

Math 403 Midterm #1, Spring 2026

Instructor: Ezra Miller

[1pt] Solutions by: Yash Dagade

I verify that I have consulted no electronic or human sources except those allowed and that I have in all other ways complied with Duke's code of academic honesty.

[1pt] I agree

Due Date: 11:59pm Saturday 14 February 2026

READING ASSIGNMENTS

/100

11. for Thu. 12 February

- [Treil, §6.1–§6.2] Spectral theorem, normal operators

12. for Tue. 17 February

- [Treil, §6.3.1–6.3.2] singular values, positive (semi)definite operators
- [Lax, Appendix 10, Theorem 1, p.334–337] Schur factorization
- [Lax, Chap.8: §Spectral Resolution, §Normal Maps]
- [Lax, Chap.9: §Polar Decomposition–§Singular Value Decomposition, p.169–170]

PROBLEMS

1. Let C be a convex cone (Homework #2.11). Prove that 0 is an extreme point of C if [/14](#) the cone C doesn't contain a straight line. Construct a convex cone that contains an entire straight line but has 0 as an extreme point.

Proof. We start by saying what the words mean. A subset $C \subseteq \mathbb{R}^n$ is convex if, whenever you pick two points in C , the whole line segment between them stays in C . A cone (in the sense of Homework #2.11) means that if $x \in C$ then every nonnegative multiple of x is also in C . A convex cone is just a set that satisfies both properties, and a useful way to think about it is that it is closed under taking nonnegative linear combinations.

Next, an extreme point of a convex set C is a point $p \in C$ that never shows up as a genuine middle point of a segment contained in C . Concretely, p is extreme if whenever $p = tx + (1 - t)y$ with $x, y \in C$ and $0 < t < 1$, the only way this can happen is that $x = y = p$. The intuition is that extreme points are the “corners”: if p lies on a line segment inside the set, then it must be at an endpoint, not in the interior of that segment.

Now assume C is a convex cone and does not contain a straight line. We claim that 0 is an extreme point of C . Suppose 0 can be written as a nontrivial convex combination of two points in C , so

$$0 = tx + (1 - t)y \quad \text{with} \quad x, y \in C, \quad 0 < t < 1.$$

Rearranging forces x to be a negative multiple of y :

$$tx = -(1-t)y \implies x = -\alpha y \text{ where } \alpha = \frac{1-t}{t} > 0.$$

Here is where the cone property matters. Since $x \in C$ and scaling by a positive number keeps you in C , we also have $\alpha^{-1}x \in C$. But $\alpha^{-1}x = -y$, so we have both $y \in C$ and $-y \in C$. Once a convex cone contains a nonzero vector and its negative, it contains the entire line through the origin in that direction, because any real multiple of y can be written using nonnegative scaling of either y or $-y$. This contradicts the assumption that C contains no straight line, unless $y = 0$. If $y = 0$, then the displayed equation forces $x = 0$ as well. Therefore the only way to express 0 as a nontrivial convex combination of points of C is impossible, which is exactly the statement that 0 is an extreme point.

For the second part, we construct a convex cone that contains an entire straight line but still has 0 as an extreme point. In \mathbb{R}^2 , let

$$C := \{(u, v) \in \mathbb{R}^2 : v > 0\} \cup \{(0, 0)\}.$$

This is a cone because scaling a point with $v > 0$ by any $\alpha > 0$ keeps v positive, and scaling by $\alpha = 0$ gives $(0, 0)$, which we included. It is convex because convex combinations of points with positive second coordinate still have positive second coordinate, and any segment from $(0, 0)$ to a point with $v > 0$ stays in $v > 0$ except at the endpoint $(0, 0)$. It contains the entire horizontal line $\{(t, 1) : t \in \mathbb{R}\}$, since every point on that line has second coordinate $1 > 0$. Finally, 0 is an extreme point of C because if x and y are nonzero points of C , then both have $v > 0$, so every nontrivial convex combination $tx + (1-t)y$ with $0 < t < 1$ also has $v > 0$ and cannot equal $(0, 0)$. Thus whenever $(0, 0)$ lies on a segment inside C , it must occur as an endpoint.

