{f}\circ e_i$ for each i.

PROPOSITION 93. *Direct sums are unique up to a unique isomorphism.*

PROOF. Suppose W' is another direct sum with system of inclusions $\{e'_i\}_{i\in I}$. Then W' is a space with system of maps, so by hypothesis there is a unique $\bar{f}'\colon W\to W'$ such that $e'_i=\bar{f}'\circ e_i$ (note that this is the key requirement from an isomorphism preserving the structure, so we see both that such a map exists and that it is unique, but we don't know it is an isomorphism yet).

By symmetry there is also $\bar{f}\colon W'\to W$ such that $e_i=\bar{f}\circ e'_i$.

Next, note that $\bar{f}\circ\bar{f}'$ and id_W are both maps $W\to W$ and they satisfy

$$\begin{aligned}\left(\bar{f}\circ\bar{f}'\right)\circ e_i &= \bar{f}\circ\left(\bar{f}'\circ e_i\right)\\ &= \bar{f}\circ e'_i && \text{choice of } \bar{f}'\\ &= e_i && \text{choice of } \bar{f}\\ &= \mathrm{id}_W\circ e_i\end{aligned}$$

so by the universal property (applied to the system of maps e_i with target W), $\bar{f}\circ\bar{f}'=\mathrm{id}_W$. By symmetry we also have $\bar{f}'\circ\bar{f}=\mathrm{id}_{W'}$ and we are done. □

Consider now direct products. Reversing all the arrows (on the board, modify the definition in red ink), we get:

DEFINITION 94 (Abstract direct product). A *direct product* of $\{V_i\}_{i\in I}$ is a space W and maps $\pi_i \in \mathrm{Hom}_F(W,V_i)$ such that for every system of maps $f_i \in \mathrm{Hom}_F(Z,V_i)$ there is a unique $\bar{f}\colon Z \to W$ such that $f_i = \pi_i \circ \bar{f}$ for each i.

PROPOSITION 95. *Direct products are unique up to a unique isomorphism.*

PROOF. Reverse all the arrows in the previous proof. □

PROPOSITION 96. *Tensor products are unique up to a unique isomorphism.*

PROOF. Let (W,ι), (W',ι') be two tensor products of U,V. Then since $\iota'\colon U\times V \to W'$ is bilinear there is a unique $\bar{f}'\colon W \to W'$ such that $\iota' = \bar{f}'\circ\iota$.

REMARK. Note that this is a basic requirement for an "isomorphism of tensor products": it must identify the vector representing $\underline{u}\otimes\underline{v}$ on both sides. Using the universal property we saw both that we can actually identify these vectors and that this identification extends to a linear map of the tensor product spaces.

Continuing with the proof, for the same reason there is $\bar{f}\colon W' \to W$ such that $\iota = \bar{f}\circ\iota'$.

REMARK. One can now finish the proof by noting, for example, that $\bar{f}\circ\bar{f}'$ fixes elements of the form $\underline{u}\otimes\underline{v}$ and that these span the tensor product. But we prefer a proof which doesn't "look under the hood" and use the vectors in the vector space.

Finally, we have

$$\begin{aligned}
\left(\bar{f}\circ\bar{f}'\right)\circ\iota &= \bar{f}\circ\left(\bar{f}'\circ\iota\right) & \\
&= \bar{f}\circ\iota' & \text{choice of } \bar{f}' \\
&= \iota & \text{choice of } \bar{f} \\
&= \mathrm{id}_W\circ\iota & .
\end{aligned}$$

We have shown that the bilinear map $\iota\colon U\times V\to W$ is represented by both homomorphisms $\bar{f}\circ\bar{f}'$ and id_W so they must be equal, and by symmetry we conclude that $\bar{f}'\circ\bar{f} = \mathrm{id}_{W'}$ as well so that $\bar{f},\bar{f}'$ are the desired isomorphisms. □

REMARK 97. This point of view leads to "category theory" where one forgets about the specific algebraic structure under consideration (here vector spaces) and considers only the statements about objects and homomorphisms. This way theorems about "direct sums", say, apply for any construction of direct sum regardless of the underlying algebraic structures.

For example, we get direct sums (and direct products) of groups, rings, modules, vector spaces. But we also get direct sums of topological spaces (this turns out to be the disjoint union) and direct products of topological spaces (this is the Tychonoff product).

1.5.4. Extension of scalars. Let K/F be a field extension.

PROPOSITION 98. *The field operations of K endow it with the structure of an F-vectorspace.*

LEMMA-DEFINITION 99. *Let V be an F-vectorspace. Then $V_K \overset{def}{=} K\otimes_F V$ has the natural structure of a K-vectorspace.*

THEOREM 100. *Let $B\subset V$ be an F-basis. Then $\{1_K\otimes\underline{v}\}_{\underline{v}\in B}\subset V_K$ is a K-basis*

EXERCISE 101. The map $V \mapsto V_K$ is functorial: every $f \in \mathrm{Hom}(U,V)$ extends naturally to a map $f_K \colon U_K \to V_K$ (take $f_K = 1_K \otimes f$), with the same matrix as f (with respect to bases at in the Theorem). If L/K if a further extension then there is a natural transformation $V_L \simeq (V_K)_L$.

EXERCISE 102. Extension of scalars respects all the constructions: direct sum, direct product, quotient, kernel, image, tensor product: $(U \otimes_F V)_K \simeq U_K \otimes_K V_K$.

1.5.5. Symmetric and antisymmetric tensor products (Lecture 13). For motivation, think of $U \otimes U$ as the state space of a *pair* of *identical* quantum particles. What happens when we swap them? Represent swapping them by the obvious map $T \in \mathrm{End}_F(U \otimes U)$.

FACT 103. *Some fundamental particles ("Bosons") always have states in the $+1$-eigenspace. Other particles ("Fermions") always have states in the -1-eigenspace. Note that a state like $\underline{u} \otimes \underline{u}$ is permitted to Bosons, but* prohibited (!) *to Fermions. This is called the "Fermi exclusion principle".*

ASSUMPTION 104. *For this section,* $\mathrm{char}(F) = 0$.

Let $(12) \in S_2$ act on $V \otimes V$ by exchanging the factors (why is this well-defined?).

LEMMA 105. *Let $T \in \mathrm{End}_F(U)$ satisfy $T^2 = \mathrm{Id}$. Then U is the direct sum of the two eigenspaces.*

DEFINITION 106. $\mathrm{Sym}^2 V$ and $\bigwedge^2 V$ are the eigenspaces.

PROPOSITION 107. *Generating sets and bases.*

In general, let S_k act on $V^{\otimes k}$.

- What do we mean by that? Well, this classifies n-linear maps $V \times \cdots \times V \to Z$. Universal property gives isom of $(U \otimes V) \otimes W$, $U \otimes (V \otimes W)$.
- Why action well-defined? After all, the set of pure tensors is nonlinear. So see first as multilinear map $V^n \to V^{\otimes n}$.
- Single out $\mathrm{Sym}^k V$, $\bigwedge^k V$. Note that there are other representations.
- Claim: bases

1.5.6. Bases of Sym^k, $\bigwedge^k$, determinants (Lecture 14).

EXAMPLE 108. $\bigwedge^2 \mathbb{R}^3$ is three-dimensional, which explains the cross-product as giving anti-symmetric 2-tensors rather than vectors.

PROPOSITION 109. *Symmetric/antisymmetric tensors are generating sets; bases coming from subsets of basis.*

Tool: the maps $P_k^{\pm} \colon V^{\otimes k} \to V^{\otimes k}$ given by $P_k^{\pm}(\underline{v}_1 \otimes \cdots \otimes \underline{v}_k) = \frac{1}{k!} \sum_{\sigma \in S_k} (\pm)^{\sigma} \left(\underline{v}_{\sigma(1)} \otimes \cdots \otimes \underline{v}_{\sigma(k)} \right)$.

LEMMA 110. *These are well defined (extensions of linear maps). Fix elements of* $\mathrm{Sym}^k V$, $\bigwedge^k V$ *respectively, images are in those subspaces (check $\tau \circ P_k^{\pm} = (\pm)^{\tau} P_k^{\pm}$). Conclude that image is spanned by image of basis.*

EXAMPLE 111. Exterior forms of top degree and determinants.

CHAPTER 2

Structure Theory: The Jordan Canonical Form

2.1. Introduction (Lecture 15)

2.1.1. The two paradigmatic problems. Fix a vector space V of dimension $n < \infty$ (in this chapter, all spaces are finite-dimensional unless stated otherwise), and a map $T \in \mathrm{End}(V)$. We will try for two kinds of structural results:

(1) ["decomposition"] $T = RS$ where $R, S \in \mathrm{End}(V)$ are "simple"
(2) ["canonical form"] There is a basis $\{\underline{v}_i\}_{i=1}^n \subset V$ in which the matrix of T is "simple".

EXAMPLE 112. (From 1st course)

(1) (Gaussian elimination) Every matrix $A \in M_n(F)$ can be written in the form $A = E_1 \cdots E_k \cdot A_{\mathrm{rr}}$ where E_i are "elementary" (row operations or rescaling) and A_{rr} is row-reduced.
(2) (Spectral theory) Suppose T is diagonable. Then there is a basis in which T is diagonal.

As an example of how to use (1), suppose $\det(A)$ is defined for matrices by column expansion. Then can show (Lemma 1) that $\det(EX) = \det(E)\det(X)$ whenever E is elementary and that (Lemma 2) $\det(AX) = \det(A)\det(X)$ whenever A is row-reduced. One can then prove

THEOREM 113. *For all A, B,* $\det(AB) = \det(A)\det(B)$.

PROOF. Let $\mathcal{D} = \{A \mid \forall X : \det(AX) = \det(A)\det(X)\}$. Then we know that all $A_{\mathrm{rr}} \in \mathcal{D}$ and that if $A \in \mathcal{D}$ then for any elementary E, $\det((EA)X) = \det(E(AX)) = \det(E)\det(AX) = \det(E)\det(A)\det(X) = \det(EA)\det(X)$ so $EA \in \mathcal{D}$ as well. It now follows from Gauss's Theorem that $\mathcal{D}$ is the set of all matrices. □

2.1.2. Triangular matrices.

DEFINITION 114. $A \in M_n(F)$ is upper (lower) triangular if ...

Significance: these are very good for computation. For example:

LEMMA 115. *The lower-triangular matrix L is invertible iff its diagonal entries are non-zero.*

The proof is:

ALGORITHM 116 (Forward-substitution). *Let L be lower-triangular with non-zero diagonal entries. Then the solution to $L\underline{x} = \underline{b}$ is given by $x_i = \frac{b_i - \sum_{j=1}^{i-1} l_{ij} x_j}{l_{ii}}$ for $i = 1, 2, \ldots, n$.*

REMARK 117. Note that the algorithm does exactly as many multiplications as non-zero entries in U. Hence better than Gaussian elimination for general matrix ($O(n^3)$), really good for sparse matrix, and doesn't require storing the matrix entries only the way to calculate u_{ij} (in particular no need to find inverse).

