Math 412: Advanced Linear Algebra
Lecture Notes

Lior Silberman

These are rough notes for the Spring 2014 course. Solutions to problem sets were posted on an internal website.

Contents

Introduction

Lior Silberman, `lior@Math.UBC.CA`, `http://www.math.ubc.ca/~lior`
Office: Math Building 229B
Phone: 604-827-3031

0.1. Administrivia

- Problem sets will be posted on the course website.
  - To the extent I have time, solutions may be posted on Connect.
  - If you create solution sets I'll certainly post them.
- Textbooks
  - Halmos
  - Algebra books

0.2. 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 1 (Euler). $v - e + f = 1$.

PROOF. Arbitrarily orient the edges. Let $\partial_E \colon \mathbb{R}^E \to \mathbb{R}^V$ 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 2. *∂_F is injective.*

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

LEMMA 3. $\operatorname{Ker} \partial_E = \operatorname{Im} \partial_F$.

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 4. $\operatorname{Im}(\partial_E)$ *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 5. Using $\mathbb{F}_2$ coefficients is even simpler.

0.3. Course plan (subject to revision) (Lecture 1, 6/1/13)

- Quick review
- Some constructions
  - Direct sum and product
  - Spaces of homomorphisms and duality
  - Quotient vector spaces
  - Multilinear algebra: Tensor products
- Structure theory for linear maps
  - Gram–Schmidt, Polar, Cartan
  - The Bruhat decompositions and LU, $LL^\dagger$ factorization; numerical applications
  - The minimal polynomial and the Cayley–Hamilton Theorem
  - The Jordan canonical form
- Analysis with vectors and matrices
  - Norms on vector spaces
  - Operator norms
  - Matrices in power series: e^{tX} and its friends.
- Other topics if time permits.

0.4. Review

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

DEFINITION 6 (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.

EXAMPLE 7. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$. $\mathbb{F}_2$ (via addition and multiplication tables; ex: show this is a field), $\mathbb{F}_p$.

EXERCISE 8. 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 9. 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 10. $0_F\cdot\underline{v}=\underline{0}$ *for all* $\underline{v}$.

PROOF. $0\underline{v}=(0+0)\underline{v}=0\underline{v}+0\underline{v}$. Now subtract $0\underline{v}$ from both sides. $\square$

0.4.2. 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 $\{a_i\}_{i=1}^r\subset F$ 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.*

0.4.3. Examples. $\{\underline{0}\}$, $\mathbb{R}^n$, F^X.

Math 412: Problem Set 1 (due 15/1/2014)

Practice problems, any sub-parts marked "OPT" (optional) and supplementary problems are not for submission.

Practice problems

P1 Show that the map $f\colon \mathbb{R}^3 \to \mathbb{R}$ given by $f(x,y,z) = x - 2y + z$ is a linear map. Show that the maps $(x,y,z) \mapsto 1$ and $(x,y,z) \mapsto x^2$ are non-linear.

P2 Let F be a field, X a set. Carefully show that pointwise addition and scalar multiplication endow the set F^X of functions from X to F with the structure of an F-vectorspace.

For submission

RMK The following idea will be used repeatedly during the course to prove that sets of vectors are linearly independent. Make sure you understand how this argument works.

1. Let V be a vector space, $S \subset V$ a set of vectors. A *minimal dependence* in S is an equality $\sum_{i=1}^{m} a_i \underline{v}_i = \underline{0}$ where $\underline{v}_i \in S$ are distinct, a_i are scalars not all of which are zero, and $m \geq 1$ is as small as possible so that such $\{a_i\}$, $\{\underline{v}_i\}$ exist.
   (a) Find a minimal dependence among $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3$.
   (b) Show that in a minimal dependence the a_i are all non-zero.
   (c) Suppose that $\sum_{i=1}^{m} a_i \underline{v}_i$ and $\sum_{i=1}^{m} b_i \underline{v}_i$ are minimal dependences in S, involving the exact same set of vectors. Show that there is a non-zero scalar c such that $a_i = c b_i$.
   (d) Let $T\colon V \to V$ be a linear map, and let $S \subset V$ be a set of (non-zero) eigenvectors of T, each corresponding to a distinct eigenvalue. Applying T to a minimal dependence in S obtain a contradiction to (b) and conclude that S is actually linearly independent.
   (**e) Let Γ be a group. The set $\mathrm{Hom}(\Gamma, \mathbb{C}^\times)$ of group homomorphisms from Γ to the multiplicative group of nonzero complex numbers is called the set of *quasicharacters* of Γ (the notion of "character of a group" has an additional, different but related meaning, which is not at issue in this problem). Show that $\mathrm{Hom}(\Gamma, \mathbb{C}^\times)$ is linearly independent in the space $\mathbb{C}^\Gamma$ of functions from Γ to $\mathbb{C}$.

2. Let $S = \{\cos(nx)\}_{n=0}^{\infty} \cup \{\sin(nx)\}_{n=1}^{\infty}$, thought of as a subset of the space $C(-\pi,\pi)$ of continuous functions on the interval $[-\pi,\pi]$.
   (a) Applying $\frac{d}{dx}$ to a putative minimal dependence in S obtain a different linear dependence of at most the same length, and use that to show that S is, in fact, linearly independent.
   (b) Show that the elements of S are an orthogonal system with respect to the inner product $\langle f,g \rangle = \int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x$ (feel free to look up any trig identities you need). This gives a different proof of their independence.
   (c) Let $W = \mathrm{Span}_{\mathbb{C}}(S)$ (this is usually called "the space of trigonometric polynomials"; a typical element is $5 - \sin(3x) + \sqrt{2}\cos(15x) - \pi\cos(32x)$). Find a ordering of S so that the matrix of the linear map $\frac{d}{dx}\colon W \to W$ in that basis has a simple form.

3. (Matrices associated to linear maps) Let V, W be vector spaces of dimensions n, m respectively. Let $T \in \operatorname{Hom}(V, W)$ be a linear map from V to W. Show that there are ordered bases $B = \{\underline{v}_j\}_{j=1}^n \subset V$ and $C = \{\underline{w}_i\}_{i=1}^m \subset W$ and an integer $d \leq \min\{n, m\}$ such that the matrix $A = (a_{ij})$ of T with respect to those bases satisfies $a_{ij} = \begin{cases} 1 & i = j \leq d \\ 0 & \text{otherwise} \end{cases}$, that is has the form

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

(Hint1: study some examples, such as the matrices $\begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$ and $\begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}$) (Hint2: start your solution by choosing a basis for the image of T).

Extra credit: Finite fields

4. Let F be a field.
   (a) Define a map $\iota\colon (\mathbb{Z}, +) \to (F, +)$ by mapping $n \in \mathbb{Z}_{\geq 0}$ to the sum $1_F + \cdots + 1_F$ n times. Show that this is a group homomorphism.
   DEF If the map ι is injective we say that F is of *characteristic zero*.
   (b) Suppose there is a non-zero $n \in \mathbb{Z}$ in the kernel of ι. Show that the smallest positive such number is a prime number p.
   DEF In that case we say that F is of *characteristic p*.
   (c) Show that in that case ι induces an isomorphism between the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ and a subfield of F. In particular, there is a unique field of p elements up to isomorphism.

5. Let F be a field with finitely many elements.
   (a) Carefully endow F with the structure of a vector space over $\mathbb{F}_p$ for an appropriately chosen p.
   (b) Show that there exists an integer $r \geq 1$ such that F has p^r elements.
   RMK For every prime power $q = p^r$ there is a field $\mathbb{F}_q$ with q elements, and two such fields are isomorphic. They are usually called *finite fields*, but also *Galois fields* after their discoverer.

Supplementary Problems I: A new field

A. Let $\mathbb{Q}(\sqrt{2})$ denote the set $\left\{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\right\} \subset \mathbb{R}$.
   (a) Show that $\mathbb{Q}(\sqrt{2})$ is a $\mathbb{Q}$-subspace of $\mathbb{R}$.
   (b) Show that $\mathbb{Q}(\sqrt{2})$ is two-dimensional as a $\mathbb{Q}$-vector space. In fact, identify a basis.
   (*c) Show that $\mathbb{Q}(\sqrt{2})$ is a field.
   (**d) Let V be a vector space over $\mathbb{Q}(\sqrt{2})$ and suppose that $\dim_{\mathbb{Q}(\sqrt{2})} V = d$. Show that $\dim_{\mathbb{Q}} V = 2d$.

Supplementary Problems II: How physicists define vectors

Fix a field F.

B. (The general linear group)
   (a) Let $\mathrm{GL}_n(F)$ denote the set of invertible $n\times n$ matrices with coefficients in F. Show that $\mathrm{GL}_n(F)$ forms a group with the operation of matrix multiplication.
   (b) For a vector space V over F let $\mathrm{GL}(V)$ denote the set of invertible linear maps from V to itself. Show that $\mathrm{GL}(V)$ forms a group with the operation of composition.
   (c) Suppose that $\dim_F V = n$ Show that $\mathrm{GL}_n(F)\simeq \mathrm{GL}(V)$ (hint: show that each of the two group is isomorphic to $\mathrm{GL}(F^n)$.

C. (Group actions) Let G be a group, X a set. An *action* of G on X is a map $\cdot\colon G\times X\to X$ such that $g\cdot(h\cdot x) = (gh)\cdot x$ and $1_G\cdot x = x$ for all $g,h\in G$ and $x\in X$ (1_G is the identity element of G).
   (a) Show that matrix-vector multiplication $(g,\underline{v})\mapsto g\underline{v}$ defines an action of $G=\mathrm{GL}_n(F)$ on $X=F^n$.
   (b) Let V be an n-dimensional vector space over F, and let $\mathcal{B}$ be the set of ordered bases of V. For $g\in \mathrm{GL}_n(F)$ and $B=\{\underline{v}_i\}_{i=1}^{\dim V}\in\mathcal{B}$ set $gB = \left\{\sum_{j=1}^n g_{ij}\underline{v}_i\right\}_{j=1}^n$. Check that $gB\in\mathcal{B}$ and that $(g,B)\mapsto gB$ is an action of $\mathrm{GL}_n(F)$ on $\mathcal{B}$.
   (c) Show that the action is *transitive*: for any $B,B'\in\mathcal{B}$ there is $g\in\mathrm{GL}_n(F)$ such that $gB=B'$.
   (d) Show that the action is *simply transitive*: that the g from part (b) is unique.

D. (From the physics department) Let V be an n-dimensional vector space, and let $\mathcal{B}$ be its set of bases. Given $\underline{u}\in V$ define a map $\phi_{\underline{u}}\colon\mathcal{B}\to F^n$ by setting $\phi_{\underline{u}}(B)=\underline{a}$ if $B=\{\underline{v}_i\}_{i=1}^n$ and $\underline{u}=\sum_{i=1}^n a_i\underline{v}_i$.
   (a) Show that $\alpha\phi_{\underline{u}}+\phi_{\underline{u}'}=\phi_{\alpha\underline{u}+\underline{u}'}$. Conclude that the set $\left\{\phi_{\underline{u}}\right\}_{\underline{u}\in V}$ forms a vector space over F.
   (b) Show that the map $\phi_{\underline{u}}\colon\mathcal{B}\to F^n$ is *equivariant* for the actions of B(a),B(b), in that for each $g\in\mathrm{GL}_n(F)$, $B\in\mathcal{B}$, $g\left(\phi_{\underline{u}}(B)\right)=\phi_{\underline{u}}(gB)$.
   (c) Physicists define a "covariant vector" to be an equivariant map $\phi\colon\mathcal{B}\to F^n$. Let Φ be the set of covariant vectors. Show that the map $\underline{u}\mapsto\phi_{\underline{u}}$ defines an isomorphism $V\to\Phi$. (Hint: define a map $\Phi\to V$ by fixing a basis $B=\{\underline{v}_i\}_{i=1}^n$ and mapping $\phi\mapsto\sum_{i=1}^n a_i\underline{v}_i$ if $\phi(B)=\underline{a}$).
   (d) Physicists define a "contravariant vector" to be a map $\phi\colon\mathcal{B}\to F^n$ such that $\phi(gB) = {}^tg^{-1}\cdot(\phi(B))$. Verify that $(g,\underline{a})\mapsto {}^tg^{-1}\underline{a}$ defines an action of $\mathrm{GL}_n(F)$ on F^n, that the set Φ' of contravariant vectors is a vector space, and that it is naturally isomorphic to the dual vector space V' of V.

Supplementary Problems III: Fun in positive characteristic

E. Let F be a field of characteristic 2 (that is, $1_F+1_F=0_F$).
   (a) Show that for all $x,y\in F$ we have $x+x=0_F$ and $(x+y)^2=x^2+y^2$.
   (b) Considering F as a vector space over $\mathbb{F}_2$ as in 5(a), show that the map given by $\mathrm{Frob}(x)=x^2$ is a linear map.
   (c) Suppose that the map $x\mapsto x^2$is actually F-linear and not only $\mathbb{F}_2$-linear. Show that $F=\mathbb{F}_2$.
   RMK Compare your answer with practice problem 1.

F. (This problem requires a bit of number theory) Now let F have characteristic $p>0$. Show that the *Frobenius endomorphism* $x\mapsto x^p$ is $\mathbb{F}_p$-linear.

CHAPTER 1

Constructions

Fix a field F.

1.1. Direct sum, direct product

1.1.1. Simplest case (Lecture 2, 8/1/2014).

CONSTRUCTION 13 (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 14. *This really is a vector space.*

REMARK 15. The Lemma serves to review the definition of vector space.

PROOF. Every property follows from the respective properties of U,V. $\square$

REMARK 16. More generally, can take the direct sum of groups.

LEMMA 17. $\dim_F(U \oplus V) = \dim_F U + \dim_F V$.

REMARK 18. 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 $U \oplus V$. $\square$

EXAMPLE 19. $\mathbb{R}^n \oplus \mathbb{R}^m \simeq \mathbb{R}^{n+m}$.

(Lecture 3, 10/1/2014) A key situation is when U,V are subspaces of an "ambient" vector space W.

LEMMA 20. *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. $\square$

DEFINITION 21. The space in the previous lemma is called the *sum* of U,V and denoted $U+V$.