Remark. A cone that contains a line through the origin can never have 0 extreme, since 0 would be the midpoint of y and $-y$. The example above contains a straight line that does not pass through the origin, so it avoids that obstruction while still satisfying the cone axioms.

□

2. Fix a finite field \mathbb{F}_q with q elements. Show that q is a power of a prime integer p . Hint: [/14](#)
 \mathbb{F}_q is a vector space over \mathbb{F}_p for some prime p ; why is that true, and how is it useful?

Proof. We begin by recalling what a finite field is. A field is a set where you can add, subtract, multiply, and divide by any nonzero element, and the usual arithmetic rules work. A finite field is simply a field with finitely many elements, and its size is called its order. The goal is to show that if $|\mathbb{F}_q| = q$, then q must be a power of a prime.

The first key idea is that every field carries a built-in notion of “counting by ones.” Inside \mathbb{F}_q we can form

$$1, \quad 1 + 1, \quad 1 + 1 + 1, \quad \dots$$

Because the field is finite, this process cannot produce infinitely many distinct elements. So eventually we must repeat, and in particular there is some positive integer m such that adding 1 to itself m times gives 0. The smallest such m is called the characteristic of the field. In a finite field the characteristic cannot be 0, so we do get such an m .

Now we use the most important property of a field: it has no zero divisors. If m were not prime, say $m = ab$ with $1 < a, b < m$, then

$$0 = m \cdot 1 = (ab) \cdot 1 = (a \cdot 1)(b \cdot 1).$$

In a field a product is zero only if one factor is zero, so we would get either $a \cdot 1 = 0$ or $b \cdot 1 = 0$. That contradicts the minimality of m . Therefore the characteristic must be a prime number; call it p . This already explains a basic obstruction: if you try to build a field “mod 6,” you fail because 6 is not prime, and in fact $\mathbb{Z}/6\mathbb{Z}$ has nonzero elements whose product is zero, so division breaks.

Once we know the characteristic is p , we can see a copy of \mathbb{F}_p sitting inside \mathbb{F}_q . Concretely, the elements

$$0, \quad 1, \quad 2 \cdot 1, \quad \dots, \quad (p-1) \cdot 1$$

form a subfield of \mathbb{F}_q and they behave exactly like arithmetic modulo p . This is the smallest subfield of \mathbb{F}_q , and it is naturally identified with \mathbb{F}_p . This is the point of the hint: \mathbb{F}_q contains \mathbb{F}_p , so we can use \mathbb{F}_p as our scalars and view \mathbb{F}_q as a vector space over \mathbb{F}_p .

Now the counting becomes linear algebra. Let $n = \dim_{\mathbb{F}_p}(\mathbb{F}_q)$. Choosing a basis means that every element of \mathbb{F}_q can be written uniquely as a linear combination of n basis vectors with coefficients in \mathbb{F}_p . Each coefficient has exactly p choices, and the choices are independent across the n coordinates. So the total number of elements in the field is p^n . Therefore $q = p^n$, which is exactly what we wanted.

This also gives the clean reason that there cannot be a field of order 6, or of any size that is not a prime power: if a finite field had q elements, then it would automatically be a finite-dimensional vector space over its prime subfield \mathbb{F}_p , so its size would have to be p^n . If q is not of that form, no field with q elements can exist.