EXERCISE 118. (1) Express this as a formula for the inverse of a lower-triangular matrix (2) Develop the backward-substitution algorithm for upper-triangular matrices and find a formula for their inverses.

COROLLARY 119. *If $A = LU$ we can efficiently solve $A\underline{x} = \underline{b}$.*

Note that we don't like to store inverses. For example, because they are generally dense matrices even if L, U are sparse.

We now try to look for a vector-space interpretation of being triangular. For this note that if $U \in M_n(F)$ is upper-triangular then

$$\begin{aligned} U\underline{e}_1 &= u_{11}\underline{e}_1 \in \mathrm{Span}\{\underline{e}_1\} \\ U\underline{e}_2 &= u_{12}\underline{e}_1 + u_{22}\underline{e}_2 \in \mathrm{Span}\{\underline{e}_1, \underline{e}_2\} \\ \vdots &= \vdots \\ U\underline{e}_k &\in \mathrm{Span}\{\underline{e}_1, \ldots, \underline{e}_k\} \\ \vdots &= \vdots. \end{aligned}$$

In particular, we found a family of subspaces $V_i = \mathrm{Span}\{\underline{e}_1, \ldots, \underline{e}_i\}$ such that $U(V_i) \subset V_i$, such that$\{\underline{0}\} = V_0 \subset V_1 \subset \cdots \subset V_n = F^n$ and such that $\dim V_i = i$.

EXERCISE 120 (Cholesky decomposition). If A is positive-definite, then $A = LL^{\mathrm{t}}$ for a lower-triangular matrix L.

THEOREM 121. $T \in \mathrm{End}(V)$ *has an upper-triangular matrix wrt some basis iff there are T-invariant subspaces $\{\underline{0}\} = V_0 \subset V_1 \subset \cdots \subset V_n = F^n$ with $\dim V_i = i$.*

PROOF. We just saw necessity. For sufficiency, given V_i choose for $1 \leq i \leq n$, $\underline{v}_i \in V_i \setminus V_{i-1}$. These exist (the dimension increases by 1), are a linearly independent set (each vector is independent of its predecessors) and the first i span V_i (by dimension count). Finally for each i, $T\underline{v}_i \in T(V_i) \subset V_i = \mathrm{Span}\{\underline{v}_1, \ldots, \underline{v}_i\}$ so the matrix of T in this basis is upper triangular. □

2.2. The minimal polynomial (Lecture 16)

Recall we have an n-dimensional F-vector space V.

- A key tool for studying linear maps is studying polynomials in the maps (we saw how to analyze maps satisfying $T^2 = \mathrm{Id}$, for example).
- We will construct a gadget (the "minimal polynomial") attached to every linear map on V. It is a polynomial, and will tell us a lot about the map.
- Computationally speaking, this polynomial cannot be found efficiently. It is a tool of theorem-proving in abstract algebra.

DEFINITION 122. Given a polynomial $f \in F[x]$, say $f = \sum_{i=0}^{d} a_i x^i$ and a map $T \in \mathrm{End}(V)$ set (with $T^0 = \mathrm{Id}$)

$$f(T) = \sum_{i=0}^{d} a_i T^i .$$

LEMMA 123. *Let $f, g \in F[x]$. Then $(f+g)(T) = f(T) + g(T)$ and $(fg)(T) = f(T)g(T)$. In other words, the map $f \mapsto f(T)$ is a linear map $F[x] \to \mathrm{End}(V)$, also respecting multiplication ("a map of F-algebras", but this is beyond our scope).*

PROOF. Do it yourself. $\square$

- Given a linear map our first instinct is to study the kernel and the image. [Aside: the kernel is an *ideal* in the algebra].
- We'll examine the kernel and leave the image for later.

LEMMA 124. *There is a non-zero polynomial $f \in F[x]$ such that $f(T) = 0$. In fact, there is such f with* $\deg f \le n^2$.

PROOF. $F[x]$ is infinite-dimensional while $\mathrm{End}_F(V)$ is finite-dimensional. Specifically, $\dim_F F[x]^{\le n^2} = n^2 + 1$ while $\dim_F \mathrm{End}_F(V) = n^2$. $\square$

REMARK 125. We will later show (Theorem of Cayley–Hamilton) that the *characteristic polynomial $P_T(x) = \det(x\,\mathrm{Id} - T)$* from basic linear algebra has this property.

- Warning: we are about to divide polynomials with remainder.

PROPOSITION 126. *Let $I = \{f \in F(x) \mid f(T) = 0\}$. Then I contains a unique non-zero monic polynomial of least degree, say $m(x)$, and $I = \{g(x)m(x) \mid g \in F[x]\}$ is the set of multiples of m.*

PROOF. Let $m \in I$ be a non-zero member of least degree. Dividing by the leading coefficient we may assume m monic. Now suppose m' is another such. Then $m - m' \in I$ (this is a subspace) is of strictly smaller degree. It must therefore be the zero polynomial, and m is unique. Clearly if $g \in F[x]$ then $(gm)(T) = g(T)m(T) = 0$. Conversely, given any $f \in I$ we can divide with remainder and write $f = qm + r$ for some $q, r \in F[x]$ with $\deg r < \deg m$. Evaluating at T we find $r(T) = 0$, so $r = 0$ and $f = qm$. $\square$

DEFINITION 127. Call $m(x) = m_T(x)$ the *minimal polynomial* of T.

REMARK 128. We will later prove directly that $\deg m_T(x) \le n$.

EXAMPLE 129. (Minimal polynomials)

(1) $T = \mathrm{Id}$, $m(x) = x - 1$.

(2) $T = \begin{pmatrix} & 1 \\ & \end{pmatrix}$, $T^2 = 0$ but $T \neq 0$ so $m_T(x) = x^2$.

(3) $T = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}$, $T^2 = \begin{pmatrix} 1 & 2 \\ & 1 \end{pmatrix}$ so $\left(T^2 - \mathrm{Id}\right) = 2(T - \mathrm{Id})$ so $T^2 - 2T + \mathrm{Id} = 0$ so $m_T(x) | (x-1)^2$. But $T - \mathrm{Id} \neq 0$ so $m_T(x) = (x-1)^2$.

(4) $T = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, $T^2 = \mathrm{Id}$ so $m_T(x) = x^2 - 1 = (x-1)(x+1)$.

- In the eigenbasis $\left\{ \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \right\}$ the matrix is $\begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ – we saw this in a previous class.

(5) $T = \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}$, $T^2 = -\mathrm{Id}$ so $m_T(x) = x^2 + 1$.

  (a) If $F = \mathbb{Q}$ or $F = \mathbb{R}$ this is irreducible. No better basis.

  (b) If $F = \mathbb{C}$ (or $\mathbb{Q}(i)$) then factor $m_T(x) = (x - i)(x + i)$ and in the eigenbasis $\left\{ \begin{pmatrix} 1 \\ \pm i \end{pmatrix} \right\}$ the matrix has the form $\begin{pmatrix} -i & \\ & i \end{pmatrix}$.

(6) $V = F[x]^{<n}$ (polynomials of degree less than n), $T = \frac{d}{dx}$. Then $T^n = 0$ but $T^{n-1} \neq 0$ (why?) so $m_T(x) = x^n$.
(7) [To be proved in problem set] Let $D = \text{diag}(a_1, \ldots, a_n)$ be diagonal, its entries being the distinct numbers $\{b_1, \cdots, b_r\}$ (perhaps with repetition). Then its minimal polynomial is $\prod_{i=1}^{r}(x - b_r)$ [cf (1),(4),(5)]

We now connect the minimal polynomial with the spectrum.

LEMMA 130 (Spectral calculus). *Suppose that $T\underline{v} = \lambda\underline{v}$. Then $f(T)\underline{v} = f(\lambda)\underline{v}$.*

PROOF. Work it out at home. □

REMARK 131. The same proof shows that if the subspace W is T-invariant ($T(W) \subset W$) then W is $f(T)$-invariant for all polynomials f.

COROLLARY 132. *If λ is an eigenvalue of T then $m_T(\lambda) = 0$. In particular, if $m_T(0) \neq 0$ then T is invertible (0 is cannot be eigenvalue)*

We now use the *minimality* of the minimal polynomial.

THEOREM 133. *T is invertible iff $m_T(0) \neq 0$.*

PROOF. Suppose that T is invertible and that $\sum_{i=1}^{d} a_i T^i = 0$ [note $a_0 = 0$ here]. Then this is not the minimal polynomial since multiplying by T^{-1} also gives

$$\sum_{i=0}^{d-1} a_{i+1} T^i = 0\,.$$

□

COROLLARY 134. *$\lambda \in F$ is an eigenvalue of T iff λ is a root of $m_T(x)$.*

PROOF. Let $S = T - \lambda\,\text{Id}$. Then $m_S(x) = m_T(x+\lambda)$. Then $\lambda \in \text{Spec}_F(T) \iff S$ not invertible $\iff$ $m_S(0) = 0 \iff m_T(\lambda) = 0$. □

REMARK 135. The characteristic polynomial $P_T(x)$ also has this property – this is how eigenvalues are found in basic linear algebra.

2.3. Generalized eigenspaces and Cayley–Hamilton (Lectures 17-18)

REMARK 136. A slower schedule covers 2.3.1 in one lecture, 2.3.2 and half of 2.3.3 in another, and finishes 2.3.3 in a third lecture.

2.3.1. Generalized Eigenspaces (Lecture 17). Continue with $T \in \text{End}_F(V)$, $\dim_F(V) = n$. Recall that T is diagonable iff V is the direct sum of the eigenspace. For non-diagonable maps we need something more sophisticated.

PROBLEM 137. Find a matrix $A \in M_2(F)$ which only has a 1-d eigenspace.

DEFINITION 138. Call $\underline{v} \in V$ a *generalized eigenvector* of T if for some $\lambda \in F$ and $k \geq 1$, $(T-\lambda)^k \underline{v} = \underline{0}$. Let $V_\lambda \subset V$ denote the set of generalized λ-eigenvectors and $\underline{0}$. Call λ a generalized eigenvalue of T if $V_\lambda \neq \{\underline{0}\}$.

In particular, if $T\underline{v} = \lambda\underline{v}$ then $\underline{v} \in V_\lambda$.

PROPOSITION 139 (Generalized eigenspaces). (1) *Each V_λ is a T-invariant subspace.*
(2) *Let $\lambda \neq \mu$. Then $(T-\mu)$ is invertible on V_λ.*
(3) *$V_\lambda \neq \{\underline{0}\}$ iff $\lambda \in \mathrm{Spec}_F(T)$.*

PROOF. Let $\underline{v}, \underline{v}' \in V_\lambda$ be killed by $(T-\lambda)^k, (T-\lambda)^{k'}$ respectively. Then $\alpha\underline{v}+\beta\underline{v}'$ is killed by $(T-\lambda)^{\max\{k,k'\}}$. Also, $(T-\lambda)^k T\underline{v} = T(T-\lambda)^k\underline{v} = \underline{0}$ so $T\underline{v} \in V_\lambda$ as well.