LEMMA 22. *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. $\square$

PROPOSITION 23 (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 22. It is surjective by Lemma 20. 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 17. $\square$

REMARK 24. This was a review of that formula.

DEFINITION 25 (Internal direct sum). We say the sum is *direct* if f is an isomorphism.

THEOREM 26. *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, $U \cap V=\operatorname{Ker} f$.
(2) $\Rightarrow$ (3): the first assumption gives existence, the second uniqueness.
(3) $\Rightarrow$ (1): existence says f is surjective, uniqueness says f is injective. □

1.1.2. Finite direct sums (Lecture 4, 13/1/2013). 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 27. 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 28. This generalizes the notion of "linear independence" from vectors to subspaces.

LEMMA 29. *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. □

PROPOSITION 30. *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. Construct. □

REMARK 31. (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 by induction.

DEFINITION 32. Abstract arbitrary direct sum.

- Block diagonality.
- Block upper-triangularity. [first structural result].

1.2. Quotients (Lecture 5, 15/1/2015)

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 33. *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. $\square$

Properties persist.

1.3. Hom spaces and duality

1.3.1. Hom spaces (Lecture 5 continued).

DEFINITION 34. $\operatorname{Hom}_F(U,V)$ will denote the space of F-linear maps $U \to V$.

LEMMA 35. $\operatorname{Hom}_F(U,V) \subset V^U$ *is a subspace, hence a vector space.*

DEFINITION 36. $V' = \operatorname{Hom}_F(V,F)$ 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.

Math 412: Problem Set 2 (due 22/1/2014)

Practice

P1 Let $\{V_i\}_{i\in I}$ be a family of vector spaces, and let $A_i \in \mathrm{End}(V_i) = \mathrm{Hom}(V_i, V_i)$.
  (a) Show that there is a unique element $\bigoplus_{i\in I} A_i \in \mathrm{End}(\bigoplus_{i\in I} V_i)$ whose restriction to the image of V_i in the sum is A_i.
  (b) Carefully show that the matrix of $\bigoplus_{i\in I} A_i$ in an appropriate basis is block-diagonal.

P2 Construct a vector space W and three subspaces U, V_1, V_2 such that $W = U \oplus V_1 = U \oplus V_2$ (internal direct sums) but $V_1 \neq V_2$.

Direct sums

1. Give an example of $V_1, V_2, V_3 \subset W$ where $V_i \cap V_j = \{\underline{0}\}$ for every $i \neq j$ yet the sum $V_1 + V_2 + V_3$ is not direct.
2. Let $\{V_i\}_{i=1}^r$ be subspaces of W with $\sum_{i=1}^r \dim(V_i) > (r-1)\dim W$. Show that $\bigcap_{i=1}^r V_i \neq \{\underline{0}\}$.
3. (Diagonability)
   (a) Let $T \in \mathrm{End}(V)$. For each $\lambda \in F$ let $V_\lambda = \mathrm{Ker}\,(T-\lambda)$. Let $\mathrm{Spec}_F(T) = \{\lambda \in F \mid V_\lambda \neq \{0\}\}$ be the set of eigenvalues of T. Show that the sum $\sum_{\lambda\in\mathrm{Spec}_F(T)} V_\lambda$ is direct (the sum equals V iff T is diagonable).
   (b) Show that a square matrix $A \in M_n(F)$ is diagonable over F iff there exist n one-dimensional subspaces $V_i \subset F^n$ such $F^n = \bigoplus_{i=1}^n V_i$ and $A(V_i) \subset V_i$ for all i.

Quotients

4. Let $\mathfrak{sl}_n(F) = \{A \in M_n(F) \mid \mathrm{Tr}A = 0\}$ and let $\mathfrak{pgl}_n(F) = M_n(F)/F \cdot I_n$ (matrices modulu scalar matrices). Suppose that n is invertible in F (equivalently, that the characteristic of F does not divide n). Show that the quotient map $M_n(F) \to \mathfrak{pgl}_n(F)$ restricts to an isomorphism $\mathfrak{sl}_n(F) \to \mathfrak{pgl}_n(F)$.
5. Recall our axiom that every vector space has a basis.
   (a) Show[1] that every linearly independent set in a vector space is contained in a basis.
   (b) Let $U \subset W$. Show that there exists another subspace V such that $W = U \oplus V$.
   (c) Let $W = U \oplus V$, and let $\pi\colon W \to W/U$ be the quotient map. Show that the restriction of W to V is an isomorphism. Conclude that if $U \oplus V_1 \simeq U \oplus V_2$ then $V_1 \simeq V_2$ (c.f. problem P2)
6. (Structure of quotients) Let $V \subset W$ with quotient map $\pi\colon W \to W/V$.
   (a) Show that mapping $U \mapsto \pi(U)$ gives a bijection between (1) the set of subspaces of W containing V and (2) the set of subspaces of W/V.
   (b) (The universal property) Let Z be another vector spaces. Show that $f \mapsto f \circ \pi$ gives a linear bijection $\mathrm{Hom}\,(W/V, Z) \to \{g \in \mathrm{Hom}(W,Z) \mid V \subset \mathrm{Ker}\,g\}$.

[1] Directly, without using any form of transfinite induction

7. For $f\colon \mathbb{R}^n \to \mathbb{R}$ the *Lipschitz constant* of f is the (possibly infinite) number

$$\|f\|_{\mathrm{Lip}} \stackrel{\mathrm{def}}{=} \sup\left\{\frac{|f(x)-f(y)|}{|x-y|} \mid x,y\in\mathbb{R}^n, x\neq y\right\}.$$

Let $\mathrm{Lip}(\mathbb{R}^n) = \left\{f\colon \mathbb{R}^n\to\mathbb{R} \mid \|f\|_{\mathrm{Lip}} < \infty\right\}$ be the space of *Lipschitz functions*.

PRA Show that $f\in\mathrm{Lip}(\mathbb{R}^n)$ iff there is C such that $|f(x)-f(y)|\leq C|x-y|$ for all $x,y\in\mathbb{R}^n$.

(a) Show that $\mathrm{Lip}(\mathbb{R}^n)$ is a vector space.

(b) Let $\mathbb{1}$ be the constant function 1. Show that $\|f\|_{\mathrm{Lip}}$ descends to a function on $\mathrm{Lip}(\mathbb{R}^n)/\mathbb{R}\mathbb{1}$.

(c) For $\bar{f}\in\mathrm{Lip}(\mathbb{R}^n)/\mathbb{R}\mathbb{1}$ show that $\left\|\bar{f}\right\|_{\mathrm{Lip}} = 0$ iff $\bar{f}=0$.

Supplement: Infinite direct sums and products

CONSTRUCTION. *Let $\{V_i\}_{i\in I}$ be a (possibly infinite) family of vector spaces.*

(1) *The* direct product $\prod_{i\in I} V_i$ *is the vector space whose underlying space is* $\{f\colon I\to\bigcup_{i\in I}V_i \mid \forall i : f(i)\in V_i\}$ *with the operations of pointwise addition and scalar multiplication.*

(2) *The* direct sum $\bigoplus_{i\in i} V_i$ *is the subspace* $\left\{f\in\prod_{i\in I}V_i \mid \#\left\{i \mid f(i)\neq \underline{0}_{V_i}\right\}<\infty\right\}$ *of finitely supported functions.*

A. (Tedium)

(a) Show that the direct product is a vector space

(b) Show that the direct sum is a subspace.

(c) Let $\pi_i\colon \prod_{i\in I}V_i\to V_i$ be the projection on the ith coordinate ($\pi_i(f)=f(i)$). Show that this is a surjective linear map.

(d) Let $\sigma_i\colon V_i\to\prod_{i\in I}V_i$ be the map such that $\sigma_i(\underline{v})(j) = \begin{cases}\underline{v} & j=i\\ \underline{0} & j\neq i\end{cases}$. Show that σ_i is an injective linear map.

B. (Meat) Let Z be another vector space.

(a) Show that $\bigoplus_{i\in I}V_i$ is the internal direct sum of the images $\sigma_i(V_i)$.

(b) Suppose for each $i\in I$ we are given $f_i\in\mathrm{Hom}(V_i,Z)$. Show that there is a unique $f\in\mathrm{Hom}(\bigoplus_{i\in I}V_i)$ such that $f\circ\sigma_i = f_i$.

(c) You are instead given $g_i\in\mathrm{Hom}(Z,V_i)$. Show that there is a unique $g\in\mathrm{Hom}(Z,\prod_i V_i)$ such that $\pi_i\circ g = g_i$ for all i.

C. (What a universal property can do) Let S be a vector space equipped with maps $\sigma_i'\colon V_i\to S$, and suppose the property of 5(b) holds (for every choice of $f_i\in\mathrm{Hom}(V_i,Z)$ there is a unique $f\in\mathrm{Hom}(S,Z)$ ...)

(a) Show that each σ_i' is injective (hint: take $Z=V_j$, f_j the identity map, $f_i=0$ if $i\neq j$).

(b) Show that the images of the σ_i' span S.

(c) Show that S is the internal direct sum of the S_i.

(d) (There is only one direct sum) Show that there is a unique isomorphism $\varphi\colon S\to\bigoplus_{i\in I}V_i$ such that $\varphi\circ\sigma_i' = \sigma_i$ (hint: construct φ by assumption, and a reverse map using the existence part of 5(b); to see that the composition is the identity use the uniqueness of the assumption and of 5(b), depending on the order of composition).

D. Now let P be a vector space equipped with maps $\pi_i'\colon P\to V_i$ such that 5(c) holds.

(a) Show that π_i' are surjective.

(b) Show that there is a unique isomorphism $\psi\colon\ : P\to\prod_{i\in I}V_i$ such that $\pi_i\circ\psi = \pi_i'$.

Supplement: universal properties

E. A *free abelian group* is a pair (F,S) where F is an abelian group, $S \subset F$, and ("universal property") for any abelian group A and any (set) map $f \colon S \to A$ there is a unique group homomorphism $\bar{f} \colon G \to A$ such that $\bar{f}(s) = f(s)$ for any $s \in S$. The size $\#S$ is called the *rank* of the free abelian group.
  (a) Show that $(\mathbb{Z}, \{1\})$ is a free abelian group.
  (b) Show that $\left(\mathbb{Z}^d, \{\underline{e}_k\}_{k=1}^d\right)$ is a free abelian group.
  (c) Let $(F,S), (F',S')$ be free abelian groups and let $f \colon S \to S'$ be a bijection. Show that f extends to a unique isomorphism $\bar{f} \colon F \to F'$.
  (d) Let (F,S) be a free abelian group. Show that S generates F.
  (e) Show that every element of a free abelian group has infinite order.

Supplement: Lipschitz functions

DEFINITION. Let $(X,d_X), (Y,d_Y)$ be metric spaces, and let $f \colon X \to Y$ be a function. We say f is a *Lipschitz function* (or is "Lipschitz continuous") if for some C and for all $x,x' \in X$ we have

$$d_Y\left(f(x), f(x')\right) \le C d_X\left(x,x'\right).$$

Write $\mathrm{Lip}(X,Y)$ for the space of Lipschitz continuous functions, and for $f \in \mathrm{Lip}(X,Y)$ write $\|f\|_{\mathrm{Lip}} = \sup\left\{\frac{d_Y(f(x),f(x'))}{d_X(x,x')} \mid x \neq x' \in X\right\}$ for its *Lipschitz constant*.

F. (Analysis)
  (a) Show that Lipschitz functions are continuous.
  (b) Let $f \in C^1(\mathbb{R}^n;\mathbb{R})$. Show that $\|f\|_{\mathrm{Lip}} = \sup\{|\nabla f(x)| : x \in \mathbb{R}^n\}$.
  (c) Show that $\|\alpha f + \beta g\|_{\mathrm{Lip}} \le |\alpha| \|f\|_{\mathrm{Lip}} + |\beta| \|g\|_{\mathrm{Lip}}$ ("$\|\cdot\|_{\mathrm{Lip}}$ is a seminorm").
  (d) Show that $D\left(\bar{f},\bar{g}\right) = \left\|\bar{f} - \bar{g}\right\|_{\mathrm{Lip}}$ defines a metric on $\mathrm{Lip}(\mathbb{R}^n;\mathbb{R})/\mathbb{R}\mathbb{1}$.
  (e) Generalize (a),(c),(d) to the case of $\mathrm{Lip}(X,\mathbb{R})$ where X is any metric space.
  (f) Show that $\mathrm{Lip}(X,\mathbb{R})/\mathbb{R}\mathbb{1}$ is complete for all metric spaces X.

1.3.2. The dual space, finite dimensions (Lecture 6, 17/1/2014).

CONSTRUCTION 37 (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 38. *These are linear functionals.*

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

EXAMPLE 39. $V = F^n$ 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 40. Alternative construction: φ_i is the unique linear map to F satisfying $\varphi_i(\underline{b}_j) = \delta_{i,j}$.

LEMMA 41. *The dual basis is linearly independent. If $\dim_F V < \infty$ it is spanning.*

PROOF. Evaluate a linear combination at $\underline{b}_j$. $\square$

REMARK 42. 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. Also, the proof used* evaluation – *let's investigate that more.*

PROPOSITION 43 (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. $\square$

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.

1.3.3. The dual space, infinite dimensions (Lecture 7, 20/1/2014). Interaction with past constructions: $(V/U)' \subset V'$ (PS2),

LEMMA 44. $(U \oplus V)' \simeq U' \oplus V'$.

PROOF. Universal property. $\square$

COROLLARY 45. *Since $(F)' \simeq F$, it follows by induction that $(F^n)' \simeq F^n$.*

What about infinite sums?

- The universal property gives a bijection $(\bigoplus_{i\in I} V_i)' \longleftrightarrow \times_{i\in I} V_i'$.
- Any Cartesian product $\times_{i\in I} W_i$ 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 $\alpha\,(w_i)_{i\in I} + (w'_i)_{i\in I} \stackrel{\text{def}}{=} (\alpha w_i + w'_i)_{i\in I}$. 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 $\left(\bigoplus_{i\in I} V_i\right)' \longleftrightarrow \prod_{i\in I} V'_i$ 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 $\{\underline{w} \in \prod_i W_i \mid \operatorname{supp}(\underline{w}) \text{ is finite}\}$.
  - 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 46. $\mathbb{R}^{\oplus\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$.