Finally, it is worth knowing the full picture. For every prime power p^n there does exist a field with p^n elements, and any two fields with p^n elements are isomorphic. Existence can be shown by taking an irreducible polynomial of degree n over \mathbb{F}_p and forming the quotient $\mathbb{F}_p[x]/(f)$, while uniqueness is a deeper but standard theorem: the field of size p^n is determined up to isomorphism by the fact that it is the splitting field of the

polynomial $x^{p^n} - x$ over \mathbb{F}_p . The main point for this problem, though, is the structural reason behind the prime power: once you locate \mathbb{F}_p inside \mathbb{F}_q , the size of \mathbb{F}_q is forced by the basic counting rule for finite-dimensional vector spaces. \square

3. Fix a 5×5 matrix A whose sole eigenvalue is λ . Either find the Jordan form of A /14 assuming that $A - \lambda I$, $(A - \lambda I)^2$, and $(A - \lambda I)^3$ have kernels of dimension 2, 3, and 5, respectively, or else prove that no such matrix A exists. Justify your response either way.

Proof. To understand what the problem is really asking, it helps to separate the “easy part” of the matrix from the “hard part.” If A has only one eigenvalue λ , then the matrix $A - \lambda I$ has only eigenvalue 0. In that situation, the entire nontrivial structure of A is captured by the nilpotent operator

$$N := A - \lambda I,$$

because A is just λI plus whatever N does.

Jordan form is the statement that, after choosing a well-adapted basis, every linear operator splits into independent “chains” called Jordan blocks. When there is only one eigenvalue λ , those blocks all look like an upper-triangular matrix with λ on the diagonal and 1’s just above the diagonal inside each block, and 0’s elsewhere. Conceptually, each block corresponds to a chain of vectors

$$v_1, v_2, \dots, v_s$$

with the property that applying N moves you one step backward along the chain:

$$Nv_1 = 0, \quad Nv_2 = v_1, \quad Nv_3 = v_2, \quad \dots, \quad Nv_s = v_{s-1}.$$

This is the intuition you should hold on to: within a single Jordan block, N behaves like a “shift” that eventually drops you to 0 once you walk back past the first vector. The length s of the chain is exactly the size of that Jordan block.

Now connect that picture to the kernels in the problem. The space $\ker(N)$ consists of vectors killed immediately by N . In the chain picture, the only vector killed immediately is the first vector v_1 in each chain. So each Jordan block contributes exactly one independent vector to $\ker(N)$. This is why $\dim \ker(N)$ equals the number of Jordan blocks.

Similarly, $\ker(N^2)$ consists of vectors killed after at most two applications of N . In a chain of length s , the first two vectors v_1 and v_2 are killed by N^2 (and no others if $s \geq 3$). So each Jordan block contributes one dimension to $\ker(N)$ no matter what, and it contributes a second dimension to $\ker(N^2)$ precisely when the block has size at

least 2. The same logic continues: a block contributes a new dimension when passing from $\ker(N^{k-1})$ to $\ker(N^k)$ exactly if the block has size at least k .

With that intuition in place, we can read the given numbers as statements about chain lengths. The condition $\dim \ker(N) = 2$ says there are exactly two Jordan blocks (two chains). Since the whole space is 5-dimensional, the sizes of these two blocks must add to 5, so they are some pair (a, b) with $a \geq b$ and $a + b = 5$.

Next, $\dim \ker(N^2) = 3$ says that when we square N , the kernel grows by only one dimension (from 2 to 3). But the growth from $\ker(N)$ to $\ker(N^2)$ counts how many blocks have size at least 2. So exactly one of the two blocks has size at least 2, meaning the other block must have size 1. Therefore the block sizes must be $(4, 1)$.

Finally, the condition $\dim \ker(N^3) = 5$ says that N^3 kills *every* vector, i.e. $\ker(N^3)$ is the entire 5-dimensional space. In the chain picture, that would mean there cannot exist any chain of length 4 or more, because in a chain of length 4 the last vector survives three steps before dying on the fourth. Concretely, a Jordan block of size 4 always produces at least one vector not killed by N^3 . But we already forced a block of size 4 from the first two conditions, so this is a contradiction.