Let $\underline{v} \in \mathrm{Ker}(T-\mu)$ be non-zero. By Lemma 130, for any k we have $(T-\lambda)^k\underline{v} = (\mu-\lambda)^k\underline{v} \neq \underline{0}$ so $\underline{v} \notin V_\lambda$.

Finally, given λ and non-zero $\underline{v} \in V_\lambda$ let k be minimal such that $(T-\lambda)^k\underline{v} = 0$. Then $(T-\lambda)^{k-1}\underline{v}$ is non-zero and is an eigenvector of eigenvalue λ. □

THEOREM 140. *The sum $\bigoplus_{\lambda\in\mathrm{Spec}_F(T)} V_\lambda \subset V$ is direct.*

PROOF. Let $\sum_{i=1}^r \underline{v}_i = \underline{0}$ be a minimal dependence with $\underline{v}_i \in V_{\lambda_i}$ for distinct λ_i. Applying $(T-\lambda_r)^k$ for k large enough to kill $\underline{v}_r$ we get the dependence.

$$\sum_{i=1}^{r-1}(T-\lambda_r)^k\underline{v}_i = \underline{0}.$$

Now $(T-\lambda_r)^k\underline{v}_i \in V_{\lambda_i}$ since these are T-invariant subspaces, and for $1 \leq i \leq r-1$ is non-zero since $T-\lambda_r$ is invertible there. This shorter dependence contradicts the minimality. □

REMARK 141. The sum may very well be empty – there are non-trivial maps without eigenvalues (for example $\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \in M_2(\mathbb{R})$).

2.3.2. Algebraically closed fields. We all know that sometimes linear maps fail to have eigenvalues, even though they "should". In this course we'll blame the field, not the map, for this deficiency.

DEFINITION 142. Call the field F *algebraically closed* if every non-constant polynomial $f \in F[x]$ has a root in F. Equivalently, if every non-constant polynomial can be written as a product of linear factors.

FACT 143 (Fundamental theorem of algebra). *$\mathbb{C}$ is algebraically closed.*

REMARK 144. Despite the title, this is a theorem of *analysis*.

Discussion. The goal is to create enough eigenvalues so that the generalized eigenspaces explain all of V. The first point of view is that we can simple "define the problem away" by restricting to the case of algebraically closed fields. But this isn't enough, since sometimes we are given maps over other fields. This already appears in the diagonable case, dealt with in 223: we can view $\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \in M_2(\mathbb{R})$ instead as $\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \in M_2(\mathbb{C})$, at which point it becomes diagonable. In other words, we can take a *constructive* point of view:

- Starting with any field F we can "close it" by repeatedly adding roots to polynomial equations until we can't, obtaining an "algebraic closure" $\bar{F}$ [the difficulty is in showing the process eventually stops].
  - This explains the "closed" part of the name – it's closure under an operation.

– [Q: do you need the full thing? A: In fact, it's enough to pass to the splitting field of the minimal polynomial]

- We now make this work for linear maps, with three points of view:
  (1) (matrices) Given $A \in M_n(F)$ view it as $A \in M_n(\bar{F})$, and apply the theory there.
  (2) (linear maps) Given $T \in \mathrm{End}_F(V)$, fix a basis $\{\underline{v}_i\}_{i=1}^n \subset V$, make the formal span $\bar{V} = \bigoplus_{i=1}^n \bar{F}\underline{v}_i$ and extends T to $\bar{V}$ by the property of having the same matrix.
  (3) (coordinate free) Given V over F set $\bar{V} = \bar{F} \otimes_F V$ (considering $\bar{F}$ as an F-vectorspace), and extend T (by $\bar{T} = \mathrm{Id}_{\bar{F}} \otimes_F T$).

2.3.3. The direct sum decomposition and Cayley–Hamilton (Lecture 18).

LEMMA 145. *Suppose F is algebraically closed and that $1 \leq \dim_F V < \infty$. Then every $T \in \mathrm{End}_F(V)$ has an eigenvector.*

PROOF. $m_T(x)$ has roots. □

We suppose now that F is algebraically closed, in other words that every linear map has an eigenvalue. The following is the key structure theorem for linear maps:

THEOREM 146. *(with F algebraically closed) We have $V = \bigoplus_{\lambda \in \mathrm{Spec}_F(T)} V_\lambda$.*

PROOF. Let $m_T(x) = \prod_{i=1}^r (x-\lambda_i)^{k_i}$ and let $W = \bigoplus_{i=1}^r V_{\lambda_i}$. Supposing that $W \neq V$, let $\bar{V} = V/W$ and consider the quotient map $\bar{T} \in \mathrm{End}_F(\bar{V})$ defined by $\bar{T}(\underline{v}+W) = T\underline{v}+W$. Since $\dim_F \bar{V} \geq 1$, $\bar{T}$ has an eigenvalue there. We first check that this eigenvalue is one of the λ_i. Indeed, for any polynomial $f \in F[x]$, $f(\bar{T})(\underline{v}+W) = (f(T)\underline{v})+W$, and in particular $m_T(\bar{T}) = 0$ and hence $m_{\bar{T}}|m_T$.

Renumbering the eigenvalues, we may assume $\bar{V}_{\lambda_r} \neq \{\underline{0}\}$, and let $\underline{v} \in V$ be such that $\underline{v}+W \in \bar{V}_{\lambda_r}$ is non-zero, that is $\underline{v} \notin W$. Since $\prod_{i=1}^{r-1}(\bar{T}-\lambda_i)^{k_i}$ is invertible on $\bar{V}_{\lambda_r}$, $\underline{u} = \prod_{i=1}^{r-1}(T-\lambda_i)^{k_i}\underline{v} \notin W$. But $(T-\lambda_r)^{k_r}\underline{u} = m_T(T)\underline{v} = \underline{0}$ means that $\underline{u} \in V_{\lambda_R} \subset W$, a contradiction. □

PROPOSITION 147. *In $m_T(x) = \prod_{i=1}^r (x-\lambda_i)^{k_i}$, the number k_i is the minimal k such that $(T-\lambda_i)^k = 0$ on V_{λ_i}.*

PROOF. Let T_i be the restriction of T to V_{λ_i}. Then $(T_i-\lambda_i)^k$ is the minimal polynomial by assumption. But $m_T(T_i) = 0$. It follows that $(x-\lambda_i)^k|m_T$ and hence that $k \leq k_i$. Conversely, since $\prod_{j\neq i}(T-\lambda_j)^{k_j}$ is invertible on V_{λ_i}, we see that $(T-\lambda_i)^{k_i} = 0$ there, so $k_i \geq k$. □

Summary of the construction so far:

- F algebraically closed field, $\dim_F V = n$, $T \in \mathrm{End}_F(V)$.
- $m_T(x) = \prod_{i=1}^r (x-\lambda_i)^{k_i}$ the minimal polynomial.
- Then $V = \bigoplus_{i=1}^r V_{\lambda_i}$ where on V_{λ_i} we have$(T-\lambda_i)^{k_i} = 0$ but $(T-\lambda_i)^{k_i-1} \neq 0$.

We now study the restriction of T to each V_{λ_i}, via the map $N = T-\lambda_i$, which is *nilpotent* of degree k_i.

DEFINITION 148. A map $N \in \mathrm{End}_F(V)$ such that $N^k = 0$ for some k is called *nilpotent*. The smallest such k is called its *degree of nilpotence*.

LEMMA 149. *Let $N \in \mathrm{End}_F(V)$ be nilpotent. Then its degree of nilpotence is at most* $\dim_F V$.

PROOF. Exercise. □

PROOF. Define subspaces V_k by $V_0 = V$ and $V_{i+1} = N(V_i)$. Then $V = V_0 \supset V_1 \cdots \supset V_i \supset \cdots$. If at any stage $V_i = V_{i+1}$ then $V_{i+j} = V_i$ for all $j \geq 1$, and in particular $V_i = \{\underline{0}\}$ (since $V_k = 0$). It follows that for $i < k$, $\dim V_{i+1} < \dim V_i$ and the claim follows. $\square$

COROLLARY 150 (Cayley–Hamilton Theorem). *Suppose F is algebraically closed. Then $m_T(x)|p_T(x)$ and, equivalently, $p_T(T) = 0$. In particular,* $\deg m_T \leq \dim_F V$.

Recall the that the *characteristic polynomial* of T is the polynomial $p_T(x) = \det(x\mathrm{Id} - T)$ of degree $\dim_F V$, and that is also has the property that $\lambda \in \mathrm{Spec}_F(T)$ iff $p_T(\lambda) = 0$.

PROOF. The linear map $x\mathrm{Id} - T$ respects the decomposition $V = \bigoplus_{i=1}^r V_{\lambda_i}$. We thus have $p_T(x) = \prod_{i=1}^r p_{T\restriction V_{\lambda_i}}(x)$. Since $p_{T\restriction V_\lambda}(x)$ has the unique root λ, it is the polynomial $(x-\lambda)^{\dim_F V_\lambda}$, so

$$p_T(x) = \prod_{i=1}^r (x - \lambda_i)^{\dim V_{\lambda_i}} .$$

Finally, k_i is the degree of nilpotence of $(T - \lambda_i)$ on V_{λ_i}. Thus $k_i \leq \dim_{\bar{F}} V_{\lambda_i}$ $\square$

We now resolve a lingering issue:

LEMMA 151. *The minimal polynomial is independent of the choice of the field. In particular, the Cayley–Hamilton Theorem holds over any field.*

PROOF. Whether $\left\{1, T, \ldots, T^{d-1}\right\} \subset \mathrm{End}_F(V)$ are linearly dependent or not does not depend on the field. $\square$

THEOREM 152 (Cayley–Hamilton). *Over any field we have $m_T(x)|p_T(x)$ or, equivalently, $p_T(T) = 0$.*

PROOF. Extend scalars to an algebraic closure. This does not change either of the polynomials m_T, p_T. $\square$

2.4. Nilpotent maps and Jordan blocks (Lectures 19-20)

2.4.1. Jordan blocks (Lecture 19). We finally turn to the problem of finding good bases for linear maps, starting with the nilpotent case. Here F can be an arbitrary field.

To start with, consider the filtration $\{0\} \subsetneq \mathrm{Ker}(N) \subsetneq \mathrm{Ker}(N^2) \subsetneq \cdots \subsetneq V = \mathrm{Ker}(N^k)$ where k is the degree of nilpotence. Choose a basis as follows:

(1) Choose a basis of size m_1 in $\mathrm{Ker}(N)$.
(2) Add m_2 vectors to get a basis of $\mathrm{Ker}(N^2) \supset \mathrm{Ker}(N)$
⋮
(3) Add m_k vectors to get a basis of $V = \mathrm{Ker}(N^k)$.