COROLLARY 47. *The dual of an infinite-dimensional space is much bigger than the sum of the duals, and the double dual is bigger yet.*

1.3.4. Question: so 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 (close to partial sums), not algebraic properties.

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

LEMMA 48 (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 functionals.*

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 can generalize limits just apply that to partial sums of a series).

DEFINITION 49. Let $c \subset \ell^{\infty} \subset \mathbb{R}^{\mathbb{N}}$ be the sets of convergent, respectively bounded sequences.

LEMMA 50. $c \subset \ell^{\infty}$ *are subspaces, and* $\lim_{n\to\infty}\colon c \to \mathbb{R}$ *is a linear functional.*

EXAMPLE 51. 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 52. *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 53. *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) $\lim_{\omega}(a_n b_n) = (\lim_{\omega} a_n)(\lim_{\omega} b_n)$.

1.3.5. The invariant pairing (Lecture 8, 22/1/2014).

- Pairing $V \times V'$
- Bilinear forms in general, equivalence to maps $V \to U'$.
- Pairing $F^n \times F^m$ via matrix; matrix representation of general bilinear form ("Gram matrix")
- Non-degeneracy: Duality = non-degen-pairing = isom to dual
- Identification of duals using pairings: Riesz representation theorems for $C(X)'$, $\mathcal{H}'$.

1.3.6. The dual of a linear map (Lecture 9, 24/1/2014).

CONSTRUCTION 54. *Let* $T \in \mathrm{Hom}(U,V)$. *Set* $T' \in \mathrm{Hom}(V',U')$ *by* $(T'\varphi)(\underline{v}) = \varphi(T\underline{v})$.

LEMMA 55. *This is a linear map* $\mathrm{Hom}(U,V) \to \mathrm{Hom}(V',U')$. *An isomorphism of* U,V *finite-dimensional.*

LEMMA 56. $(TS)' = S'T'$

Math 412: Problem Set 3 (due 29/1/2014)

Practice

P1 Let $\underline{u}_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $\underline{u}_2 = \begin{pmatrix}1\\1\\-1\end{pmatrix}$, $\underline{u}_3 = \begin{pmatrix}1\\-1\\-1\end{pmatrix}$, $\underline{u} = \begin{pmatrix}0\\1\\0\end{pmatrix}$ as vectors in $\mathbb{R}^3$.

(a) Construct an explicit linear functional $\varphi \in \left(\mathbb{R}^3\right)'$ vanishing on $\underline{u}_1, \underline{u}_2$.

(b) Show that $\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$ is a basis on $\mathbb{R}^3$ and find its dual basis.

(c) Evaluate the dual basis at $\underline{u}$.

P2 Let V be n-dimensional and let $\{\varphi_i\}_{i=1}^m \in V'$.

(a) Show that if $m < n$ there is a non-zero $\underline{v} \in V$ such that $\varphi_i(\underline{v}) = 0$ for all i. Interpret this as a statement about linear equations.

(b) When is it true that for each $\underline{x} \in F^m$ there is $\underline{v} \in V$ such that for all i, $\varphi_i(\underline{v}) = x_i$?

Banach limits

Recall that $\ell^\infty \subset \mathbb{R}^{\mathbb{N}}$denote the set of *bounded* sequences (the sequences $\underline{a}$ such that for some M we have $|a_i| \leq M$ for all i). Let $S\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ be the *shift* map $(S\underline{a})_n = \underline{a}_{n+1}$. A subspace $U \subset \mathbb{R}^{\mathbb{N}}$ is *shift-invariant* if $S(U) \subset U$. If U is shift-invariant a function F with domain U is called *shift-invariant* if $F \circ S = F$ (example: the subset $c \subset \mathbb{R}^{\mathbb{N}}$ of convergent sequences is shift-invariant, as is the functional $\lim\colon c \to \mathbb{R}$ assigning to every sequence its limit).

P3 (Useful facts)

(a) Show that ℓ^∞ is a subspace of $\mathbb{R}^{\mathbb{N}}$.

(b) Show that $S\colon \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$ is linear and that $S(\ell^\infty) = \ell^\infty$.

(c) Let $U \subset \mathbb{R}^{\mathbb{N}}$ be a shift-invariant subspace. Show that the set $U_0 = \{S\underline{a} - \underline{a} \mid \underline{a} \in U\}$ is a subspace of U.

(d) In the case $U = \mathbb{R}^{\oplus\mathbb{N}}$ of sequences of finite support, show that $U_0 = U$.

(e) Let Z be an auxiliary vector space. Show that $F \in \operatorname{Hom}(U, Z)$ is shift-invariant iff F vanishes on U_0.

1. Let $W = \{S\underline{a} - \underline{a} \mid \underline{a} \in \ell^\infty\} \subset \ell^\infty$. Let $\mathbb{1}$ be the sequences everywhere equal to 1.

(a) Show that the sum $W + \mathbb{R}\mathbb{1} \subset \ell^\infty$ is direct and construct an S-invariant functional $\varphi\colon \ell^\infty \to \mathbb{R}$ such that $\varphi(\mathbb{1}) = 1$.

(b) (Strengthening) For $\underline{a} \in \ell^\infty$ set $\|\underline{a}\|_\infty = \sup_n |a_n|$. Show that if $\underline{a} \in W$ and $x \in R$ then $\|\underline{a} + x\mathbb{1}\|_\infty \geq |x|$. (Hint: consider the average of the first N entries of the vector $\underline{a} + x\mathbb{1}$).

SUPP Let $\varphi \in (\ell^\infty)'$ be shift-invariant, positive, and satisfy $\varphi(\mathbb{1}) = 1$. Show that $\liminf_{n\to\infty} a_n \leq \varphi(\underline{a}) \leq \limsup_{n\to\infty} a_n$ and conclude that the restriction of φ to c is the usual limit.

2. ("choose one") Let $\varphi \in (\ell^\infty)'$ satisfy $\varphi(\mathbb{1}) = 1$. Let $\underline{a}$ be the sequence $a_n = \frac{1+(-1)^n}{2}$.

(a) Suppose that φ is shift-invariant. Show that $\varphi(\underline{a}) = \frac{1}{2}$.

(b) Suppose that φ respects pointwise multiplication (if $z_n = x_n y_n$ then $\varphi(\underline{z}) = \varphi(\underline{x})\varphi(\underline{y})$). Show that $\varphi(\underline{a}) \in \{0, 1\}$.

Duality and bilinear forms

3. (The dual map) Let U,V,W be vector spaces, and let $T \in \mathrm{Hom}(U,V)$, and let $S \in \mathrm{Hom}(V,W)$.
   (a) (The abstract meaning of transpose) Suppose U,V be finite-dimensional with bases $\left\{\underline{u}_j\right\}_{j=1}^{m} \subset U$, $\{\underline{v}_i\}_{i=1}^{n} \subset V$, and let $A \in M_{n,m}(F)$ be the matrix of T in those bases. Show that the matrix of the dual map $T' \in \mathrm{Hom}(V',U')$ with respect to the dual bases $\left\{\underline{u}'_j\right\}_{j=1}^{m} \subset U'$, $\{\underline{v}'_i\}_{i=1}^{n} \subset V'$ is the transpose ${}^{\mathrm{t}}A$.
   (b) Show that $(ST)' = T'S'$. It follows that ${}^{\mathrm{t}}(AB) = {}^{\mathrm{t}}B\,{}^{\mathrm{t}}A$.

4. Let $F^{\oplus\mathbb{N}}$ denote the space of sequences of finite support. Construct a non-degenerate pairing $F^{\oplus\mathbb{N}} \times F^{\mathbb{N}} \to F$, giving a concrete realization of $\left(F^{\oplus\mathbb{N}}\right)'$.

5. Let $C_c^\infty(\mathbb{R})$ be the space of smooth functions on $\mathbb{R}$ with compact support, and let $D\colon C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$ be the differentiation operator $\frac{\mathrm{d}}{\mathrm{d}x}$. For a reasonable function f on $\mathbb{R}$ define a functional φ_f on $C_c^\infty(\mathbb{R})$ by $\varphi_f(g) = \int_{\mathbb{R}} fg\,\mathrm{d}x$ (note that f need only be integrable, not continuous).
   (a) Show that if f is continuously differentiable then $D'\varphi_f = \varphi_{-Df}$.
   DEF For this reason one usually extends the operator D to the dual space by $D\varphi = -D'\varphi$, thus giving a notion of a "derivative" for non-differentiable and even discontinuous functions.
   (b) Let the "Dirac delta" $\delta \in C_c^\infty(\mathbb{R})'$ be the evaluation functional $\delta(f) = f(0)$. Express $(D\delta)(f)$ in terms of f.
   (c) Let φ be a linear functional such that $D'\varphi = 0$. Show that for some constant c, $\varphi = \varphi_{c\mathbb{1}}$.

Supplement: The support of distributions

A. (This is a problem in analysis unrelated to 412) Let $\varphi \in C_c^\infty(\mathbb{R})'$.
   DEF Let $U \subset \mathbb{R}$ be open. Say that φ is *supported away from* U if for any $f \in C_c^\infty(U)$, $\varphi(f) = 0$. The *support* $\mathrm{supp}(\varphi)$ is the complement the union of all such U.
   (a) Show that $\mathrm{supp}(\varphi)$ is closed, and that φ is supported away from $\mathbb{R} \setminus \mathrm{supp}(\varphi)$.
   (b) Show that $\mathrm{supp}(\delta) = \{0\}$ (see problem 5(b)).
   (c) Show that $\mathrm{supp}(D\varphi) \subset \mathrm{supp}(\varphi)$ (note that this is well-known for functions).
   (d) Show that $D\delta$ is not of the form φ_f for any function f.

1.4. Multilinear algebra and tensor products

1.4.1. Multilinear forms (Lecture 9, 24/1/2014).

DEFINITION 57. 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 58 (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 59. A *multilinear 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 60. 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 58(3): Fix bases $\{\underline{u}_i\},\{\underline{v}_j\}$. Then

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

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

$$B = \sum_{i,j} B\left(\underline{u}_i,\underline{v}_j\right)\varphi_i\psi_j\,. \tag{1.4.1}$$

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

PROPOSITION 61. *The set* $\{\varphi_i\psi_j\}_{i,j}$ *is a basis of the space of bilinear forms* $U\times V\to F$.

COROLLARY 62. *The space of bilinear forms on* $U\times V$ *has dimension* $\dim_F U\cdot\dim_F V$.

REMARK 63. Also works in infinite dimensions, since can have the sum (1.4.1) be infinite – every pair of vectors only has finite support in the respective bases.

1.4.2. The tensor product (Lecture 10, 27/1/2014). 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 64 (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 65. *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 66. *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: for $\underline{u} \in B_U$ let $\left\{\varphi_{\underline{u}}\right\}_{\underline{u} \in B_U} \subset U'$, $\left\{\psi_{\underline{v}}\right\}_{\underline{v} \in B_V} \subset V'$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 $\left\{\widetilde{\varphi_{\underline{u}}\psi_{\underline{v}}}\right\}_{(\underline{u}, \underline{v}) \in B_U \times B_V}$ are dual bases. □

COROLLARY 67. $\dim_F (U \otimes V) = \dim_F U \cdot \dim_F V$.

1.4.3. (Lecture 11, 29/1/2014).

1.4.4. Symmetric and antisymmetric tensor products (Lecture 13, 3/2/2014). In this section we assume $\mathrm{char}(F) = 0$.

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

LEMMA 68. *Let $T \in \mathrm{End}_F(U)$ satisfy $T^2 = \mathrm{Id}$. Then U is the direct sum of the two eigenspaces.*

DEFINITION 69. $\mathrm{Sym}^2 V$ and $\bigwedge^2 V$ are the eigenspaces.

PROPOSITION 70. *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

1.4.5. Bases of Sym^k, $\bigwedge^k$, determinants (Lecture 14, 5/2/2014).

PROPOSITION 71. *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 72. *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.*

- Exterior forms of top degree and determinants.

Math 412: Problem Set 4 (due 7/2/2014)

Practice

P1. Let U,V be vector spaces and let $U_1 \subset U$, $V_1 \subset V$ be subspaces.

(a) "Naturally" embed $U_1 \otimes V_1$ in $U \otimes V$.

(b) Is $(U \otimes V)/(U_1 \otimes V_1)$ isomorphic to $(U/U_1) \otimes (V/V_1)$?

P2. Let $(\cdot,\cdot)$ be a non-degenerate bilinear form on a finite-dimensional vector space U, defined by the isomorphism $g\colon U \to U'$ such that $(\underline{u},\underline{v}) = (g\underline{u})(\underline{v})$.

(a) For $T \in \mathrm{End}(U)$ define $T^\dagger = g^{-1}T'g$ where T' is the dual map. Show that $T^\dagger \in \mathrm{End}(U)$ satisfies $(\underline{u}, T\underline{v}) = \left(T^\dagger \underline{u}, \underline{v}\right)$ for all $\underline{u},\underline{v} \in V$.

(b) Show that $(TS)^\dagger = S^\dagger T^\dagger$.

(c) Show that the matrix of $T^\dagger$ wrt an $(\cdot,\cdot)$-orthonormal basis is the transpose of the matrix of T in that basis.

Bilinear forms

In problems 1,2 we assume 2 is invertible in F, and fix F-vector spaces V,W.

1. (Alternating pairings and symplectic forms) Let V,W be vector spaces, and let $[\cdot,\cdot] : V \times V \to W$ be a bilinear map.

(a) Show that $(\forall \underline{u},\underline{v} \in V : [\underline{u},\underline{v}] = -[\underline{v},\underline{u}]) \leftrightarrow (\forall \underline{u} \in V : [\underline{u},\underline{u}] = 0)$ (Hint: consider $\underline{u}+\underline{v}$).

DEF A form satisfying either property is *alternating*. We now suppose $[\cdot,\cdot]$ is alternating.

(b) The *radical* of the form is the set $R = \{\underline{u} \in V \mid \forall \underline{v} \in V : [\underline{u},\underline{v}] = 0\}$. Show that the radical is a subspace.

(c) The form $[\cdot,\cdot]$ is called *non-degenerate* if its radical is $\{\underline{0}\}$. Show that setting $[\underline{u}+R, \underline{v}+R] \stackrel{\text{def}}{=} [\underline{u},\underline{v}]$ defines a non-degenerate alternating bilinear map $(V/R) \times (V/R) \to W$.