What is the matrix of N in this basis? It is block-upper triangular. On the diagonal the blocks are $m_i \times m_i$, all zero. Above each block is a block of size $m_{i-1} \times m_i$ where every column is non-zero (every if $\underline{v} \in \mathrm{Ker}(N^r) \setminus \mathrm{Ker}(N^{r-1})$ then $N\underline{v} \in \mathrm{Ker}(N^{r-1}) \setminus \mathrm{Ker}(N^{r-2})$). Above that no information.

We now try to improve this by carefully choosing the basis to encode more of the action of N.

LEMMA 153. *Let $N \in \mathrm{End}(V)$ be nilpotent. Let $B \subset V$ be a set of vectors such that $N(B) \subset B \cup \{\underline{0}\}$. Then B is linearly independent iff $B \cap \mathrm{Ker}(N)$ is.*

PROOF. One direction is clear. For the converse, let $\sum_{i=1}^{r} a_i \underline{v}_i = \underline{0}$ be a minimal dependence in B. Apply N^s where s is maximal such that for some i, $N^s \underline{v}_i \neq \underline{0}$ (perhaps $s = 0$). Then

$$\sum_{i=1}^{r} a_i N^s \underline{v}_i = \underline{0}$$

where each $N^s \underline{v}_i \in B \cup \{\}$. If some $N^s \underline{v}_i = \underline{0}$ the new dependence is shorter. Thus they are all non-zero then by the maximality of s they are all in $\mathrm{Ker}(N)$, so we obtain a dependence among $B \cap \mathrm{Ker}(N)$, a contradiction. $\square$

COROLLARY 154. *Let $N \in \mathrm{End}(V)$ and let $\underline{v} \in V$ be non-zero such that $N^k \underline{v} = \underline{0}$ for some k (wlog minimal). Then $\left\{N^i \underline{v}\right\}_{i=0}^{k-1}$ is linearly independent.*

PROOF. N is nilpotent on Span $\left\{N^i \underline{v}\right\}_{i=0}^{k-1}$, this set is invariant, and its intersection with $\mathrm{Ker} N$ is exactly $\left\{N^{k-1} \underline{v}\right\} \neq \{\underline{0}\}$. $\square$

2.4.2. Jordan canonical form for nilpotent maps (Lecture 20). Our goal is now to decompose V as a direct sum of N subspaces ("Jordan blocks") each of which has a basis as in the Corollary.

THEOREM 155 (Jordan form for nilpotent maps). *Let $N \in \mathrm{End}_F(V)$ be nilpotent. We then have a decomposition $V = \bigoplus_{j=1}^{r} V_j$ where each V_j is an N-invariant Jordan block.*

EXAMPLE 156. $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$.

- $A^2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = 0$, so A is nilpotent. The characterstic polynomial must be x^3.
- The image of A is Span $\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$. Since $A \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, Span $\left\{ \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}$ is a block.
- Taking any other vector in the kernel (say, $\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$) we get the basis $\left\{ \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right\}$ in which A has the matrix

$$\begin{pmatrix} \begin{pmatrix} 0 & 1 \\ & 0 \end{pmatrix} & \\ & (0) \end{pmatrix}.$$

PROOF. Let N have degree of nilpotence d and kernel W. For $1 \leq k \leq d$ define $W_k = \mathrm{Im}(N^k) \cap W$, so that $W_0 = W \supset W_1 \supset W_d = \{0\}$. Now choose a basis C of W compatible with this decomposition – in other words choose subsets $C_k \subset W_k$ such that $\cup_{k \geq k'} C_k$ is a basis for $W_{k'}$. Let $C = \cup_{k=0}^{d-1} C_k = \{\underline{v}_i\}_{i \in I}$ and for each i define k_i by $\underline{v}_i \in C_{k_i}$. Choose $\underline{u}_i$ such that $N^{k_i} \underline{u}_i = \underline{v}_i$, and for

$1 \le j \le k_i$ set $\underline{v}_{i,j} = N^{k_i - j}\underline{u}_i$ so that $\underline{v}_{i,1} = \underline{u}_i$ and in general $N\underline{v}_{i,j} = \begin{cases} \underline{v}_{i,j-1} & j \ge 1 \\ \underline{0} & j = 1 \end{cases}$. It is clear that $\mathrm{Span}_F \left\{\underline{v}_{i,j}\right\}_{j=1}^{k_i}$ is a Jordan block, and that $B = \left\{\underline{u}_{i,j}\right\}_{i,j}$ is a union of Jordan blocks. □

- **The set B is linearly independent**: by construction, $N(B) \subset B \cup \{\underline{0}\}$ and $B \cap W = C$ is independent.
- **The set B is spanning**: We prove by induction on $k \le d$ that $\mathrm{Span}_F(B) \supset \mathrm{Ker}(N^k)$. This is clear for $k = 0$; suppose the result for $0 \le k < d$, and let $\underline{v} \in \mathrm{Ker}(N^{k+1})$. Then $N^k\underline{v} \in W_k$, so we can write

$$\begin{aligned} N^k\underline{v} &= \sum_{i:k_i \ge k} a_i \underline{v}_i \\ &= \sum_{i:k_i \ge k} a_i N^k \left(\underline{v}_{i,k+1}\right) . \end{aligned}$$

  It follows that

$$N^k \left(\underline{v} - \sum_{i:k_i \ge k} a_i \underline{v}_{i,k+1} \right) = \underline{0} .$$

  By induction, $\underline{v} - \sum_{i:k_i \ge k} a_i \underline{v}_{i,k} \in \mathrm{Span}_F(B)$, and it follows that $\underline{v} \in \mathrm{Span}_F(B)$.

DEFINITION 157. A *Jordan basis* is a basis as in the Theorem.

LEMMA 158. *Any Jordan basis for N has exactly* $\dim_F W_{k-1} - \dim_F W_k$ *blocks of length k. Equivalently, up to permuting the blocks, N has a unique matrix in Jordan form.*

PROOF. Let $\left\{\underline{v}_{i,j}\right\}$ be a Jordan basis. Then $\mathrm{Ker}\, N = \mathrm{Span}\left\{\underline{v}_{i,1}\right\}$, while $\left\{\underline{v}_{i,j} \mid k_i \ge k,\ j \le k_i - k\right\}$ is a basis for $\mathrm{Im}(N^k)$. Clearly $\left\{\underline{v}_{i,1} \mid k_i \ge k\right\}$ then spans W_k and the claim follows. □

2.5. The Jordan canonical form (Lecture 21)

THEOREM 159 (Jordan canonical form). *Let $T \in \mathrm{End}_F(V)$ and suppose that m_T splits into linear factors in F (for example, that F is algebraically closed). Then there is a basis $\left\{\underline{v}_{\lambda,i,j}\right\}_{\lambda,i,j}$ of V such that $\left\{\underline{v}_{\lambda,i,j}\right\}_{i,j} \subset V_\lambda$ is a basis, and such that $(T-\lambda)\,\underline{v}_{\lambda,i,j} = \begin{cases} \underline{v}_{\lambda,i,j-1} & j \ge 1 \\ \underline{0} & j = 1 \end{cases}$. Furthermore, writing $W_\lambda = \mathrm{Ker}(T-\lambda)$ for the eigenspace, we have for each λ, that $1 \le i \le \dim_F W_\lambda$ and that the number of i such that $1 \le j \le k$ is exactly* $\dim_F \left((T-\lambda)^{k-1}V_\lambda \cap W_\lambda\right) - \dim_F \left((T-\lambda)^k V_\lambda \cap W_\lambda\right)$. *Equivalently, T has a unique matrix in Jordan canonical form up to permuting the blocks.*

COROLLARY 160. *The* algebraic *multiplicity of λ is* $\dim_F V_\lambda$. *The* geometric *multiplicity is the number of blocks.*

EXAMPLE 161 (Jordan forms). (1) $A_1 = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{pmatrix} = I + A$. This has characteristic polynomial $(x-1)^3$, $A_1 - I = A$ and we are back in example 156.

(2) (taken from Wikibooks:Linear Algebra) $p_B(x) = (x-6)^4$.

Let $B = \begin{pmatrix} 7 & 1 & 2 & 2 \\ 1 & 4 & -1 & -1 \\ -2 & 1 & 5 & -1 \\ 1 & 1 & 2 & 8 \end{pmatrix}$, $B' = B - 6I = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & -2 & -1 & -1 \\ -2 & 1 & -1 & -1 \\ 1 & 1 & 2 & 2 \end{pmatrix}$. Gaussian elimination shows $B' = E \begin{pmatrix} 3 & -3 & 0 & 0 \\ 1 & -2 & -1 & -1 \\ -3 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $B'^2 = \begin{pmatrix} 0 & 3 & 3 & 3 \\ 0 & 3 & 3 & 3 \\ 0 & -6 & -6 & -6 \\ 0 & 3 & 3 & 3 \end{pmatrix}$ and $B'^3 = 0$. Thus $\operatorname{Ker} B' = \{(x,y,z,w)^t \mid x = y = -(z+w)\}$ is two-dimensional. We see that the image of B'^2 is spanned by $(3,3,-6,3)^t$, which is (say) $B'(2,-1,-1,2)^t$ which (being the last column) was $B'(0,0,0,1)^t$. Another vector in the kernel is $(-1,-1,1,0)^t$, and we get the Jordan basis $\left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ -6 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\}$.

(3) $C = \begin{pmatrix} 4 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 \\ -1 & 0 & 2 & 0 \\ 4 & 0 & 1 & 2 \end{pmatrix}$ acting on $V = \mathbb{R}^4$ with $p_C(x) = (x-2)^2(x-3)^2$. Then $C - 2I = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 2 & 0 & 3 & 0 \\ -1 & 0 & 0 & 0 \\ 4 & 0 & 1 & 0 \end{pmatrix}$, $C - 3I = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & -1 & 3 & 0 \\ -1 & 0 & -1 & 0 \\ 4 & 0 & 1 & -1 \end{pmatrix}$, $(C-3I)^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -3 & 1 & -4 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & 1 \end{pmatrix}$. Thus $\operatorname{Ker}(C-2I) = \operatorname{Span}\{\underline{e}_2, \underline{e}_4\}$, which must be the 2d generalized eigenspace V_2 giving two 1×1 blocks. For $\lambda = 3$, $\operatorname{Ker}(C-3I) = \{(x,y,z,w)^t \mid z = y = -x, w = 3x\} = \operatorname{Span}\{(1,-1,-1,3)^t\}$. This isn't the whole generalized eigenspace, and

$$\operatorname{Ker}(C-3I)^2 = \left\{(x,y,z,w)^t \mid y = 3x + 4z,\, w = x - 2z\right\} = \operatorname{Span}\left\{(1,-1,-1,3)^t, (1,3,0,1)^t\right\}.$$

This must be the generalized eigenspace V_3, since it's 2d. We need to find the image of $(C-3I)[V_3]$. One vector is in the kernel, so we try the other one, and indeed $(C-3I)(1,3,0,1)^t = (1,-1,-1,3)$. This gives us a 2x2 block, so in the basis $\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\}$ the matrix has the form $\begin{pmatrix} (2) & & \\ & (2) & \\ & & \begin{pmatrix} 3 & 1 \\ & 3 \end{pmatrix} \end{pmatrix}$. Note how the image of $(C-3I)^2$ is exactly V_2 (why?)