RMK Note that you need to justify each claim, starting with "defines".

2. (Darboux's Theorem) Suppose now that V is finite-dimensional, and that $[\cdot,\cdot] : V \times V \to F$ is a non-degenerate alternating form.

DEF The *orthogonal complement* of a subspace $U \subset V$ is a set $U^\perp = \{\underline{v} \in V \mid \forall \underline{u} \in U : [\underline{u},\underline{v}] = 0\}$.

(a) Show that $U^\perp$ is a subspace of V.

(b) Show that the restriction of $[\cdot,\cdot]$ to U is non-degenerate iff $U \cap U^\perp = \{\underline{0}\}$.

(*c) Suppose that the conditions of (b) hold. Show that $V = U \oplus U^\perp$, and that the restriction of $[\cdot,\cdot]$ to $U^\perp$ is non-degenerate.

(d) Let $\underline{u} \in V$ be non-zero. Show that there is $\underline{u}' \in V$ such that $[\underline{u},\underline{u}'] \neq 0$. Find a basis $\{\underline{u}_1,\underline{v}_1\}$to $U = \mathrm{Span}\{\underline{u},\underline{u}'\}$ in which the matrix of $[\cdot,\cdot]$ is $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(e) Show that $\dim_F V = 2n$ for some n, and that V has a basis $\{\underline{u}_i,\underline{v}_i\}_{i=1}^n$ in which the matrix of $[\cdot,\cdot]$ is block-diagonal, with each 2×2 block of the form from (d).

RECAP Only even-dimensional spaces have non-degenerate alternating forms, and up to choice of basis, there is only one such form.

Tensor product

3. (Preliminary step) Let U,V be finite-dimensional.

(a) Construct a natural isomorphism $\mathrm{End}(U \otimes V) \to \mathrm{Hom}(U, U \otimes \mathrm{End}(V))$.

(b) Generalize this to a natural isomorphism $\mathrm{Hom}(U \otimes V_1, U \otimes V_2) \to \mathrm{Hom}(U, U \otimes \mathrm{Hom}(V_1,V_2))$.

4. Let U,V be vector spaces with U finite-dimensional, and let $A \in \mathrm{Hom}(U, U \otimes V)$. Given a basis $\left\{\underline{u}_j\right\}_{j=1}^{\dim U}$ of U let $\underline{v}_{ij} \in V$ be defined by $A\underline{u}_j = \sum_i \underline{u}_i \otimes \underline{v}_{ij}$ and define $\mathrm{Tr} A = \sum_{i=1}^{\dim U} \underline{v}_{ii}$. Show that this definition is independent of the choice of basis.

5. (Inner products) Let U,V be inner product spaces (real scalars, say).
   (a) Show that $\langle u_1 \otimes v_1, u_2 \otimes v_2\rangle_{U\otimes V} \stackrel{\mathrm{def}}{=} \langle u_1, u_2\rangle_U \langle v_1, v_2\rangle_V$ extends to an inner product on $U \otimes V$.
   (b) Let $A \in \mathrm{End}(U)$, $B \in \mathrm{End}(V)$. Show that $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$ (for a definition of the adjoint see practice problem P2).
   (c) Let $P \in \mathrm{End}(U \otimes V)$, interpreted as an element of $\mathrm{Hom}(U, U \otimes \mathrm{End}(V))$ as in 1(b). Show that $(\mathrm{Tr}_U P)^\dagger = \mathrm{Tr}_U \left(P^\dagger\right)$.
   (*d) [Thanks to J. Karczmarek] Let $\underline{w} \in U \otimes V$ be non-zero, and let $P_{\underline{w}} \in \mathrm{End}(U \otimes V)$ be the orthogonal projection on $\underline{w}$. It follows from 3(a) that $\mathrm{Tr}_U P_{\underline{w}} \in \mathrm{End}(V)$ and $\mathrm{Tr}_V P_{\underline{w}} \in \mathrm{End}(U)$ are both Hermitian. Show that their non-zero eigenvalues are the same.

Supplementary problems

A. (Extension of scalars) Let $F \subset K$ be fields. Let V be an F-vectorspace.
   (a) Considering K as an F-vectorspace (see PS1), we have the tensor product $K \otimes_F V$ (the subscript means "tensor product as F-vectorspaces"). For each $x \in K$ defining a $x(\alpha \otimes \underline{v}) \stackrel{\mathrm{def}}{=} (x\alpha) \otimes \underline{v}$. Show that this extends to an F-linear map $K \otimes_F V \to K \otimes_F V$ giving $K \otimes_F V$ the structure of a K-vector space. This construction is called "extension of scalars"
   (b) Let $B \subset V$ be a basis. Show that $\{1 \otimes \underline{v}\}_{\underline{v}\in B}$ is a basis for $K \otimes_F V$ as a K-vectorspace. Conclude that $\dim_K (K \otimes_F V) = \dim_F V$.
   (c) Let $V_N = \mathrm{Span}_{\mathbb{R}} \left(\{1\} \cup \{\cos(nx), \sin(nx)\}_{n=1}^N\right)$. Then $\frac{d}{dx}\colon V_N \to V_N$ is not diagonable. Find a different basis for $\mathbb{C} \otimes_{\mathbb{R}} V_N$ in which $\frac{d}{dx}$ is diagonal. Note that the elements of your basis are not "pure tensors", that is not of the form $af(x)$ where $a \in \mathbb{C}$and $f = \cos(nx)$ or $f = \sin(nx)$.

B. DEF: An *F-algebra* is a triple $(A, 1_A, \times)$ such that A is an F-vector space, $(A, 0_A, 1_A+, \times)$ is a ring, and (compatibility of structures) for any $a \in F$ and $x,y \in A$ we have $a \cdot (x \times y) = (a \cdot x) \times y$. Because of the compatibility from now on we won't distinguish the multiplication in A and scalar multiplication by elements of F.
   (a) Verify that $\mathbb{C}$ is an $\mathbb{R}$-algebra, and that $M_n(F)$ is an F-algebra for all F.
   (b) More generally, verify that if R is a ring, and $F \subset R$ is a subfield then R has the structure of an F-algebra. Similarly, that $\mathrm{End}_F(V)$ is an F-algebra for any vector space V.
   (c) Let A,B be F-algebras. Give $A \otimes_F B$ the structure of an F-algebra.
   (d) Show that the map $F \to A$ given by $a \mapsto a \cdot 1_A$ gives an embedding of F-algebars $F \hookrightarrow A$.
   (e) (Extension of scalars for algebras) Let K be an extension of F. Give $K \otimes_F A$ the structure of a K-algebra.
   (f) Show that $K \otimes_F \mathrm{End}_F(V) \simeq \mathrm{End}_K (K \otimes_F V)$.

C. The *center* $Z(A)$ of a ring is the set of elements that commute with the whole ring.
   (a) Show that the center of an F-algebra is an F-subspace, containing the subspace $F \cdot 1_A$.
   (b) Show that the image of $Z(A) \otimes Z(B)$ in $A \otimes B$ is exactly the center of that algebra.

Math 412: Supplementary Problem Set on Categories

In the second half of the 20^{th} century it became clear that, in some sense, it is the *functions* that are important in the theory of an algebraic structure more than the structures themselves. This has been formalized in *Category Theory*, and the categorical point of view has been underlying much of the constructions in 412. Here's a taste of the ideas.

First examples

1. Some ideas of linear algebra can be expressed purely in terms of linear maps.
   (a) Show that $\{\underline{0}\}$ is the unique (up to isomorphism) vector space U such that for all vector spaces Z, $\mathrm{Hom}_F(U,Z)$ is a singleton.
   (b) Show that $\{\underline{0}\}$ is the unique (up to isomorphism) vector space U such that for all vector spaces Z, $\mathrm{Hom}_F(Z,U)$ is a singleton.
   (c) Let $f \in \mathrm{Hom}_F(U,V)$. Show that f is injective iff for all vector spaces Z, and all $g_1, g_2 \in \mathrm{Hom}(Z,U)$, $f \circ g_1 = f \circ g_2$ iff $g_1 = g_2$.
   (d) Let $f \in \mathrm{Hom}_F(U,V)$. Show that f is surjective iff for all vector spaces Z, and all $g_1, g_2 \in \mathrm{Hom}(V,Z)$, $g_1 \circ f = g_2 \circ f$ iff $g_1 = g_2$.
   (e) Show that $U \oplus V$ (the vector standard space structure on $U \times V$, together with the map $\underline{u} \mapsto (\underline{u},\underline{0})$and $\underline{v} \mapsto (\underline{0},\underline{v})$) has the property that for any vector space Z, the map $\mathrm{Hom}_F(U \oplus V,Z) \to \mathrm{Hom}_F(U,Z) \times \mathrm{Hom}_F(V,Z)$ given by restriction is a linear isomorphism.
   (f) Suppose that the triple (W, ι_U, ι_V) of a vector space W and maps from U,V to W respectively satisfies the property of (e). Show that there is a unique isomorphism $\varphi \colon W \to U \oplus V$ such that $\varphi \circ \iota_U$ is the inclusion of U in $U \oplus V$, and similarly for V.

2. (The category of sets)
   (a) Show that $\emptyset$ is the unique set U such that for all sets X, X^U is a singleton.
   (b) Show that $1 = \{\emptyset\}$ is the (up to bijection) set U such that for all sets X, U^X is a singleton.
   (c) Let $f \in Y^X$. Show that f is 1-1 iff for all sets Z, and all $g_1, g_2 \in X^Z$, $f \circ g_1 = f \circ g_2$ iff $g_1 = g_2$.
   (d) Let $f \in Y^X$. Show that f is onto iff for all sets Z, and all $g_1, g_2 \in Z^Y$, $g_1 \circ f = g_2 \circ f$ iff $g_1 = g_2$.
   (e) Given sets X_1, X_2 show that the disjoint union $X_1 \sqcup X_2 = X_1 \times \{1\} \cup X_2 \times \{2\}$ together with the maps $\iota_j(x) = (x, j)$ has the property that for any Z, the map $Z^{X_1 \sqcup X_2} \to Z^{X_1} \times Z^{X_2}$ given by restriction: $f \mapsto (f \circ \iota_1, f \circ \iota_2)$ is a bijection. If X_1, X_2 are disjoint, show that $X_1 \cup X_2$ with ι_j the identity maps has the same property.
   (f) Suppose that the triple (U, ι_1', ι_2') of a set U and maps $\iota_j' \colon X_j \to U$ satisfies the property of (e). Show that there is a unique bijection $\varphi \colon U \to X_1 \sqcup X_2$ such that $\varphi \circ \iota_j' = \iota_j$.

Categories

Roughly speaking, the "category of Xs" consists of all objects of type X, for each two such objects of all relevant maps between them, and of the *composition rule* telling us who to compose maps between Xs. We formalize this as follows:

DEFINITION. A *category* is a triple $\mathcal{C} = (V, E, h, t, \circ, \mathrm{Id})$ where: V is a class called the *objects* of $\mathcal{C}$, E is a class called the *arrows* of $\mathcal{C}$, $h,t\colon E \to V$ are maps assigning to each arrow its "head" and "tail" objects, $\mathrm{Id}\colon V \to E$ is map, and $\circ \subset (E\times E)\times E$ is a partially defined function (see below) called *composition*. We suppose that for each two objects $X,Y \in V$, $\mathrm{Hom}_{\mathcal{C}}(X,Y) \stackrel{\mathrm{def}}{=} \{e \in E \mid t(e) = X, h(e) = Y\}$ is a set, and then have:

- For $f,g \in E$, $f\circ g$ is defined iff $h(g) = t(e)$, in which case $f \circ g \in \mathrm{Hom}_{\mathcal{C}}(t(g), h(f))$.
- For each $X \in V$, $\mathrm{Id}_X \in \mathrm{Hom}_{\mathcal{C}}(X,X)$ and for all f, $f\circ \mathrm{Id}_{t(f)} = \mathrm{Id}_{h(f)} \circ f = f$.
- $\circ$ is associative, in the sense that one of $(f\circ g)\circ h$ and $f\circ(g\circ h)$ is defined then so is the other, and they are equal.

In other words, for each three objects X,Y,Z we have a map $\circ\colon \mathrm{Hom}(X,Y)\times\mathrm{Hom}(Y,Z) \to \mathrm{Hom}(X,Z)$ which is associative, and respects the distinguished "identity" map.

EXAMPLE. Some familiar categories:

- **Set**: the category of sets. Here $\mathrm{Hom}_{\mathbf{Set}}(X,Y) = Y^X$ is the set of set-theoretic maps from X to Y, composition is composition of functions, and Id_X is the identity map $X \to X$.
- **Top**: the category of topological spaces with continuous maps. Here $\mathrm{Hom}_{\mathbf{Top}}(X,Y) = C(X,Y)$ is the set of continuous maps $X \to Y$.
- **Grp**: the category of groups with group homomorphims. $\mathrm{Hom}_{\mathbf{Grp}}(G,H)$ is the set of group homomorphisms.
- **Ab**: the category of abelian groups. Note that for abelian groups A,B we have $\mathrm{Hom}_{\mathbf{Ab}}(A,B) = \mathrm{Hom}_{\mathbf{Grp}}(A,B)$ [the word for this is "full subcategory"]
- $\mathbf{Vect}_F$: the category of vector spaces over the field F. Here $\mathrm{Hom}_{\mathbf{Vect}_F}(U,V) = \mathrm{Hom}_F(U,V)$ is the space of linear maps $U \to V$.

3. (Formalization of familiar words) For each category above (except **Set**) express the statement $\mathrm{Id}_X \in \mathrm{Hom}_{\mathcal{C}}(X,X)$ $\circ\colon$ $\mathrm{Hom}(X,Y)\times\mathrm{Hom}(Y,Z)\to\mathrm{Hom}(X,Z)$ as a familiar lemma. For example, "the identity map $X \to X$ is continuous" and "the composition of continuous functions is continuous".

Properties of a single arrow and a single object

DEFINITION. Fix a category $\mathcal{C}$, objects $X,Y \in \mathcal{C}$, and an arrow $f \in \mathrm{Hom}_{\mathcal{C}}(X,Y)$.

- Call f a *monomorphism* if for every object Z and every two arrows $g_1, g_2 \in \mathrm{Hom}(Z,X)$ we have $f \circ g_1 = f \circ g_2$ iff $g_1 = g_2$.
- Call f an *epimorphism* if for every object Z and every two arrows $g_1, g_2 \in \mathrm{Hom}(Y,Z)$ we have $g_1 \circ f = g_2 \circ f$ iff $g_1 = g_2$.
- Call f an *isomorphism* if there is an arrow $f^{-1} \in \mathrm{Hom}_{\mathcal{C}}(Y,X)$ such that $f^{-1} \circ f = \mathrm{Id}_X$ and $f \circ f^{-1} = \mathrm{Id}_Y$.

4. Show that two sets are isomorphic iff they have the same cardinality,

5. Suppose that f is an isomorphism.
   (a) Show that f^{-1} is an isomorphism.
   (b) Show that f is a monomorphism and an epimorphism.
   (c) Show that there is a unique $g \in \mathrm{Hom}_{\mathcal{C}}(Y,X)$ satisfying the properties of f^{-1}.
   (d) Show that composition with f gives bijections $\mathrm{Hom}_{\mathcal{C}}(X,Z) \to \mathrm{Hom}_{\mathcal{C}}(Y,Z)$ and $\mathrm{Hom}_{\mathcal{C}}(W,X) \to \mathrm{Hom}_{\mathcal{C}}(W,Y)$ which respect composition.

   RMK Part (d) means that isomorphic objects "are the same" as far as the category is concerned.

6. For each category in the example above:
   (a) Show that f is a monomorphism iff it is injective set-theoretically.
   (b) Show that f is an epimorphism iff it is surjective set-theoretically, *except* in **Top**.
   (c) Which continuous functions are epimorphisms in **Top**?

DEFINITION. Call an object $I \in C_c$ *initial* if for every object X, $\mathrm{Hom}_{\mathcal{C}}(I,X)$ is a singleton. Call $F \in \mathcal{C}$ *final* if for every $X \in \mathcal{C}$, $\mathrm{Hom}_{\mathcal{C}}(X,F)$ is a singleton.

7. (Uniqueness)
   (a) Let I_1, I_2 be initial. Show that there is a unique isomorphism $f \in \mathrm{Hom}_{\mathcal{C}}(I_1, I_2)$.
   (b) The same for final objects.**In this section we assume** $\mathrm{char}(F) = 0$**.**

8. (Existence)
   (a) Show that the $\emptyset$ is initial and $\{\emptyset\}$ is final in **Set**. Why is $\{\emptyset\}$ not an initial object?
   (b) Show that $\{\underline{0}\}$ is both initial and final in $\mathbf{Vect}_F$.
   (c) Find the initial and final objects in the categories of groups and abelian groups.

Sums and products

DEFINITION. Let $\{X_i\}_{i\in I} \subset \mathcal{C}$ be objects.

- Their *coproduct* is an object $U \in \mathcal{C}$ together with maps $u_i \in \mathrm{Hom}_{\mathcal{C}}(X_i, U)$ such that for every object Z the map $\mathrm{Hom}(U,Z) \to \times_{i\in I} \mathrm{Hom}_{\mathcal{C}}(X_i, U)$ given by $f \mapsto (f \circ u_i)_{i\in I}$ is a bijection.
- Their *product* is an object $P \in \mathcal{C}$ together with maps $p_i \in \mathrm{Hom}_{\mathcal{C}}(P, X_i)$ such that for every object Z the map $\mathrm{Hom}(Z,U) \to \times_{i\in I} \mathrm{Hom}_{\mathcal{C}}(U, X_i)$ given by $f \mapsto (p_i \circ f)_{i\in I}$ is a bijection.

9. (Uniqueness)
   (a) Show that if U, U' are coproducts then there is a unique isomorphism $\varphi \in \mathrm{Hom}_{\mathcal{C}}(U, U')$ such that $\varphi \circ u_i = u_i'$.
   (b) Show that if P, P' are products then there is a unique isomorphism $\varphi \in \mathrm{Hom}_{\mathcal{C}}(P, P')$ such that $p_i' \circ \varphi = p_i$.

10. (Existence)
   (a) In the category **Set**.
      (i) Show that $\bigcup_{i\in I}(X_i \times \{i\})$ with maps $u_i(x) = (x, i)$ is a coproduct. In particular, if X_i are disjoint show that $\bigcup_{i\in I} X_i$ is a coproduct.
      (ii) Show that $\times_{i\in I} X_i$ with maps $p_j((x_i)_{i\in I}) = x_j$ is a product.
   (b) In the category **Top**.
      (i) Show that $[0,2) = [0,1) \cup [1,2)$ (with the inclusion maps) is a coproduct in **Set** but not in **Top** (subspace topologies from $\mathbb{R}$).
      (ii) Show that $\bigcup_{i\in I}(X_i \times \{i\})$ with the topology $\mathcal{T} = \{\bigcup_{i\in I} A_i \times \{i\} \mid A_i \subset X_i \text{ open}\}$ is a coproduct .
      (iii) Show that the product topology on $\times_{i\in I} X_i$ makes it into a product.
   (c) In the category $\mathbf{Vect}_F$.
      (i) Show that $\oplus_{i\in I} X_i$ is a coproduct .
      (ii) Show that $\prod_{i\in I} X_i$ is a product.
   (d) In the category **Grp**.
      (i) Show that the "coordinatewise" group structure on $\times_{i\in I} G_i$ is a product.
      – The coproduct exists, is called the *free product* of the groups G_i, and is denoted $*_{i\in I} G_i$.

Challenge

A category can be thought of as a "labelled graph" – it has a set of vertices (the objects), a set of directed edges (the arrows), and a composition operator and marked identity morphism, but in fact every vertex has a "label" – the object it represents, and every arrow similarly has a label. Suppose you are only given the combinatorial data, without the "labels" (imagine looking at the category of groups as a graph and then deleting the labels that say which vertex is which group). Can you restore the labels on the objects? Given that, can you restore the labels on the arrows? [up to automorphism of each object]?

This is easy in **Set**, not hard in **Top** and $\mathbf{Vect}_F$, a challenge in **Ab** and really difficult in **Grp**.

Math 412: Problem Set 5 (due 14/2/2014)

Tensor products of maps

1. Let U,V be finite-dimensional spaces, and let $A \in \mathrm{End}(U)$, $B \in \mathrm{End}(V)$.
   (a) Show that $(\underline{u},\underline{v}) \mapsto (A\underline{u}) \otimes (B\underline{v})$ is bilinear, and obtain a linear map $A \otimes B \in \mathrm{End}(U \otimes V)$.
   (b) Suppose A,B are diagonable. Using an appropriate basis for $U \otimes V$, Obtain a formula for $\det(A \otimes B)$ in terms of $\det(A)$ and $\det(B)$.
   (c) Extending (a) by induction, show that $A^{\otimes k}$ induces maps $\mathrm{Sym}^k A \in \mathrm{End}(\mathrm{Sym}^k V)$ and $\bigwedge^k A \in \mathrm{End}(\bigwedge^k V)$.
   (**d) Show that the formula of (b) holds for all A,B.

2. Suppose $\frac{1}{2} \in F$, and let U be finite-dimensional. Construct isomorphisms

$$\{\text{ symmetric bilinear forms on } U\} \leftrightarrow \left(\mathrm{Sym}^2 U\right)' \leftrightarrow \mathrm{Sym}^2\left(U'\right).$$

Structure Theory

3. Let L be a lower-triangular square matrix with non-zero diagonal entries.
   (a) Give a "forward substitution" algorithm for solving $L\underline{x} = \underline{b}$ efficiently.
   (b) Give a formula for L^{-1}, proving in particular that L is invertible and that L^{-1} is again lower-triangular.
   RMK We'll see that if $\mathcal{A} \subset M_n(F)$ is a subspace containing the identity matrix and closed under matrix multiplication, then the inverse of any matrix in $\mathcal{A}$ belongs to $\mathcal{A}$, giving an abstract proof of the same result).

4. Let $U \in M_n(F)$ be *strictly upper-triangular*, that is upper triangular with zeroes along the diagonal. Show that $U^n = 0$ and construct such U with $U^{n-1} \neq 0$.

5. Let V be a finite-dimensional vector space, $T \in \mathrm{End}(V)$.
   (*a) Show that the following statements are equivalent:
   (1) $\forall \underline{v} \in V : \exists k \geq 0 : T^k \underline{v} = \underline{0}$; (2) $\exists k \geq 0 : \forall \underline{v} \in V : T^k \underline{v} = \underline{0}$.
   DEF A linear map satisfying (2) is called *nilpotent*. Example: see problem 4.
   (b) Find nilpotent $A,B \in M_2(F)$ such that $A+B$ isn't nilpotent.
   (c) Suppose that $A,B \in \mathrm{End}(V)$ are nilpotent and that A,B commute. Show that $A+B$ is nilpotent.

Supplementary problems

A. (The tensor algebra) Fix a vector space U.

(a) Extend the bilinear map $\otimes\colon U^{\otimes n} \times U^{\otimes m} \to U^{\otimes n} \otimes U^{\otimes m} \simeq U^{\otimes (n+m)}$ to a bilinear map $\otimes\colon \bigoplus_{n=0}^{\infty} U^{\otimes n} \times \bigoplus_{n=0}^{\infty} U^{\otimes n} \to \bigoplus_{n=0}^{\infty} U^{\otimes n}$.

(b) Show that this map $\otimes$ is associative and distributive over addition. Show that $1_F \in F \simeq U^{\otimes 0}$ is an identity for this multiplication.

DEF This algebra is called the *tensor algebra* $T(U)$.

(c) Show that the tensor algebra is *free:* for any F-algebra A and any F-linear map $f\colon U \to A$ there is a unique F-algebra homomorphism $\bar{f}\colon T(U) \to A$ whose restriction to $U^{\otimes 1}$ is f.

B. (The symmetric algebra). Fix a vector space U.

(a) Endow $\bigoplus_{n=0}^{\infty} \mathrm{Sym}^n U$ with a product structure as in 3(a).

(b) Show that this creates a commutative algebra $\mathrm{Sym}(U)$.

(c) Fixing a basis $\{\underline{u}_i\}_{i \in I} \subset U$, construct an isomorphism $F\left[\{x_i\}_{i \in I}\right] \to \mathrm{Sym}^* U$.

RMK In particular, $\mathrm{Sym}^*(U')$ gives a coordinate-free notion of "polynomial function on U".

(d) Let $I \lhd T(U)$ be the two-sided ideal generated by all elements of the form $\underline{u} \otimes \underline{v} - \underline{v} \otimes \underline{u} \in U^{\otimes 2}$. Show that the map $\mathrm{Sym}(U) \to T(U)/I$ is an isomorphism.

RMK When the field F has finite characteristic, the correct definition of the symmetric algebra (the definition which gives the universal property) is $\mathrm{Sym}(U) \stackrel{\text{def}}{=} T(U)/I$, not the space of symetric tensors.

C. Let V be a (possibly infinite-dimensional) vector space, $A \in \mathrm{End}(V)$.

(a) Show that the following are equivalent for $\underline{v} \in V$: (1) $\dim_F \mathrm{Span}_F \{A^n \underline{v}\}_{n=0}^{\infty} < \infty$;
(2) there is a finite-dimensional subspace $\underline{v} \in W \subset V$ such that $AW \subset W$.

DEF Call such $\underline{v}$ *locally finite*, and let V_{fin} be the set of locally finite vectors.

(b) Show that V_{fin} is a subspace of V.

(c) A A is called *locally nilpotent* for every $\underline{v} \in V$ there is $n \geq 0$ such that $A^n \underline{v} = \underline{0}$ (condition (1) of 5(a)). Find a vector space V and a locally nilpotent map $A \in \mathrm{End}(V)$ which is not nilpotent.

(*d) A is called *locally finite* if $V_{\text{fin}} = V$, that is if every vector is contained in a finite-dimensional A-invariant subspace. Find a space V and locally finite linear maps $A, B \in \mathrm{End}(V)$ such that $A + B$ is not locally finite.

CHAPTER 2

Structure Theory

2.1. Introduction (Lecture 15,7/2/14)

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) ["form"] There is a basis $\{\underline{v}_i\}_{i=1}^n \subset V$ in which the matrix of T is "simple".

EXAMPLE 73. (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 $\det(EX) = \det(X)$ whenever E is elementary and that (Lemma 2) $\det(AX) = \det(X)$ whenever A is row-reduced. One can then prove

THEOREM 74. *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 $A_{\mathrm{rr}} \in cD$ 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. $\square$

2.1.2. Triangular matrices.

DEFINITION 75. $A \in M_n(F)$ is upper (lower) triangular if ...

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

LEMMA 76. *The upper-triangular matrix U is invertible iff its diagonal entries are non-zero.*

ALGORITHM 77 (Back-substitution). *Suppose upper-triangular U is invertible. Them the solution to $U\underline{x} = \underline{b}$ is given by setting $x_i = \frac{b_i - \sum_{j=k+1}^n u_{ij}x_j}{u_{ii}}$ for $i = n, n-1, n-2, \cdots, 1$.*

REMARK 78. 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 79. (1) Give formula for inverse of upper-triangular matrix (2) Develop forward-substitution algorithm for lower-triangular matrices.

COROLLARY 80. *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 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 $\dim V_i = i$.

THEOREM 81. *$T \in \mathrm{End}(V)$ 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. Jordan Canonical Form

2.2.1. The minimal polynomial. 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 82. 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 83. *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. □

- 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 84. *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 finite-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 85. 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 86. *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 87. Call $m(x) = m_T(x)$ the *minimal polynomial* of T.

REMARK 88. We will later prove directly that $\deg m_T(x) \le n$.

EXAMPLE 89. (Minimal polynomials)

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

(2) $T = \begin{pmatrix} & 1 \\ & \end{pmatrix}$, $T^2 = 0$ 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 $T - \mathrm{Id} \ne 0$ 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 $T^{n-1} \ne 0$ (why?) so $m_T(x) = x^n$.