(4) $V = \mathbb{R}^6$. $p_D(x) = t^6 + 3t^5 - 10t^3 - 15t^2 - 9t - 2 = (t+1)^5(t-2)$:

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & -1 \\ 0 & -8 & 4 & -3 & 1 & -3 \\ -3 & 13 & -8 & 6 & 2 & 9 \\ -2 & 14 & -7 & 4 & 2 & 10 \\ 1 & -18 & 11 & -11 & 2 & -6 \\ -1 & 19 & -11 & 10 & -2 & 7 \end{pmatrix}, D+I = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & -7 & 4 & -3 & 1 & -3 \\ -3 & 13 & -7 & 6 & 2 & 9 \\ -2 & 14 & -7 & 5 & 2 & 10 \\ 1 & -18 & 11 & -11 & 3 & -6 \\ -1 & 19 & -11 & 10 & -2 & 8 \end{pmatrix},$$

$$(D+I)^2 = \begin{pmatrix} 1 & -1 & 0 & 1 & -2 & -3 \\ -2 & -16 & 9 & -11 & 4 & -3 \\ -1 & 37 & -18 & 17 & 2 & 21 \\ 1 & 35 & -18 & 19 & -2 & 15 \\ -1 & -53 & 27 & -28 & 2 & -24 \\ 2 & 52 & -27 & 29 & -4 & 21 \end{pmatrix}, (D+I)^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -54 & 27 & -27 & 0 & -27 \\ 0 & 108 & -54 & 54 & 0 & 54 \\ 0 & 108 & -54 & 54 & 0 & 54 \\ 0 & -162 & 81 & -81 & 0 & -81 \\ 0 & 162 & -81 & 81 & 0 & 81 \end{pmatrix}.$$

(5) First, V_2must be a 1-dimensional eigenspace. Gaussian elimination finds the eigenvector $(01,-2,-2,3,-3)^t$. Next, V_{-1} must be 5-dimensional. Row-reduction gives: $D+I \to$ $\begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & -1/2 \\ 0 & 1 & 0 & 0 & 1 & 3/2 \\ 0 & 0 & 1 & 0 & 2 & 3/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $(D+I)^2 \to \begin{pmatrix} 2 & 0 & -1 & 3 & -4 & -5 \\ 0 & 2 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. So the $\mathrm{Ker}(D+I)$ is two-dimensional (since $(D+I)^2 \neq 0$ there will be a block of size at least 3; since $(D+I)^3$has rank one, it has the 5d kernel $V_{-1} = \{\underline{x} \mid x_3 = 2x_2 + x_4 + x_6\}$ so the largest block is 3, and so the other block must have size 2. We need a vector from the generalized eigenspace in the image of $(D+I)^2$. Since $(D+I)^3\,\underline{e}_1 = \underline{0}$ but the first column of $(D+I)^2$ is non-zero, we see that $(D+I)^2\,\underline{e}_1 = (1,-2,-1,1,-1,2)^t$ has preimage $(D+I)\,\underline{e}_1 = (1,0,-3,-2,1,-1)^t$, and we obtain our first block. Next, we need an eigenvector in the kernel and image of $D+I$, but any vector in the kernel is also in the image (no blocks of size 1), so we cam take any vector in $\mathrm{Ker}(D+I)$ independent of the one we already have. Using the row-reduced form we see that $(1,-1,-2,0,1,0)^t$ is such a vector. Then we solve

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & -7 & 4 & -3 & 1 & -3 \\ -3 & 13 & -7 & 6 & 2 & 9 \\ -2 & 14 & -7 & 5 & 2 & 10 \\ 1 & -18 & 11 & -11 & 3 & -6 \\ -1 & 19 & -11 & 10 & -2 & 8 \end{pmatrix} \underline{x} = \begin{pmatrix} 1 \\ -1 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

finding for example the vector $(1,0,-1,-1,0,0)^t$ and our second block. We conclude that in the basis $\left\{ \begin{pmatrix} 0 \\ 1 \\ -2 \\ -2 \\ 3 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ -1 \\ 1 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -3 \\ -2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\}$ the matrix has the form

$$\begin{pmatrix} (2) & & \\ & \begin{pmatrix} -1 & 1 & \\ & -1 & 1 \\ & & -1 \end{pmatrix} & \\ & & \begin{pmatrix} -1 & 1 \\ & -1 \end{pmatrix} \end{pmatrix}$$

CHAPTER 3

Vector and matrix norms

For the rest of the course our field of scalars is either $\mathbb{R}$ or $\mathbb{C}$.

3.1. Norms on vector spaces (Lecture 22)

3.1.1. Review of metric spaces.

DEFINITION 162. A *metric space* is a pair (X, d_X) where X is a set, and $d_X\colon X \times X \to \mathbb{R}_{\geq 0}$ is a function such that for all $x,y,z \in X$, $d_X(x,y) = 0$ iff $x = y$, $d_X(x,y) = d_X(y,z)$ and (the *triangle inequality*) $d_X(x,z) \leq d_X(x,y) + d_X(y,z)$.

NOTATION 163. For $x \in X$ and $r \geq 0$ we write $B_X(x,r) = \{y \in X \mid d_X(x,y) \leq r\}$ for the closed ball of radius r around x, $B_X^{\circ}(x,r) = \{y \in X \mid d_X(x,y) < r\}$ for the open ball.

DEFINITION 164. Let $(X, d_X), (Y, d_Y)$ be metric spaces and let $f\colon X \to Y$ be a function.

(1) We say f is *continuous* if $\forall x \in X : \forall \varepsilon > 0 : \exists \delta > 0 : f(B_X(x,\delta)) \subset B_Y(f(x),\varepsilon)$.
(2) We say f is *uniformly continuous* $\forall \varepsilon > 0 : \exists \delta > 0 : \forall x \in X : f(B_X(x,\delta)) \subset B_Y(f(x),\varepsilon)$.
(3) We say f is *Lipschitz continuous* if in (2) we can take $\delta = \varepsilon/L$, in other words if for all $x \neq x' \in X$,

$$d_Y\left(f(x), f(x')\right) \leq L d_X\left(x, x'\right).$$

In that case we let $\|f\|_{\mathrm{Lip}}$ denote the smallest L for which this holds.

Clearly $(3) \Rightarrow (2) \Rightarrow (1)$.

LEMMA 165. *The composition of two functions of type (1),(2),(3) is again a function of that type. In particular,* $\|f \circ g\|_{\mathrm{Lip}} \leq \|f\|_{\mathrm{Lip}} \|g\|_{\mathrm{Lip}}$.

DEFINITION 166. We call the metric space (X, d_X) *complete* if every Cauchy sequence converges.

3.1.2. Norms. Fix a vector space V.

DEFINITION 167. A *norm* on V is a function $\|\cdot\|: V \to \mathbb{R}_{\geq 0}$ such that $\|\underline{v}\| = 0$ iff $\underline{v} = \underline{0}$, $\|\alpha \underline{v}\| = |\alpha| \, \|\underline{v}\|$ and $\|\underline{u} + \underline{v}\| \leq \|\underline{u}\| + \|\underline{v}\|$. A *normed space* is a pair $(V, \|\cdot\|)$.

LEMMA 168. *Let $\|\cdot\|$ be a norm on V. Then the function $d(\underline{u},\underline{v}) = \|\underline{u} - \underline{v}\|$ is a metric.*

EXERCISE 169. The map $\| \mapsto \| \, d$ is a bijection between norms on V and metrics on V which are (1) translation-invariant $d(\underline{u},\underline{v}) = d(\underline{u}+\underline{w}, \underline{v}+\underline{w})$ and (2) 1-homogenous: $d(\alpha \underline{u}, \alpha \underline{v}) = |\alpha| \, d(\underline{u},\underline{v})$.

The restriction of a norm to a subspace is a norm.

3.1.3. Finite-dimensional examples.

EXAMPLE 170. Standard norms on $\mathbb{R}^n$ and $\mathbb{C}^n$:

(1) The *supremum norm* $\|\underline{v}\|_\infty = \max\{|v_i|\}_{i=1}^n$, parametrizing *uniform convergence*.
(2) $\|\underline{v}\|_1 = \sum_{i=1}^n |v_i|$.
(3) The *Euclidean norm* $\|\underline{v}\|_2 = \left(\sum_{i=1}^n |v_i|^2\right)^{1/2}$, connected to the *inner product* $\langle \underline{u}, \underline{v}\rangle = \sum_{i=1}^n \overline{u_i} v_i$ (prove $\triangle$ inequality from this by squaring norm of sum).
(4) For $1 < p < \infty$, $\|\underline{v}\|_p = \left(\sum_{i=1}^n |v_i|^p\right)^{1/p}$.

PROOF. These functions are clearly homogeneous, and clearly are non-zero if $\underline{v} \neq 0$; the only non-trivial part is the triangle inequality ("Minkowsky's inequality"). This is easy for $p = 1, \infty$, well-known for $p = 2$. Other cases resolved in supplement to PS8. □

EXERCISE 171. Show that $\lim_{p\to\infty} \|\underline{v}\|_p = \|\underline{v}\|_\infty$.

We have a geometric interpretation. The *unit ball* of a norm is the set $B = B(\underline{0}, 1) = \{\underline{v} \in V \mid \|\underline{v}\| \leq 1\}$. This determines the norm ($\frac{1}{\|\underline{v}\|}$ is the largest α such that $\alpha\underline{v} \in B$). Now applying a linear map to B gives a the ball of a new norm.

EXERCISE 172. Draw the unit balls for $\|\cdot\|_p$ on $\mathbb{R}^2, \mathbb{R}^3$ for $p = 1, 2, \infty$.

PROPOSITION 173 (Pullback). *Let $T \colon U \hookrightarrow V$ be an injectivel linear map. Let $\|\cdot\|_V$ be a norm on V. Then $\|\underline{u}\| \stackrel{def}{=} \|T\underline{u}\|_V$ defines a norm on U.*

PROOF. Easy check. □

3.1.4. Infinite-dimensional examples. Now the norm comes first, the space second.

EXAMPLE 174. For a set X let $\ell^\infty(X) = \left\{ f \in F^X \mid \sup\{|f(x)| : x \in X\} < \infty \right\}$, $\|f\|_\infty = \sup_{x\in X} |f(x)|$.

PROOF. The map $\|\cdot\|_\infty \colon F^X \to [0, \infty]$ satisfies the axioms of a norm, suitably extended to include the value ∞. That the set of vectors of finite norm is a subspace follows from the scaling and triangle inequalities. □

REMARK 175. A vector space with basis B can be embedded into $\ell^\infty(B)$ (we've basically seen this).