(7) [To be proved in problem set] Let $D = \mathrm{diag}(a_1, \ldots, a_n)$ be diagonal, entries the distinct numbers $\{b_1, \cdots, b_r\}$. 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 90 (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. $\square$

REMARK 91. 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 92. *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 93. *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

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

$\square$

COROLLARY 94. *$\lambda \in F$ is an eigenvalue of T iff λ is a root of $m_T(x)$.*

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

REMARK 95. The characteristic polynomial $P_T(x)$ also has this property – this is how eigenvalues are found in basic linear algebra.

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

PROBLEM 96. Find a matrix $A \in M_2(F)$ which only has a 1-d eigenspace.

DEFINITION 97. 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 98 (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 90, 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. $\square$

THEOREM 99. *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 $1 \le i \le r-1$ is non-zero since $T-\lambda_r$ is invertible there. This shorter dependence contradicts the minimality. $\square$

REMARK 100. The sum may very well be empty – there are non-trivial maps without eigenvalues (for example $\begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \in M_2(\mathbb{R})$).

2.2.3. Algebraically closed field. 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 101. 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 102 (Fundamental theorem of algebra). $\mathbb{C}$ *is algebraically closed.*

REMARK 103. 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 $\bar{T} = \mathrm{Id}_K \otimes_F T$).

Back to the stream of the course.

LEMMA 104. *Suppose F is algebraically closed and that* $\dim_F V \ge 1$. *Then every* $T \in \mathrm{End}_F(V)$ *has an eigenvector.*

PROOF. $m_T(x)$ has roots. $\square$

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 105. *(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 106. *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.

2.2.4. Nilpotent maps. We return to the case of a general field F.

DEFINITION 107. 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 108. *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. □

COROLLARY 109 (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 110. *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 111 (Cayley–Hamilton). *Over any field we have $m_T(x) | p_T(x)$ and, equivalently, $p_T(T) = 0$.*

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

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.

LEMMA 112. *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. Applying N we obtain the dependence

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

If all $N\underline{v}_i = \underline{0}$ then we had a dependence in $B \cap \mathrm{Ker}(N)$. Otherwise, no $N\underline{v}_i = \underline{0}$ (this would shorten the dependence), and by uniqueness it follows that, up to rescaling, N permutes the $\underline{v}_i$. But then the same is true for any power of N, contradicting the nilpotence. $\square$

COROLLARY 113. *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 $\mathrm{Span}\left\{N^i \underline{v}\right\}_{i=0}^{k-1}$, this set is invariant, and its intersection with $\mathrm{Ker} N$ is exactly $\left\{N^{k-1}\underline{v}\right\} \neq \{\underline{0}\}$. $\square$

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 114 (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 115. $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ -1 & -2 & -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}$.

- $A^2 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = 0$, so A is nilpotent. The characterstic polynomial must be x^3.

- The image of A is $\mathrm{Span}\left\{\begin{pmatrix}1\\1\\-1\end{pmatrix}\right\}$. Since $A\begin{pmatrix}3\\-1\\0\end{pmatrix}=\begin{pmatrix}1\\1\\-1\end{pmatrix}$, $\mathrm{Span}\left\{\begin{pmatrix}3\\-1\\0\end{pmatrix},\begin{pmatrix}1\\1\\-1\end{pmatrix}\right\}$ 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 $1\le k\le d$ 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 of W compatible with this decomposition – in other words choose subsets $B_k\subset W_k$ such that $\cup_{k\ge k'}B_k$ is a basis for $W_{k'}$. Let $B=\cup_{k=0}^{d-1}B_k=\{\underline{v}_i\}_{i\in I}$ and for each i define k_i by $\underline{v}_i\in B_{k_i}$. 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 $C=\left\{\underline{u}_{i,j}\right\}_{i,j}$ is a union of Jordan blocks. $\square$

- **The set C is linearly independent**: by construction, $N(C)\subset C\cup\{\underline{0}\}$ and $C\cap W=B$ is independent.
- **The set C is spanning**: We prove by induction on $k\le d$ that $\mathrm{Span}_F(C)\supset\mathrm{Ker}(N^k)$. 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

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

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

DEFINITION 116. A *Jordan basis* is a basis as in the Theorem.

LEMMA 117. *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. $\square$

2.2.5. The Jordan canonical form.

THEOREM 118 (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 $(T-\lambda)\,\underline{v}_{\lambda,i,j} = \begin{cases} \underline{v}_{\lambda,i,j-1} & j \geq 1 \\ \underline{0} & j = 1 \end{cases}$. Furthermore, writing $W_\lambda = \mathrm{Ker}(T-\lambda)$ for the eigenspace, we have for each λ, that $1 \leq i \leq \dim_F W_\lambda$ and that the number of i such that $1 \leq j \leq k$ 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 119. *The* algebraic *multipicity of λ is* $\dim_F V_\lambda$. *The* geometric *multiplicity is the number of blocks.*

EXAMPLE 120 (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 115.

(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 $\mathrm{Ker}\,B' = \left\{(x,y,z,w)^t \mid x = y = -(z+w)\right\}$ 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 $\mathrm{Ker}\,(C-2I) = \mathrm{Span}\,\{\underline{e}_2, \underline{e}_4\}$, which must be the 2d generalized eigenspace V_2 giving two 1×1 blocks. For $\lambda = 3$, $\mathrm{Ker}\,(C-3I) = \{(x,y,z,w)^t \mid z = y = -x, w = 3x\} = \mathrm{Span}\,\{(1,-1,-1,3)^t\}$. This isn't the whole generalized eigenspace, and

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

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 would 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

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

Math 412: Problem Set 6 (due 28/2/2014)

P1. (Minimal polynomials)

(a) Find the minimal polynomial of $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

(b) Show that the minimal polynomial of $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -2 & -2 & 2 & 1 \\ 1 & 1 & -1 & 0 \end{pmatrix}$ is $x^2(x-1)^2$.

(c) Find a 3×3 matrix whose minimal polynomial is x^2.

P2. (Generalized eigenspaces) Let A be as in P1(b).

(a) What are the eigenvalues of A?

(b) Find the generalized eigenspaces.

The minimal polynomial

1. Let $D \in M_n(F) = \operatorname{diag}(a_1, \ldots, a_n)$ be diagonal.
   (a) For any polynomial $p \in F[x]$ show that $p(D) = \operatorname{diag}(p(a_1), \ldots, p(a_n))$.
   (b) Show that the minimal polynomial of D is $m_D(x) = \prod_{j=1}^{r}(x - a_{i_j})$ where $\left\{a_{i_j}\right\}_{j=1}^{r}$ is an enumeration of the distinct values among the a_i.
   (c) Show that (over any field) the matrix of P1(b) is not similar to a diagonal matrix.
   (d) Now suppose that U is an upper-triangular matrix with diagonal D. Show that for any $p \in F[x]$, $p(U)$ has diagonal $p(D)$. In particular, $m_D | m_U$.

2. Let $T \in \operatorname{End}(V)$ be diagonable. Show that every generalized eigenspace is simply an eigenspace.

3. Let $S \in \operatorname{End}(U)$, $T \in \operatorname{End}(V)$. Let $S \oplus T \in \operatorname{End}(U \oplus V)$ be the "block-diagonal map".
   (a) For $f \in F[x]$ show that $f(S \oplus T) = f(S) \oplus f(T)$.
   (b) Show that $m_{T \oplus S} = \operatorname{lcm}(m_S, m_T)$ ("least common multiple": the polynomial of smallest degree which is a multiple of both).
   (c) Conclude that $\operatorname{Spec}_F(S \oplus T) = \operatorname{Spec}_F(S) \cup \operatorname{Spec}_F(T)$.
   RMK See also problem A below.

Supplementary problems

A. Let $R \in \mathrm{End}(U \oplus V)$ be "block-upper-triangular", in that $R(U) \subset U$.
   (a) Define a "quotient linear map" $\bar{R} \in \mathrm{End}(U \oplus V/U)$.
   (b) Let S be the restriction of R to U. Show that both m_S, $m_{\bar{R}}$ divide m_R.
   (c) Let $f = \mathrm{lcm}[m_S, m_{\bar{R}}]$ and set $T = f(R)$. Show that $T(U) = \{\underline{0}\}$ and that $T(V) \subset U$.
   (d) Show that $T^2 = 0$ and conclude that $f \mid m_R \mid f^2$.
   (e) Show that $\mathrm{Spec}_F(R) = \mathrm{Spec}_F(S) \cup \mathrm{Spec}_F(\bar{R})$.

B. Let $T \in \mathrm{End}(V)$. For monic irreducible $p \in F[x]$ define $V_p = \left\{\underline{v} \in V \mid \exists k : p(T)^k \underline{v} = \underline{0}\right\}$.
   (a) Show that V_p is a T-invariant subspace of V and that $m_{T\restriction V_p} = p^k$ for some $k \geq 0$, with $k \geq 1$ iff $V_p \neq \{\underline{0}\}$. Conclude that $p^k | m_T$.
   (b) Show that if $\{p_i\}_{i=1}^r \subset F[x]$ are distinct monic irreducibles then the sum $\bigoplus_{i=1}^r V_{p_i}$ is direct.
   (c) Let $\{p_i\}_{i=1}^r \subset F[x]$ be the prime factors of $m_T(x)$. Show that $V = \bigoplus_{i=1}^r V_{p_i}$.
   (d) Suppose that $m_T(x) = \prod_{i=1}^r p_i^{k_i}(x)$ is the prime factorization of the minimal polynomial. Show that $V_{p_i} = \mathrm{Ker}\, p_i^{k_i}(T)$.

Math 412: Problem set 7 (due 10/3/2014)

Practice

P1. Find the characteristic and minimal polynomial of each matrix:

$$\begin{pmatrix}1&1&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix},\ \begin{pmatrix}5&0&0&0&0&0\\0&4&0&0&0&0\\0&0&2&0&0&0\\0&0&0&2&1&0\\0&0&0&0&2&1\\0&0&0&0&0&2\end{pmatrix},\ \begin{pmatrix}5&0&0&0&0&0\\0&2&1&0&0&0\\0&0&2&0&0&0\\0&0&0&2&1&0\\0&0&0&0&2&1\\0&0&0&0&0&2\end{pmatrix}.$$

P2. Show that $\begin{pmatrix}0&1&\alpha\\0&0&1\\0&0&0\end{pmatrix}$, $\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$ are similar. Generalize to higher dimensions.

The Jordan Canonical Form

1. For each of the following matrices, (i) find the characteristic polynomial and eigenvalues (over the complex numbers), (ii) find the eigenspaces and generalized eigenspaces, (iii) find a Jordan basis and the Jordan form.

$$A=\begin{pmatrix}1&2&1&0\\-2&1&0&1\\0&0&1&2\\0&0&-2&1\end{pmatrix},\ B=\begin{pmatrix}0&1&-1&0\\0&0&0&-1\\1&0&0&1\\0&1&0&0\end{pmatrix},\ C=\begin{pmatrix}1&0&0&0&0&0\\0&0&0&0&-1&1\\-1&-1&1&1&-1&1\\0&0&0&1&0&0\\0&1&0&0&2&0\\0&0&0&0&0&1\end{pmatrix}.$$

RMK I suggest computing by hand first even if you later check your answers with a CAS.

2. Suppose the characteristic polynomial of T is $x(x-1)^3(x-3)^4$.
   (a) What are the possible minimal polynomials?
   (b) What are the possible Jordan forms?

3. Let $T,S \in \mathrm{End}_F(V)$.
   (a) Suppose that T,S are similar. Show that $m_T(x) = m_S(x)$.
   (b) Prove or disprove: if $m_T(x) = m_S(x)$ and $p_T(x) = p_S(x)$ then T,S are similar.

4. Let F be algebraically closed of characteristic zero. Show that every $g \in \mathrm{GL}_n(F)$ has a square root, that is $g = h^2$ for some $h \in \mathrm{GL}_n(F)$.

5. Let V be finite-dimensional, and let $\mathcal{A} \subset \mathrm{End}_F(V)$ be an F-subalgebra, that is a subspace containing the identity map and closed under multiplication (composition). Suppose that $T \in \mathcal{A}$ is invertible in $\mathrm{End}_F(V)$. Show that $T^{-1} \in \mathcal{A}$.

(extra credit problem on reverse)

Extra credit

6. (The additive Jordan decomposition) Let V be a finite-dimensional vector space, and let $T \in \mathrm{End}_F(V)$.

   DEF An *additive Jordan decomposition* of T is an expression $T = S + N$ where $S \in \mathrm{End}_F(V)$ is diagonable, $N \in \mathrm{End}_F(V)$ is nilpotent, and S, N commute.

   (a) Suppose that F is algebraically closed. Separating the Jordan form into its diagonal and off-diagonal parts, show that T has an additive Jordan decomposition.

   (b) Let $S, S' \in \mathrm{End}_F(V)$ be diagonable and suppose that S, S' commute. Show that $S + S'$ is diagonable.

   (c) Show that a nilpotent diagonable linear transformation vanishes.

   (d) Suppose that T has two additive Jordan decompositions $T = S + N = S' + N'$. Show that $S = S'$ and $N = N'$.

Supplementary problems

A. (extension of scalars for linear algebra) Let $F \subset K$ be fields and let V be an F-vectorspace. Let $V_K = K \otimes_F V$, where we consider K as an F-vectorspace in the natural way.

   (a) Show that setting $\alpha(u \otimes \underline{v}) = (\alpha u) \otimes \underline{v}$ extends to a map $K \times V_K \to V_K$ satisfying the axioms of scalar multiplication for a K-vectorspace and compatible with the structure of $V_K$as an F-vectorspace coming from the tensor product.

   (b) Let $\{\underline{v}_i\}_{i \in I} \subset V$ be a set of vectors. Show that it is linearly independent (resp. spanning) iff $\{1_K \otimes \underline{v}_i\}_{i \in I} \subset V_K$ is linearly independent (resp. spanning).

   RMK This is how we show that the minimal polynomial does not depend on the field.

   (c) For $T \in \mathrm{End}_F(V)$ let $T_K = \mathrm{Id}_K \otimes_F T \in \mathrm{End}_F(V_K)$ be the tensor product map. Show that T_K is in fact K-linear.