EXAMPLE 176. $\ell^p(\mathbb{N}) = \left\{ \underline{a} \in F^{\mathbb{N}} : \sum_{i=1}^\infty |a_i|^p < \infty \right\}$ with the obvious norm.

In the continuous case we a construction from earlier in the course:

DEFINITION 177. $L^p(\mathbb{R}) = \left\{ f \colon \mathbb{R} \to F \text{ [measurable]} \mid \int_{\mathbb{R}} |f(x)|^p \, dx < \infty \right\} \Big/ \{f \mid f = 0 \text{ a.e.}\}$ with the natural norm.

REMARK 178. The quotient is *essential*: for actual functions, can have $\int |f(x)|^p \, dx = 0$ without $f = 0$ exactly. In particular, elements of $L^p(\mathbb{R})$ don't have specific values.

FACT 179. *In each equivalence class in $L^p(\mathbb{R})$ there is* at most one *continuous representative.*

So part of PDE is about whether an L^p solution can be promoted to a continuous functiuon. We give an example theorem:

THEOREM 180 (Elliptic regularity). *Let $\Omega \subset \mathbb{R}^2$ be a domain, and let $f \in L^2(\Omega)$ satisfy $\Delta f = \lambda f$ distributionally:: for $g \in C_c^\infty(\Omega)$, $\int_\Omega f \Delta g = \lambda \int fg$. Then there is a smooth function $\bar{f}$ such that $\Delta f = \lambda f$ pointwise and such that $f = \bar{f}$ almost everywhere.*

3.1.5. Converges in the norm. While there are many norms on $\mathbb{R}^n$, it turns out that there is only one notion of convergence.

LEMMA 181. *Every norm on $\mathbb{R}^n$ is a continuous function.*

PROOF. Let $M = \max_i \|\underline{e}_i\|$. Then

$$\|\underline{x}\| = \left\| \sum_{i=1}^{n} x_i \underline{e}_i \right\| \leq \sum_{i=1}^{n} |x_i| \, \|\underline{e}_i\| \leq M \|\underline{x}\|_1 .$$

In particular,

$$\left| \|\underline{x}\| - \|\underline{y}\| \right| \leq \left\| \underline{x} - \underline{y} \right\| \leq M \left\| \underline{x} - \underline{y} \right\|_1 .$$

□

DEFINITION 182. Call two norms equivalent *if there are* $0 < m \leq M$ *such that* $m \|\underline{x}\| \leq \|\underline{x}\|' \leq M \|\underline{x}\|$ *holds for all* $\underline{x} \in V$.

EXERCISE 183. This is an equivalence relation. The norms are equivalent iff the same sequences of vectors satisfy $\lim_{n\to\infty} \underline{x}_n = \underline{0}$.

THEOREM 184. *All norms on $\mathbb{R}^n$ (and $\mathbb{C}^n$) are equivalent.*

PROOF. It is enough to show that they are all equivalent to $\|\cdot\|_1$. Accordingly let $\|\cdot\|$ be any other norm. Then the Lemma shows that there is M such that

$$\|\underline{x}\| \leq M \|\underline{x}\|_1 .$$

Next, the "sphere"$\{\underline{x} \mid \|\underline{x}\|_1 = 1\}$ is closed and bounded, hence compact. Accordingly let $m = \min \{\|\underline{x}\| \mid \|\underline{x}\|_1 = 1\}$. Then $m > 0$ since $\|\underline{0}\|_1 = 0 \neq 1$. Finally, for any $\underline{x} \neq 0$ we have

$$\frac{\|\underline{x}\|}{\|\underline{x}\|_1} = \left\| \frac{x}{\|\underline{x}\|_1} \right\| \geq m$$

since $\left\| \frac{x}{\|\underline{x}\|_1} \right\|_1 = 1$. It follows that

$$m \|\underline{x}\|_1 \leq q \|\underline{x}\| \leq M \|\underline{x}\|_1$$

□

3.2. Norms on matrices (Lectures 23-24)

DEFINITION 185. Let U, V be normed spaces. A map $T : U \to V$ is called *bounded* if there is $M \geq 0$ such that $\|T\underline{u}\|_V \leq M \|\underline{u}\|_U$ for all $\underline{u} \in U$. The smallest such M is called the *(operator) norm* of T.

REMARK 186. Motivation: Let U be the space of initial data for an evolution equation (say wave, or heat). Let V be the space of possible states at time t. Let T be "time evolution". Then a key part of PDE is finding norms in which T is bounded as a map from U to V. This shows that solution exist, and that they are unique.

EXAMPLE 187. The identity map has norm 1. Now consider the matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ acting on $\mathbb{R}^2$.

(1) As a map from $\ell^1 \to \ell^1$ we have

$$\left\| A \begin{pmatrix} x \\ y \end{pmatrix} \right\|_1 = |x+y| + |y| \le 2 \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_1 ,$$

with equality if $x = 0$. Thus $\|A\|_1 = 2$.

(2) Next,

$$\left\| A \begin{pmatrix} x \\ y \end{pmatrix} \right\|_2^2 = |x+y|^2 + |y|^2 \le \frac{3+\sqrt{5}}{2} \left| x^2 + y^2 \right| .$$

(3) Finally,

$$\left\| A \begin{pmatrix} x \\ y \end{pmatrix} \right\|_\infty = \max\left\{|x+y|, |y|\right\} \le 2 \max\left\{|x|, |y|\right\}$$

with equality if $x = y$, Thus $\|A\|_\infty = 2$.

EXAMPLE 188. Consider $D_x \colon C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$. This is not bounded in any norm (consider $f(x) = e^{2\pi i k x}$).

LEMMA 189. *Every map of finite-dimensional spaces is bounded.*

PROOF. Identify U with $\mathbb{R}^n$. Then the $\|\cdot\|_U$ is equivalent with $\|\cdot\|_1$, so there is A such that $\|\underline{u}\|_1 \le A \|\underline{u}\|_U$. Now the map $\underline{u} \mapsto \|T\underline{u}\|_V$ is 1-homogenous and satisfies the triangle inequality, so by the proof of Lemma 181 there is B so that $\|T\underline{u}\|_V \le B \|\underline{u}\|_1 \le (AB) \|\underline{u}\|_U$. □

LEMMA 190. *Let T, S be bounded and composable. Then ST is bounded and $\|ST\| \le \|S\| \|T\|$.*

PROOF. For any $\underline{u} \in U$, $\|ST\underline{u}\|_W \le \|S\| \|T\underline{u}\|_V \le \|S\| \|T\| \|\underline{u}\|_U$. □

PROPOSITION 191. *The operator norm is a norm on* $\mathrm{Hom}_b(U,V)$*, the space of bounded maps* $U \to V$.

PROOF. For any $S, T \in \mathrm{Hom}_b(U,V)$, $|\alpha| \|T\| + \|S\|$ is a bound for $\alpha T + S$. Since the zero map is bounded it follows that $\mathrm{Hom}_b(U,V) \subset \mathrm{Hom}(U,V)$ is a subspace, and setting $\alpha = 1$ gives the triangle inequality. If $T \ne 0$ then there is $\underline{u}$ such that $T\underline{u} \ne \underline{0}$ at which point

$$\|T\| \ge \frac{\|T\underline{u}\|}{\|\underline{u}\|} > 0 .$$

Finally, $\|(\alpha T)\underline{u}\| = |\alpha| \|T\underline{u}\| \le |\alpha| \|T\| \|\underline{u}\|$ so $\|\alpha T\| \le |\alpha| \|T\|$. But then

$$\|T\| = \left\| \frac{1}{\alpha} \alpha T \right\| \le \frac{1}{|\alpha|} \|\alpha T\|$$

gives the reverse inequality. □

3.3. Example: eigenvalues and the power method (Lecture 25)

Let A be diagonable. Want eigenvalues of A. Raising A to large powers selects the eigenvalue with largest component.

- Algorithm: multiply by A and renormalize.
- Advantage: if A sparse only need to multiply by A.
- Rate of convergence related to spectral gap.

3.4. Sequences and series of vectors and matrices (Lectures 26-27)

3.4.1. Completeness (Lecture 26).

DEFINITION 192. A metric space is complete if every Cauchy sequence in it converges.

EXAMPLE 193. $\mathbb{R}$. $\mathbb{R}^n$ in any norm. $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ (because isom to $\mathbb{R}^{mn}$).

FACT 194. *Any metric space has a* completion. *[note associated universal property and hence uniqueness]*

THEOREM 195. *Let* $(U, \|\cdot\|_U), (V, \|\cdot\|_V)$ *be normed spaces with* V *complete. Then* $\mathrm{Hom}_b(U,V)$ *is complete with respect to the operator norm.*

PROOF. Let $\{T_n\}_{n=1}^{\infty}$ be a Cauchy sequences of linear maps. For fixed $\underline{u} \in U$, the sequence $\{T_n\underline{u}\}$ is Cauchy: $\|(T_n\underline{u} - T_m\underline{u})\|_V \leq \|T_n - T_m\| \|\underline{u}\|$. It is therefore convergent – call the limit $T\underline{u}$. This is linear since $\alpha T_n\underline{u} + T_n\underline{u}'$ converges to $\alpha T\underline{u} + T\underline{u}'$ while $T_n(\alpha\underline{u} + \underline{u}')$ converges to $T(\alpha\underline{u} + \underline{u}')$.

Since $|\|T_n\| - \|T_m\|| \leq \|T_n - T_m\|$, the norms themselves are a Cauchy sequences of real numbers, in particular a convergent sequence. Now for fixed $\underline{u}$, we have $\|T\underline{u}\|_V = \lim_{n\to\infty} \|T_n\underline{u}\|_V$. We have the pointwise bound $\|T_n\underline{u}\| \leq \|T_n\| \|\underline{u}\|_U$. Passing to the limit we find

$$\|T\underline{u}\|_V \leq \left(\lim_{n\to\infty} \|T_n\|\right) \|\underline{u}\|_U$$

so T is bounded. Finally, given ε let N be such that if $m,n \geq N$ then $\|T_n - T_m\| \leq \varepsilon$. Then for any $\underline{u} \in U$,

$$\|T_n\underline{u} - T_m\underline{u}\| \leq \|T_n - T_m\| \|\underline{u}\|_U \leq \varepsilon \|\underline{u}\|_U .$$

Letting $m \to \infty$ and using the continuity of the norm, we get that if $n \geq N$ then

$$\|T_n\underline{u} - T\underline{u}\| \leq \varepsilon \|\underline{u}\|_U .$$

Since $\underline{u}$ was arbitrary this shows that $\|T_n - T\| \leq \varepsilon$ for $n \geq N$ and we are done. □

EXAMPLE 196. Let K be a compact space. Then $C(K)$, the space of continuous functions on K, is complete wrt $\|\cdot\|_\infty$.