   (d) Show that $T_K \in \mathrm{End}_K(V_K)$ is the unique K-linear map such that for any basis $\{\underline{v}_i\}_{i \in I} \subset V$, the matrix of T_K in the basis $\{1_K \otimes_F \underline{v}_i\}_{i \in I}$ is the matrix of T in the basis $\{\underline{v}_i\}$ (identification of the matrices under the inclusion $F \subset K$).

B. (conjugacy classes in $\mathrm{GL}_n(F)$) Let F be a field, and let $G = \mathrm{GL}_n(F)$.

   (a) Construct a bijection between conjugacy classes in G and certain Jordan forms. Note that the spectrum can lie in an extension field.

   (b) Enumerate the conjugacy classes in $\mathrm{GL}_2(\mathbb{F}_p)$.

   (c) Enumerate the conjugacy classes of $\mathrm{GL}_3(\mathbb{F}_p)$.

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. Review of metric spaces

DEFINITION 121. A *metric space* is a pair (X, d_X) where X is a set, and $d_X : X \times X \to \mathbb{R}_{\geq 0}$ 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 122. 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 123. Let $(X, d_X), (Y, d_Y)$ be metric spaces and let $f : X \to Y$ 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 $\delta = L\varepsilon$, 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 124. *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 125. We call the metric space (X, d_X) *complete* if every Cauchy sequence converges.

3.2. Norms on vector spaces

Fix a vector space V.

DEFINITION 126. 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 127. *Let* $\|\cdot\|$ *be a norm on* V. *Then the function* $d(\underline{u}, \underline{v}) = \|\underline{u} - \underline{v}\|$ *is a metric.*

EXERCISE 128. 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.

3.2.1. Finite-dimensional examples.

EXAMPLE 129. 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 left to supplement. $\square$

EXERCISE 130. 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 131. Draw the unit balls of

PROPOSITION 132 (Pullback). *Let $T: U \hookrightarrow V$ 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. $\square$

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

EXAMPLE 133. 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 $\|\cdot\|_\infty : X^F \to [0, \infty]$ 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. $\square$

REMARK 134. A vector space with basis B can be embedded into $\ell^\infty(B)$ (we've basically seen this).

EXAMPLE 135. $\ell^p(\mathbb{N}) = \left\{\underline{a} \in F^{\mathbb{N}} : \sum_{i=1}^\infty |a_i|^p < \infty\right\}$ with the obvious norm.

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

DEFINITION 136. $L^p(\mathbb{R}) = \left\{f: \mathbb{R} \to F \text{ [measurable]} \mid \int_{\mathbb{R}} |f(x)|^p \,\mathrm{d}x < \infty\right\} \Big/ \{f \mid f = 0 \text{ a.e.}\}$ with the natural norm.

REMARK 137. The quotient is *essential*: for actualy functions, can have $\int |f(x)|^p \,\mathrm{d}x = 0$ without $f = 0$ exactly. In particular, elements of $L^p(\mathbb{R})$ don't have specific values.

FACT 138. *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 139 (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.*

3.2.3. 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 140. *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 141. 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 142. 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 143. *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

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

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

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

□

3.3. Norms on matrices

DEFINITION 144. Let U, V be normed spaces. A map $T \colon U \to V$ 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 145. 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 146. 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,
$$\left\| A \begin{pmatrix} x \\ y \end{pmatrix} \right\|_\infty = \max\{|x+y|, |y|\} \le 2 \max\{|x|, |y|\}$$
with equality if $x = y$, Thus $\|A\|_\infty = 2$.

EXAMPLE 147. 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 148. *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 140 there is B so that $\|T\underline{u}\|_V \le B \|\underline{u}\|_1 \le (AB) \|\underline{u}\|_U$. □

LEMMA 149. *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 150. *The operator norm is a norm on* $\mathrm{Hom}_b(U,V)$*, the space of bounded maps* $U \to V$.

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

Math 412: Problem set 8, due 19/3/2014

Practice: Norms

P1. Call two norms $\|\cdot\|_1, \|\cdot\|_2$ on V *equivalent* if there are constants c, C such that for all $\underline{v} \in V$,
$$c\|\underline{v}\|_1 \le \|\underline{v}\|_2 \le C\|\underline{v}\|_1 .$$
(a) Show that this is an equivalence relation.
(b) Suppose the two norms are equivalent and that $\lim_{n\to\infty}\|\underline{v}_n\|_1 = 0$ (that is, that $\underline{v}_n \xrightarrow[n\to\infty]{\|\cdot\|_1} \underline{0}$). Show that $\lim_{n\to\infty}\|\underline{v}_n\|_2 = 0$ (that is, that $\underline{v}_n \xrightarrow[n\to\infty]{\|\cdot\|_2} \underline{0}$).
(**c) Show the converse of (b) also holds. In other words, two norms are equivalent iff they determine the same notion of convergence.

P2. Constructions
(a) Let $\{(V_i, \|\cdot\|_i)\}_{i=1}^n$ be normed spaces, and let $1 \le p \le \infty$. For $\underline{v} = (\underline{v}_i) \in \bigoplus_{i=1}^n V_i$ define
$$\|\underline{v}\| = \left(\sum_{i=1}^n \|\underline{v}_i\|_i^p\right)^{1/p} .$$
Show that this defines a norm on $\bigoplus_{i=1}^n V_i$.
DEF This operation is called the *L^p-sum* of the normed spaces.
DEF Let $(V, \|\cdot\|)$ be a normed space, and let $W \subset V$ be a subspace. For $\underline{v} + W \in V/W$ set $\|\underline{v}+W\|_{V/W} = \inf\{\|\underline{v}+\underline{w}\| : \underline{w} \in W\}$. Show
(b) Show that $\|\cdot\|_{V/W}$ is 1-homogenous and satisfies the triangle inequality (it's not always a norm because it can be zero for non-zero vectors).

Norms

1. Let $f(x) = x^2$ on $[-1,1]$.
(a) For $1 \le p < \infty$. Calculate $\|f\|_{L^p} = \left(\int_{-1}^1 |f(x)|^p \,\mathrm{d}x\right)^{1/p}$.
(b) Calculate $\|f\|_{L^\infty} = \sup\{|f(x)| : -1 \le x \le 1\}$. Check that $\lim_{p\to\infty}\|f\|_{L^p} = \|f\|_\infty$.
(c) Calculate $\|f\|_{H^2} = \left(\|f\|_{L^2}^2 + \|f'\|_{L^2}^2 + \|f''\|_{L^2}^2\right)^{1/2}$.
SUPP Show that the H^2 norm is equivalent to the norm $\left(\|f\|_{L^2}^2 + \|f''\|_{L^2}^2\right)^{1/2}$.

2. Let $A \in M_n(\mathbb{R})$.
(a) Show $\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|$ (hint: we basically did this in class).
(b) Show that $\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|$.
RMK See below on *duality*.

3. The *spectral radius* of $A \in M_n(\mathbb{C})$ is the magnitude of its largest eigenvalue: $\rho(A) = \max\{|\lambda| \,|\, \lambda \in \mathrm{Spec}(A)\}$.
(a) Show that for any norm $\|\cdot\|$ on F^n and any $A \in M_n(F)$, $\rho(A) \le \|A\|$.
(b) Suppose that A is diagonable. Show that there is a norm on F^n such that $\|A\| = \rho(A)$.
(*c) Show that if A is Hermitian then $\|A\|_2 = \rho(A)$.

(d) Show that if A,B are similar, and $\|\cdot\|$ is any norm in $\mathbb{C}^n$, then $\lim_{m\to\infty}\|A^m\|^{1/m} = \lim_{m\to\infty}\|B^m\|^{1/m}$ (in the sense that, if one limit exists, then so does the other, and they are equal).

(**e) Show that for any norm on $\mathbb{C}^n$ and any $A \in M_n(\mathbb{C})$, we have $\lim_{m\to\infty}\|A^m\|^{1/m} = \rho(A)$.

4. The *Hilbert–Schmidt* norm on $M_n(\mathbb{C})$ is $\|A\|_{\mathrm{HS}} = \left(\sum_{i,j=1}^n |a_{ij}|^2\right)^{1/2}$.
   - Show that $\|A\|_{\mathrm{HS}} = \left(\mathrm{Tr}(A^\dagger A)\right)^{1/2}$.

   (a) Show that this is, indeed, a norm.

   (b) Show that $\|A\|_2 \leq \|A\|_{\mathrm{HS}}$.

Supplementary problems

A. A *seminorm* on a vector space V is a map $V \to \mathbb{R}_{\geq 0}$ that satisfies all the conditions of a norm except that it can be zero for non-zero vectors.

   (a) Show that for any $f \in V'$, $\varphi(\underline{v}) = |f(\underline{v})|$ is a seminorm.

   (b) Construct a seminorm on $\mathbb{R}^2$ not of this form.

   (c) Let Φ be a family of seminorms on V which is pointwise bounded. Show that $\bar{\varphi}(\underline{v}) = \sup\{\varphi(\underline{v}) \mid \varphi \in \Phi\}$ is again a seminorm.

B. For $\underline{v} \in \mathbb{C}^n$ and $1 \leq p \leq \infty$ let $\|\underline{v}\|_p$ be as defined in class.

   (a) For $1 < p < \infty$ define $1 < q < \infty$ by $\frac{1}{p} + \frac{1}{q} = 1$ (also if $p = 1$ set $q = \infty$ and if $p = \infty$ set $q = 1$). Given $x \in \mathbb{C}$ let $y(x) = \frac{\bar{x}}{|x|}|x|^{p/q}$ (set $y = 0$ if $x = 0$), and given a vector $\underline{x} \in \mathbb{C}^n$ define a vector $\underline{y}$analogously.

   (i) Show that $\|\underline{y}\|_q = \|\underline{x}\|_p^{p/q}$.

   (ii) Show that $|\sum_{i=1}^n x_i y_i| = \|\underline{x}\|_p \|\underline{y}\|_q$

   (b) Now let $\underline{u}, \underline{v} \in \mathbb{C}^n$ and let $1 \leq p \leq \infty$. Show that $|\sum_{i=1}^n u_i v_i| \leq \|\underline{u}\|_p \|\underline{v}\|_q$ (this is called *Hölder's inequality*).

   (c) Conlude that $\|\underline{u}\|_p = \max\left\{|\sum_{i=1}^n u_i v_i| \mid \|\underline{v}\|_q = 1\right\}$.

   (d) Show that $\|\underline{u}\|_p$ is a norm (hint: A(c)).

   (e) Show that $\lim_{p\to\infty}\|\underline{v}\|_p = \|\underline{v}\|_\infty$ (this is why the supremum norm is usually called the L^∞ norm).

C. Let $\{\underline{v}_n\}_{n=1}^\infty$ be a Cauchy sequence in a normed space. Show that $\{\|\underline{v}_n\|\}_{n=1}^\infty \subset \mathbb{R}_{\geq 0}$ is a Cauchy sequence.

D. Let X be a set. For $1 \leq p < \infty$ set $\ell^p(X) = \{f\colon X \to \mathbb{C} \mid \sum_{x\in X}|f(x)|^p < \infty\}$, and also set $\ell^\infty(X) = \{f\colon X \to \mathbb{C} \mid f \text{ bounded}\}$.

   (a) Show that for $f \in \ell^p(X)$ and $g \in \ell^q(X)$ we have $fg \in \ell^1(X)$ and $|\sum_{x\in X} f(x)g(x)| \leq \|f\|_p \|g\|_q$.

   (b) Show that $\ell^p(X)$ are subspaces of $\mathbb{C}^X$, and that $\|f\|_p = (\sum_{x\in X}|f(x)|^p)^{1/p}$ is a norm on $\ell^p(X)$

   (c) Let $\{f_n\}_{n=1}^\infty \subset \ell^p(X)$ be a Cauchy sequence. Show that $\{f_n(x)\}_{n=1}^\infty \subset \mathbb{C}$ is a Cauchy sequence.

(d) Let $\{f_n\}_{n=1}^{\infty} \subset \ell^p(X)$ be a Cauchy sequence and let $f(x) = \lim_{n\to\infty} f_n(x)$. Show that $f \in \ell^p(X)$.

(e) Let $\{f_n\}_{n=1}^{\infty} \subset \ell^p(X)$ be a Cauchy sequence. Show that it is convergent in $\ell^p(X)$.

E. Let V, W be normed vector spaces, equipped with the metric topology coming from the norm. Let $T \in \mathrm{Hom}_F(V,W)$. Show that the following are equivalent:
(1) T is continuous.
(2) T is continuous at zero.
(3) T is *bounded*: $\|T\|_{V\to W} < \infty$, that is: for some $C > 0$ and all $\underline{v} \in V$, $\|T\underline{v}\|_W \le C\|\underline{v}\|_V$.
Hint: the same idea is used in problem P1

F. Let V, W be normed spaces, and let $\mathrm{Hom}_{\mathrm{cts}}(V,W)$ be the set of bounded linear maps from V to W.

(a) Show that the operator norm is a norm on $\mathrm{Hom}_{\mathrm{cts}}(V,W)$.

(b) Suppose that W is complete with respects to its norm. Show that $\mathrm{Hom}_{\mathrm{cts}}(V,W)$ is also complete.

DEF The norm on $V^* \stackrel{\text{def}}{=} \mathrm{Hom}_{\mathrm{cts}}(V,F)$ is called the *dual norm*.

(c) Let $V = \mathbb{R}^n$ and identify V^* with $\mathbb{R}^n$ via the basis of δ-functions. Show that the norm on V^* dual to the ℓ^1-norm is the ℓ^∞ norm and vice versa. Show that the ℓ^2-norm is self-dual.

G. (The completion) Let (X,d) be a metric space.

(a) Let $\{x_n\}, \{y_n\} \subset X$ be two Cauchy sequences. Show that $\{d(x_n,y_n)\}_{n=1}^{\infty} \subset \mathbb{R}$ is a Cauchy sequence.

DEF Let $(\tilde{X}, \tilde{d})$ denote the set of Cauchy sequences in X with the distance $\tilde{d}(\underline{x},\underline{y}) = \lim_{n\to\infty} d(x_n,y_n)$.

(b) Show that $\tilde{d}$ satisfies all the axioms of a metric except that it can be non-zero for distinct sequences.

(c) Show that the relation $\underline{x} \sim \underline{y} \iff \tilde{d}(\underline{x},\underline{y}) = 0$ is an equivalence relation.