PROOF. Continuous functions on a compact space are bounded. Let $\{f_n\}_{n=1}^{\infty} \subset C(K)$ be a Cauchy sequence. Then for fixed $x \in X$, $\{f_n(x)\}_{n=1}^{\infty} \subset \mathbb{C}$ is a Cauchy sequence, hence convergent to some $f(x) \in \mathbb{C}$. To see the convergence is in the norm, give $\varepsilon > 0$ let N be such that $\|f_n - f_m\|_\infty \leq \varepsilon$ for $n,m \geq N$. Then for any x,

$$|f_n(x) - f_m(x)| \leq \varepsilon .$$

Letting $m \to \infty$ we find for all $n \leq N$ that $|f_n(x) - f(x)| \leq \varepsilon$, that is

$$\|f_n - f\|_\infty \leq \varepsilon .$$

Finally, we need to show that f is continuous. Given $x \in X$ and $\varepsilon > 0$ let N be as above and let $n \geq N$. For any x, the continuity of f_n gives a neighbourhood of x where $|f_n(x) - f_n(y)| \leq \varepsilon$. Then

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \leq 3\varepsilon$$

in that neighbourhood, so f is continuous at x. □

EXERCISE 197. Generalize this example:

(1) Show that $\ell^\infty(X)$ is complete for any set X.

(2) For a general (topological) space X, show that $C_b(X) = C(X) \cap \ell^\infty(X)$ is complete with respect to the supremum norm.

Now fix a complete normed space V.

(3) For a set X write $\ell^\infty(X;V)$ for the space of bounded functions $X \to V$. Then $\ell^\infty(X;V)$ is complete.

(4) $C_b(X;V) = C(X;V) \cap \ell^\infty(X;V)$ is complete.

3.4.2. Series of vectors and matrices (Lecture 27). Fix a complete normed space V.

DEFINITION 198. Say the series $\sum_{n=1}^\infty \underline{v}_n$ converges *absolutely* if $\sum_{n=1}^\infty \|\underline{v}_n\|_V < \infty$.

PROPOSITION 199. *If $\sum_{n=1}^\infty \underline{v}_n$ converges absolutely it converges, and $\|\sum_{n=1}^\infty \underline{v}_n\|_V \le \sum_{n=1}^\infty \|\underline{v}_n\|_V$.*

PROOF. Standard. □

EXAMPLE 200 (Exponential series). Let V be a complete normed space and let $T \in \mathrm{End}_b(V)$. Then $\exp(T) \stackrel{\text{def}}{=} \sum_{n=0}^\infty \frac{1}{n!} T^n$ converges absolutely.

Another instance of this phenomenon:

THEOREM 201 (Weierstrass's M-test). *Let X be a (topological) space, $f_n \colon X \to V$ continuous. Suppose that we have M_n such that $\|f_n(x)\|_V \le M_n$ holds for all $x \in X$. Suppose that $M = \sum_{n=1}^\infty M_n < \infty$. Then $\sum_n f_n$ converges uniformly to a continuous function $F \colon X \to V$.*

PROOF. By the Proposition this amounts to showing that the space $C_b(X,V)$ (continuous functions $X \to V$ with $\|f(x)\|_V$ bounded) is complete with respect to the supremum norm $\|f\|_\infty = \sup\{\|f(x)\|_V : x \in X\}$. □

We will apply this to power series of matrices.

EXAMPLE 202. Let $\|\cdot\|$ be some operator norm on $M_n(\mathbb{R})$, and let $A \in M_n(\mathbb{R})$. For $0 < R < \frac{1}{\|A\|}$ (any $R > 0$ if $A = 0$) and $z \in \mathbb{C}$ with $|z| \le R$ consider the series

$$\sum_{n=0}^\infty z^n A^n .$$

We have $\|A^n\| \le \|A\|^n$ (operator norm!) so that $\|z^n A^n\| \le (R\|A\|)^n$. Since $\sum_{n=0}^\infty (R\|A\|)^n$ converges, we see that our series converges and the sum is continuous in z (and in A). Taking the union we get convergence in $|z| < \frac{1}{\|A\|}$. The limit is $(\mathrm{Id} - zA)^{-1}$ (incidentally showing this is invertible).

REMARK 203. In fact, the radius of convergence is $\frac{1}{\rho(A)}$.

3.4.3. Vector-valued limits and derivatives. We recall facts about vector-valued limits.

LEMMA 204 (Limit arithmetic). *Let U,V,W be normed spaces. Let $\underline{u}_i(x) \colon X \to U$, $\alpha_i(x) \colon X \to F$, $T(x) \colon X \to \mathrm{Hom}_b(U,V)$, $S(x) \colon X \to \mathrm{Hom}_b(V,W)$. Then, in each case supposing the limits on the right exist, the limits on the left exist and equality holds:*

(1) $\lim_{x\to x_0} (\alpha_1(x)\underline{u}_1(x) + \alpha_2(x)\underline{u}_2(x)) = (\lim_{x\to x_0} \alpha_1(x))(\lim_{x\to x_0} \underline{u}_1(x)) + (\lim_{x\to x_0} \alpha_2(x))(\lim_{x\to x_0} \underline{u}_2(x))$.

(2) $\lim_{x\to x_0} T(x)\underline{u}(x) = (\lim_{x\to x_0} T(x))(\lim_{x\to x_0} \underline{u}(x))$.

(3) $\lim_{x\to x_0} S(x)T(x) = (\lim_{x\to x_0} S(x))(\lim_{x\to x_0} T(x))$.

PROOF. Same as in $\mathbb{R}$, replacing $|\cdot|$ with $\|\cdot\|_V$. □

We can also differentiate vector-valued functions (see Math 320 for details)

DEFINITION 205. Let $X \subset \mathbb{R}^n$ be open. Say that $f\colon X \to V$ is *strongly differentiable* at x_0 if there is a bounded linear map $L\colon \mathbb{R}^n \to V$ such that

$$\lim_{h \to \underline{0}} \frac{\|f(x_0+h) - f(x_0) - Lh\|_V}{\|h\|_{\mathbb{R}^n}} = 0\,.$$

In that case we write $Df(x_0)$ for L.

It is clear that differentiability at x_0 implies continuity at x_0.

LEMMA 206 (Derivatives). *Let U,V,W be normed spaces. Let $\underline{u}_i(x)\colon X \to U$, $T(x)\colon X \to \mathrm{Hom}_b(U,V)$, $S(x)\colon X \to \mathrm{Hom}_b(V,W)$ be differentiable at x_0. Then the derivatives on the left exist and take the following values:*

(1) $D(\underline{u}_1 + \underline{u}_2)(x_0) = D\underline{u}_1(x_0) + D\underline{u}_2(x_0)$.
(2) $D(T\underline{u})(x_0)(\underline{h}) = (DT(x_0)(\underline{h}) \cdot \underline{u}(x_0)) + T(x_0) \cdot D\underline{u}(x_0)(\underline{h})$.
(3) $D(ST)(x_0)(\underline{h}) = (DS(x_0)(\underline{h}) \cdot T(x_0)) + (S(x_0) \cdot DT(x_0)(\underline{h}))$.

PROOF. Same as in $\mathbb{R}$, replacing $|\cdot|$ with $\|\cdot\|_V$. □

3.5. The exponential series (Lecture 28)

We apply Weierestrass's M-test to *power series.*

THEOREM 207. *Let X be a (topological) space, $f_n\colon X \to V$ continuous. Suppose that we have M_n such that $\|f_n(x)\|_V \leq M_n$ holds for all $x \in X$ with $M = \sum_{n=1}^{\infty} M_n < \infty$. Then $\sum_n f_n$ converges uniformly to a continuous function $F\colon X \to V$ and $\|F(x)\|_V \leq M$ for all $x \in X$.*

COROLLARY 208. *Let V be a complete normed space, and let $\sum_n a_n z^n$ be a power series with radius of convergence R. Then for any $A \in \mathrm{End}_b(V)$, $\sum_n a_n A^n$ converges absolutely if $\|A\| < R$, uniformly in $\{\|A\| \leq R - \varepsilon\}$*

PROOF. Let $X = V = \mathrm{End_b}(V)$, $f_n(A) = a_n A^n$, so that $\|f_n(A)\| \leq |a_n|\,\|A\|^n$. For $T < R$ we have $\sum_n |a_n| T^n < \infty$ and hence uniform convergence in $\{\|A\| \leq T\}$. □

We therefore fix a normed space V, and and plug matrices $A \in \mathrm{End_b}(V)$ into power series.

EXAMPLE 209. $\exp(A) = \sum_k \frac{A^k}{k!}$ converges everywhere.

REMARK 210. We'll look at two kinds of matrix-valued series:

(1) Power series with matrix coefficients: $f(t) = \sum_{n=0}^{\infty} A_n t^n$. Here, t is a scalar and $A_n \in \mathrm{End}_b(V)$.
(2) Plugging in matrices into power series: given $f(z) = \sum_{n=0}^{\infty} a_n z^n$ set $f(A) = \sum_{n=0}^{\infty} a_n A^n$.

3.5.1. Basic properties.

LEMMA 211. $\exp(tA)\exp(sA) = \exp((t+s)A)$.

PROOF. The series converge absolutely, so the product converges in any order. We thus have

$$\begin{aligned}\exp(tA)\exp(sA) &= \left(\sum_{k=0}^{\infty}\frac{(tA)^k}{k!}\right)\left(\sum_{l=0}^{\infty}\frac{(sA)^l}{l!}\right)=\sum_{k,l}\frac{t^k s^\ell A^{k+\ell}}{k!\ell!}\\ &= \sum_{m=0}^{\infty}\sum_{k+l=m}\frac{t^k s^\ell A^{k+\ell}}{k!\ell!}=\sum_{m=0}^{\infty}\frac{A^m}{m!}\sum_{k+l=m}\frac{m!}{k!\ell!}t^k s^\ell\\ &= \sum_{m=0}^{\infty}\frac{A^m}{m!}(t+s)^m=\exp\left((t+s)A\right).\end{aligned}$$

□

Recap: multiplication of absolutely convergent series.