(d) Let $\hat{X} = \tilde{X}/\sim$ be the set of equivalence classes. Show that $\tilde{d} \colon \tilde{X} \times \tilde{X} \to \mathbb{R}_{\ge 0}$ descends to a well-defined function $\hat{d} \colon \hat{X} \times \hat{X} \to \mathbb{R}_{\ge 0}$ which is a metric.

(e) Show that $(\hat{X}, \hat{d})$ is a complete metric space.

DEF For $x \in X$ let $\iota(x) \in \hat{X}$ be the equivalence class of the constant sequence x.

(f) Show that $\iota \colon X \to \hat{X}$ is an isometric embedding with dense image.

(g) (Universal property) Show that for any complete metric space (Y, d_Y) and any uniformly continuous $f \colon X \to Y$ there is a unique extension $\hat{f} \colon \hat{X} \to Y$ such that $\hat{f} \circ \iota = f$.

(h) Show that triples $(\hat{X}, \hat{d}, \iota)$ satisfying the property of (g) are unique up to a unique isomorphism.

Hint for D(d): Suppose that $\|f\|_p = \infty$. Then there is a finite set $S \subset X$ with $\left(\sum_{x\in S} |f(x)|^p\right)^{1/p} \ge \lim_{n\to\infty} \|f_n\| + 1$.

3.4. Example: eigenvalues and the power method (Lecture, 17/

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.

3.5. Sequences and series of vectors and matrices

3.5.1. Completeness (Lecture 19/3/2014).

DEFINITION 151. A metric space is complete if

EXAMPLE 152. $\mathbb{R}$. $\mathbb{R}^n$ in any norm. $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ (because isom to $\mathbb{R}^{mn}$).

FACT 153. *Any metric space has a* completion*. [note associated universal property and hence uniqueness]*

THEOREM 154. *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 155. 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. □

3.5.2. Series of vectors and matrices (Lecture 21/3/2014). Fix a complete normed space V.

DEFINITION 156. Say the series $\sum_{n=1}^{\infty} \underline{v}_n$ converges *absolutely* if $\sum_{n=1}^{\infty} \|\underline{v}_n\|_V < \infty$.

PROPOSITION 157. *If $\sum_{n=1}^{\infty} \underline{v}_n$ converges absolutely it converges, and $\|\sum_{n=1}^{\infty} \underline{v}_n\|_V \leq \sum_{n=1}^{\infty} \|\underline{v}_n\|_V$.*

PROOF. Standard. □

THEOREM 158 (M-test). *Let X be a (topological) space, $f_n \colon X \to V$ continuous. Suppose that we have M_n such that $\|f_n(x)\|_V \leq M_n$ 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 $F \colon X \to V$.*

REMARK 159. This can be interpreted as $C_b(X,V)$ (continuous functions $X \to V$ with $\|f(x)\|_V$ bounded) being complete with respect to the norm $\|f\|_\infty = \sup\{\|f(x)\|_V : x \in X\}$.

We will apply this to power series of matrices.

EXAMPLE 160. Let $\|\cdot\|$ be some operator norm on $M_n(\mathbb{R})$, and let $A \in M_n(\mathbb{R})$. For $0 < T < \frac{1}{\|A\|}$ (any $T > 0$ if $A = 0$) and $z \in \mathbb{C}$ with $|z| \leq T$ consider the series

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

We have $\|A^n\| \leq \|A\|^n$ (operator norm!) so that $\|z^n A^n\| \leq (T\|A\|)^n$. Since $\sum_{n=0}^{\infty} (T\|A\|)^n$ 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 161. In fact, the radius of convergence is $\frac{1}{\rho(A)}$.

3.5.3. Vector-valued limits and derivatives (24/3/2014). We recall facts about vector-valued limits.

LEMMA 162 (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 R, replacing $|\cdot|$ with $\|\cdot\|_V$. □

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

DEFINITION 163. 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 164 (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 R, replacing $|\cdot|$ with $\|\cdot\|_V$. $\square$

Math 412: Problem set 9, due 26/3/2014

1. Let $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Let $\underline{v}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
   (a) Find S invertible and D diagonal such that $A = S^{-1}DS$.
   – Prove for yourself the formula $A^k = S^{-1}D^kS$.
   (b) Find a formula for $\underline{v}_k = A^k\underline{v}_0$, and show that $\frac{\underline{v}_k}{\|\underline{v}_k\|}$ converges for any norm on $\mathbb{R}^2$.
   RMK You have found a formula for Fibbonacci numbers (why?), and have shown that the real number $\frac{1}{2}\left(\frac{1+\sqrt{5}}{2}\right)^n$ is exponentially close to being an integer.
   RMK This idea can solve any *difference equation*. We will also apply this to solving *differential* equations.

2. Let $A = \begin{pmatrix} z & 1 \\ 0 & z \end{pmatrix}$ with $z \in \mathbb{C}$.
   (a) Find (and prove) a simple formula for the entries of A^n.
   (b) Use your formula to decide the set of z for which $\sum_{n=0}^{\infty} A^n$ converge, and give a formula for the sum.
   (c) Show that the sum is $(\mathrm{Id} - A)^{-1}$ when the series converges.

3. For each n construct a projection $E_n \colon \mathbb{R}^2 \to \mathbb{R}^2$ of norm at least n ($\mathbb{R}^n$ is equipped with the Euclidean norm unless specified otherwise).
   RMK Prove for yourself that the norm of an *orthogonal* projection is 1.

Supplementary problems

A. Consider the map $\mathrm{Tr}\colon M_n(F) \to F$.
   (a) Show that this is a continuous map.
   (b) Find the norm of this map when $M_n(F)$ is equipped with the $L^1 \to L^1$ operator norm (see PS8 Problem 2(a)).
   (c) Find the norm of this map when $M_n(F)$ is equipped with the Hilbert–Schmidt norm (see PS8 Problem 4).
   (*d) Find the norm of this map when $M_n(F)$ is equipped with the $L^p \to L^p$ operator norm. Find the matrices A with operator norm 1 and trace maximal in absolute value.

B. Call $T \in \mathrm{End}_F(V)$ *bounded below* if there is $K > 0$ such that $\|T\underline{v}\| \geq K\|\underline{v}\|$ for all $\underline{v} \in V$.
   (a) Let T be boudned below. Show that T is invertible, and that T^{-1} is a bounded operator.
   (*b) Suppose that V is finite-dimensional. Show that every invertible map is bounded below.

CHAPTER 4

The Holomorphic Calculus

4.1. The exponential series (24/3/2014)

We prove in the last lecture:

THEOREM 165. *$f_n\colon X \to V$ cts, $\|f_n(x)\|_V \leq M_n$. Then if $\sum_n M_n < \infty$, $\sum_n f_n$ converges uniformly to a cts function $X \to V$.*

We apply this to *power series*:

COROLLARY 166. *Let $\sum_n a_n z^n$ be a power series with radius of convergence R. Then $\sum_n a_n A^n$ 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 167. $\exp(A) = \sum_k \frac{A^k}{k!}$ converges everywhere.

LEMMA 168. $\exp(tA)\exp(sA) = \exp((t+s)A)$.

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

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

□

COROLLARY 169. $\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA) = A\exp(tA) = \exp(tA)A$.

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

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

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 $A\exp(tA) = \exp(tA)A$. □

4.1.1. Application: differential equations with constant coefficients. 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 170. *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 171. *Let* $f\colon [0,1] \to V$ *be strongly 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 $\varphi\left(f(t_0) - f(0)\right) \neq 0$. Then $\varphi \circ f\colon [0,1] \to \mathbb{R}$is differentiable and its derivative is 0:

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

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

REMARK 172. If V is finite-dimensional, every linear functional is bounded. If V is infinite-dimensional the existence of φ is a serious fact.

Now consider a *linear ODE with constant coefficients*:

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

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

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

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.

4.1.2. Diagonal matrices.

HW: $f(\mathrm{diag}(a_1,\ldots,a_n)) = \mathrm{diag}(f(a_1),\ldots,f(a_n))$.

4.2. 26/3/2014

DEFINITION 173. Let $f(z) = \sum_n a_n z^n$. Define $f(A) = \sum_{n=0}^{\infty} a_n A^n$.

LEMMA 174. $Sf(A)S^{-1} = f(SAS^{-1})$.

PROPOSITION 175. $(f \circ g)(A) = f(g(A))$ *if it all works.*

THEOREM 176. $\det(\exp(A)) = \exp(\mathrm{Tr}(A))$.

Math 412: Problem set 10, due 7/4/2014

Differenetial Equations

1. We will analyze the differential equation $u'' = -u$ with initial data $u(0) = u_0$, $u'(0) = u_1$.
   (a) Let $\underline{v}(t) = \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}$. Show that u is a solution to the equation iff $\underline{v}$ solves
   $$\underline{v}'(t) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \underline{v}(t)\,.$$
   (b) Let $W = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Find formulas for W^n and express $\exp(Wt) = \sum_{k=0}^{\infty} \frac{W^k t^k}{k!}$ as a matrix whose entries are standard power series.
   (c) Show that $u(t) = u_0 \cos(t) + u_1 \sin(t)$.
   (d) Find a matrix S such that $W = S \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} S^{-1}$. Evaluate $\exp(Wt)$ again, this time using $\exp(Wt) = S \left(\exp \begin{pmatrix} it & 0 \\ 0 & -it \end{pmatrix} \right) S^{-1}$.

2. Consider the differential equation $\frac{\mathrm{d}}{\mathrm{d}t}\underline{v} = B\underline{v}$ where B is at in PS7 problem 1.
   (a) Find matrices S, D so that D is in Jordan form, and such that $B = SDS^{-1}$.
   (b) Find $\exp(tD)$ directly (as in 1(b)).
   (c) Find the solution such that $\underline{v}(0) = \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}^{\mathrm{t}}$.

Power series

3. Products of absolutely convergent series.
   (a) Let V be a normed space, and let $T, S \in \mathrm{End}_{\mathrm{b}}(V)$ commute. Show that $\exp(T+S) = \exp(T)\exp(S)$.
   (b) Show that, for appropriate values of t, $\exp(A)\exp(B) \neq \exp(A+B)$ where $A = \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ -t & 0 \end{pmatrix}$.

Companion matrices

PRAC Find the Jordan canonical form of $\begin{pmatrix} & 1 & \\ & & 1 \\ 0 & 0 & 2 \end{pmatrix}$.

4. Let $C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 1 \\ a_0 & a_1 & \cdots & a_{n-2} & a_{n-1} \end{pmatrix}$ be the companion matrix associated with the polynomial $p(x) = x^n - \sum_{k=0}^{n-1} a_k x^k$.
   (a) Show that $p(x)$ is, indeed, the characteristic polynomial of C.
   – For parts (b),(c) fix a non-zero root λ of $p(x)$.
   (b) Find (with proof) an eigenvector with eigenvalue λ.
   (**c) Let g be a polynomial, and let $\underline{v}$ be the vector with entries $v_k = \lambda^k g(k)$ for $0 \leq k \leq n-1$. Show that, if the degree of g is small enough (depending on p, λ), then $((C-\lambda)\underline{v})_k = \lambda\,(g(k+1) - g(k))\,\lambda^k$ and (the hard part) that
   $$((C-\lambda)\underline{v})_{n-1} = \lambda\,(g(n) - g(n-1))\,\lambda^{n-1}\,.$$
   (**d) Find the Jordan canonical form of C.

Holomorphic calculus

Let $f(z) = \sum_{m=0}^{\infty} a_m z^m$ be a power series with radius of convergence R. For a matrix A define $f(A) = \sum_{m=0}^{\infty} a_m A^m$ if the series converges absolutely in some matrix norm.

5. Let $D = \operatorname{diag}(\lambda_1, \cdots, \lambda_n)$ be diagonal with $\rho(D) < R$ (that is, $|\lambda_i| < R$ for each i). Show that $f(D) = \operatorname{diag}(f(\lambda_1), \cdots, f(\lambda_n))$.

6. Let $A \in M_n(\mathbb{C})$ be a matrix with $\rho(A) < R$.
   (a) [review of power series] Choose R' such that $\rho(A) < R' < R$. Show that $|a_m| \leq C(R')^{-m}$ for some $C > 0$.
   (b) Using PS8 problem 3(a) show that $f(A)$ converges absolutely with respect to any matrix norm.
   (*c) Suppose that $A = S(D+N)S^{-1}$ where $D+N$ is the Jordan form (D is diagonal, N upper-triangular nilpotent). Show that
   $$f(A) = S\left(\sum_{k=0}^{n} \frac{f^{(k)}(D)}{k!} N^k\right) S^{-1}\,.$$
   Hint: D, N commute.
   RMK1 This gives an alternative proof that $f(A)$ converges absolutely if $\rho(A) < R$, using the fact that $f^{(k)}(D)$ can be analyzed using single-variable methods.
   RMK2 Compare your answer with the Taylor expansion $f(x+y) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x)}{k!} y^k$.
   (d) Apply this formula to find $\exp(tB)$ where B is as in problem 2.

7. Let $A \in M_n(\mathbb{C})$. Prove that $\det(\exp(A)) = \exp(\mathrm{Tr} A)$.

4.3. Invertibility and the resolvent (31/3/2014)

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).

4.3.1. Application: Gauss-Seidel and Jacobi iteration.

4.3.2. 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 $T - z\mathrm{Id}$ 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 $\mathrm{Ker}\,(T - \lambda)$ 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 177. *The resolvent set is open, and the function ("resolvent function") $\rho(T) \to \mathrm{End}_b(V)$ given by $z \mapsto R(z) = (z\mathrm{Id} - T)^{-1}$ 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 178. 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).

CHAPTER 5

Vignettes

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

5.1. The exponential map and structure theory for $\mathrm{GL}_n(\mathbb{R})$ (2/4/2014)

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

REMARK 180. 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 R).

THEOREM 181. *$\mathrm{Lie}(H)$ is a subspace of $M_n(\mathbb{R})$, closed under $[X,Y]$.*

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

THEOREM 182. *Bijection between closed connected subgroups of G and subalgebras of the Lie algebra.*

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

5.2. Representation Theory of Groups

EXAMPLE 183 (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 $L^2\left(\mathbb{R}^3\right)$.

Bibliography