LEMMA 212. *Let $\sum_n \underline{u}_n$ converge absolutely, Then it converges in any reordering and its sum is unchanged.*

PROOF. Let $\sigma\in S_{\mathbb{N}}$. Given $N>0$ let $K>\max\sigma^{-1}\left(\{0,1,\ldots,N\}\right)$. Then

$$\begin{aligned}\left\|\sum_{k=0}^{K}\underline{u}_{\sigma(k)}-\sum_{n=0}^{N}\underline{u}_n\right\| &= \left\|\sum_{\substack{0\le k\le K\\ \sigma(k)>N}}\underline{u}_{\sigma(k)}\right\|\\ &\le \sum_{\substack{0\le k\le K\\ \sigma(k)>N}}\left\|\underline{u}_{\sigma(k)}\right\|\\ &\le \sum_{n>N}\|\underline{u}_n\|.\end{aligned}$$

Now given $\varepsilon>0$ let N be large enough such that $\sum_{n>N}\|\underline{u}_n\|<\varepsilon$ (exists by absolute convergence). Then for K large enough as above,

$$\begin{aligned}\left\|\sum_{k=0}^{K}\underline{u}_{\sigma(k)}-\sum_{n=0}^{\infty}\underline{u}_n\right\| &\le \left\|\sum_{n>N}\underline{u}_n\right\|+\sum_{n>N}\|\underline{u}_n\|\\ &\le 2\sum_{n>N}\|\underline{u}_n\|\le 2\varepsilon.\end{aligned}$$

□

PROPOSITION 213. *Let $A=\sum_n a_n$, $B=\sum_m b_m$ be convergent series of positive real numbers. Then $\sum_{n,m}a_nb_m$ converges to AB.*

PROOF. Let $S\subset\mathbb{N}^2$ be finite so that $\sum_{(n,m)\in S}a_nb_m$ is a partial sum. Then for N large enough we have $S\subset\{0,1,\ldots,N\}^2$ so that

$$\sum_{(m,n)\in S}a_nb_m\le\sum_{0\le n,m\le N}a_nb_m=\left(\sum_{n\le N}a_n\right)\left(\sum_{m\le N}b_m\right)\le AB$$

and the series converges. To evaluate the limit it's enough to note that the subsequence of partial sums

$$\sum_{0\le n,m\le N} a_n b_m = \left(\sum_{n\le N} a_n\right)\left(\sum_{m\le N} b_m\right)$$

evidently converges to AB. □

THEOREM 214. *Let $A=\sum_{n=0}^{\infty} A_n$, $B=\sum_{m=0}^{\infty} B_m$ be absolutely convergent ($A_n \in \mathrm{Hom}_b(V,W)$, $B_m \in \mathrm{Hom}_b(U,V)$). Then $\sum_{m,n\ge 0} A_n B_m$ converges absolutely to AB.*

PROOF. Since $\|A_n B_m\| \le \|A_n\|\,\|B_m\|$ absolute convergence follows from the Proposition and the convergence of $\sum_n \|A_n\|$ and $\sum_m \|B_m\|$. To evaluate the sum we again take the "square" partial sums:

$$\sum_{0\le n,m\le N} A_n B_m = \left(\sum_{n\le N} A_n\right)\left(\sum_{m\le N} B_m\right) \xrightarrow[N\to\infty]{} AB\,.$$

□

3.5.2. Differentiation and application to constant-coefficient differential equations.

COROLLARY 215. $\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA) = A\exp(tA) = \exp(tA)A$.

PROOF. At $t=0$ we have $\frac{\exp(hA)-\mathrm{Id}}{h} = A + \sum_{k=1}^{\infty}\frac{h^k}{(k+1)!}A^{k+1}$ and

$$\left\|\sum_{k=1}^{\infty}\frac{h^k}{(k+1)!}A^{k+1}\right\| \le \sum_{k=1}^{\infty}\frac{|h|^k}{(k+1)!}\|A\|^{k+1} \le \frac{\exp(|h|\,\|A\|-1-\|A\|\,|h|}{|h|} \xrightarrow[h\to 0]{} 0\,.$$

In general we have

$$\frac{\exp\left((t+h)A\right)-\exp(tA)}{h} = \exp(tA)\frac{\exp(hA)-\mathrm{Id}}{h} \xrightarrow[h\to 0]{} \exp(tA)A\,.$$

That $A\exp(tA)=\exp(tA)A$ follows from considering partial sums. □

Consider the system of differential equations

$$\begin{cases}\frac{\mathrm{d}}{\mathrm{d}t}\underline{v}(t) = A\underline{v}(t)\\ \underline{v}(0)=\underline{v}_0\end{cases}$$

where A is a bounded map.

PROPOSITION 216. *The system has the unique solution $\underline{v}(t)=\exp(At)\underline{v}_0$.*

PROOF. We saw $\frac{\mathrm{d}}{\mathrm{d}t}\exp(At)\underline{v}_0 = A\left(\exp(At)\underline{v}_0\right)$. Conversely, suppose $\underline{v}(t)$ is any solution. Then

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-At}\underline{v}(t)\right) &= \left(e^{-At}(-A)\right)(\underline{v}(t)) + \left(e^{-At}\right)(A\underline{v}(t))\\ &= e^{-At}(-A+A)\,\underline{v}(t) = 0\,.\end{aligned}$$

It remains to prove: □

LEMMA 217. *Let $f\colon [0,1]\to V$ be differentiable. If $f'(t)=0$ for all t then f is constant.*

PROOF. Suppose $f(t_0) \neq f(0)$. Let $\varphi \in V'$ be a *bounded* linear functional such that $\varphi(f(t_0) - f(0)) \neq 0$. Then $\varphi \circ f\colon [0,1] \to \mathbb{R}$is differentiable and its derivative is 0:

$$\lim_{h\to 0} \frac{\varphi(f(t+h)) - \varphi(f(t))}{h} = \lim_{h\to 0} \varphi\left(\frac{f(t+h)-f(t)}{h}\right) = \varphi\left(\lim_{h\to 0} \frac{f(t+h)-f(t)}{h}\right) = \varphi(f'(t)).$$

But $(\varphi \circ f)(t_0) - (\varphi \circ f)(0) = \varphi(f(t_0) - f(0)) \neq 0$, a contradiction. □

REMARK 218. If V is finite-dimensional, every linear functional is bounded. If V is infinite-dimensional the existence of φ is a serious fact ("Hahn–Banach Theorem")

Now consider a *linear ODE with constant coefficients*:

$$\begin{cases} \frac{d^n}{dt^n} u(t) = \sum_{k=0}^{n-1} a_k u^{(k)}(t) & \\ u^{(k)}(0) = w_k & 0 \leq k \leq n-1\,. \end{cases}$$

We solve this system via the auxilliary vector

$$\underline{v}(t) = \left(u(t), u'(t), \cdots, u^{(n-1)}(t)\right).$$

We then have

$$\frac{d\underline{v}(t)}{dt} = A\underline{v}$$

where A is the *companion matrix*

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}.$$

(companion to the polynomial $x^n - \sum_{k=0}^{n-1} a_k x^k$). It follows that

$$\underline{v}(t) = e^{At}\underline{w}.$$

Idea: bring A to Jordan form so easier to take exponential.

3.6. Invertibility and the resolvent (Lecture 29)

Say we have a matrix A we'd like to invert. Idea: write $A = D + E$ where we know to invert D. Then $A = D(I + D^{-1}E)$, so if $\left\|D^{-1}E\right\| < 1$ we have

$$\left(I + D^{-1}E\right)^{-1} = \sum_{n=0}^{\infty} \left(-D^{-1}E\right)^n$$

and

$$A^{-1} = \sum_{n=0}^{\infty} \left(-D^{-1}E\right)^n D^{-1}$$

(in particular, A is invertible).

3.6.1. Application: Gauss-Seidel and Jacobi iteration.

3.6.2. Application: the resolvent. Let V be a complete normed space. Let T be an operator on V. Define the *resolvent set* of T to be the set of $z \in \mathbb{C}$for which $T - z\text{Id}$ has a bounded inverse. Define the *spectrum* $\sigma(T)$ to be the complement of the resolvent set. This contains the actual eigenvalues (λ such that $\text{Ker}(T-\lambda)$ is non-trivial) but also λ where $T-\lambda$ is not surjective, and λ where an inverse to $T-\lambda$ exists but is unbounded).

THEOREM 219. *The resolvent set is open, and the function ("resolvent function") $\rho(T) \to \text{End}_b(V)$ given by $z \mapsto R(z) = (z\text{Id} - T)^{-1}$ is holomorphic.*

PROOF. Suppose $z_0 - T$ has a bounded inverse. We need to invert $z - T$ for z close to z_0. Indeed, if $|z - z_0| < \frac{1}{\left\|(z_0-T)^{-1}\right\|}$ then

$$-\sum_{n=0}^{\infty} (T - z_0)^{n+1} (z - z_0)^n$$

converges and furnishes the requisite inverse. It is evidently holomorphic in z in the indicated ball. □

EXAMPLE 220. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain with nice boundary, $\Delta = \frac{d^2}{dx^2} + \frac{d^2}{dy^2}$ the Laplace operator (say defined on $f \in C^\infty(\Omega)$ vanishing on the boundary). Then Δ is unbounded, but its resolvent is nice. For example, $R(i\varepsilon)$ only has eigenvalues. It follows that the spectrum of Δ consists of eigenvalues, that is for $\lambda \in \sigma(\Delta)$ there is $f \in L^2(\Omega)$ with $\Delta f = \lambda f$ (and $f \in C^\infty$ by elliptic regularity).

3.7. Holomorphic calculus

DEFINITION 221. Let $f(z) = \sum_n a_n z^n$. Define $f(A) = \sum_{n=0}^{\infty} a_n A^n$.

LEMMA 222. $Sf(A)S^{-1} = f(SAS^{-1})$.

PROPOSITION 223. $(f \circ g)(A) = f(g(A))$ *if it all works.*

THEOREM 224. $\det(\exp(A)) = \exp(\text{Tr}(A))$.

CHAPTER 4

Vignettes

Sketches of applications of linear algebra to group theory.
Key Idea: *linearization* – use linear tools to study non-linear objects.

4.1. The exponential map and structure theory for $\mathrm{GL}_n(\mathbb{R})$

Our goal is to understand the (topologically) closed subgroups of $G = \mathrm{GL}_n(\mathbb{R})$.
Idea: to a subgroup H assign the logarithms of the elements of H. If H was commutative this would be a subspace.

DEFINITION 225. $\mathrm{Lie}(H) = \{X \in M_n(\mathbb{R}) \mid \forall t : \exp(tX) \in H\}$.

REMARK 226. Clearly this is invariant under scaling. In fact, enough to take small t, and even just a sequence of t tending to zero (since $\{t \mid \exp(tX) \in H\}$ is a closed subgroup of $\mathbb{R}$).

THEOREM 227. $\mathrm{Lie}(H)$ *is a subspace of* $M_n(\mathbb{R})$*, closed under* $[X,Y]$.

PROOF. For $t \in \mathbb{R}$and $m \in \mathbb{Z}_{\geq 1}$, $\left(\exp\left(\frac{tX}{m}\right)\exp\left(\frac{tY}{m}\right)\right)^m = \left(\mathrm{Id} + \frac{tX+tY}{m} + O(\frac{1}{m^2})\right)^m \xrightarrow[m\to\infty]{} \exp\left(tX + tY\right)$. Thus If $X,Y \in \mathrm{Lie}(H)$ then also $X+Y \in \mathrm{Lie}(H)$. □

- Classify subgroups of G containing A by action on Lie algebra and finding eigenspaces.

4.2. Representation Theory of Groups

EXAMPLE 228 (Representations). (1) Structure of $\mathrm{GL}_n(\mathbb{R})$: let A act on $M_n(\mathbb{R})$.
(2) M manifold, G acting on M, thus acting on $H_k(M)$ and $H^k(M)$.
(3) *Angular momentum*: $O(3)$ acting by rotation on $L^2\left(\mathbb{R}^3\right)$.