Therefore no 5×5 matrix A with sole eigenvalue λ can satisfy

$$\dim \ker(A - \lambda I) = 2, \quad \dim \ker(A - \lambda I)^2 = 3, \quad \dim \ker(A - \lambda I)^3 = 5.$$

In other words, the kernel-growth data is asking for an impossible pattern of Jordan chains: it simultaneously forces a chain of length 4 (because only one chain grows at the second power) and forbids any chain of length 4 (because everything is killed by the third power). \square

4. Suppose $\varphi : V \rightarrow W$ is an invertible linear transformation of Hermitian inner product spaces. Prove or construct a counterexample: if $\mathbf{e}_1, \dots, \mathbf{e}_n$ is an orthogonal basis of V and $\|\varphi(\mathbf{e}_i)\| = \|\mathbf{e}_i\|$ for all i , then φ is an isometry: $\|\varphi(\mathbf{v}) - \varphi(\mathbf{v}')\| = \|\mathbf{v} - \mathbf{v}'\|$. /14

Proof. First, let me pin down the vocabulary in the exact way this problem wants you to think. A *Hermitian inner product space* is a complex vector space together with a function $\langle \cdot, \cdot \rangle$ that behaves like “dot product, but complex”: it is linear in one slot (your course will specify which), conjugate-symmetric (swapping the two entries conjugates the value), and positive in the sense that $\langle v, v \rangle$ is a nonnegative real number and equals 0 only when $v = 0$. From this inner product we *define* the norm by the geometric rule “length is the square root of self-inner-product,” namely $\|v\| = \sqrt{\langle v, v \rangle}$. A basis e_1, \dots, e_n is *orthogonal* if different basis vectors have inner product 0, meaning they meet at right angles in the geometry coming from $\langle \cdot, \cdot \rangle$; it is *orthonormal* if it is orthogonal and each vector has length 1. Finally, an *isometry* is a map that preserves

distances, meaning $\|\varphi(v) - \varphi(v')\| = \|v - v'\|$ for all v, v' . Because φ is linear, this is the same as saying it preserves the norm of every vector: $\|\varphi(x)\| = \|x\|$ for all x (just take $x = v - v'$).

Now the statement we are testing says: if e_1, \dots, e_n is orthogonal and $\|\varphi(e_i)\| = \|e_i\|$ for each i , then φ must preserve *all* distances. The intuitive issue is that length information on a basis tells you nothing about the *angles* between the images $\varphi(e_i)$. But in an inner product space, the length of a sum (or difference) depends not just on the lengths of the pieces but also on how aligned they are. In other words, preserving lengths of basis vectors does not force φ to preserve the inner product, and without preserving the inner product you should not expect distances to be preserved.

Here is a concrete counterexample inside a Hermitian inner product space. Let $V = W = \mathbb{C}^2$ with the standard Hermitian inner product

$$\langle (z_1, z_2), (w_1, w_2) \rangle = z_1 \overline{w_1} + z_2 \overline{w_2},$$

so the norm is the usual Euclidean length $\|(z_1, z_2)\| = \sqrt{|z_1|^2 + |z_2|^2}$. Take the standard orthonormal basis $e_1 = (1, 0)$ and $e_2 = (0, 1)$, which is certainly orthogonal. Define a linear map φ by declaring

$$\varphi(e_1) = e_1, \quad \varphi(e_2) = \frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2,$$

and extending linearly to all of \mathbb{C}^2 . Both images have the same length as the originals: $\|\varphi(e_1)\| = \|e_1\| = 1$ and $\|\varphi(e_2)\| = \sqrt{(1/2)^2 + (\sqrt{3}/2)^2} = 1 = \|e_2\|$. Also, φ is invertible because $\varphi(e_1)$ and $\varphi(e_2)$ are linearly independent (they are not multiples of each other).

However, φ is not an isometry. The simplest way to see this is to test one vector whose length is controlled by an angle. Consider $v = e_1 + e_2$. In V , $\|v\| = \|(1, 1)\| = \sqrt{2}$. But

$$\varphi(v) = \varphi(e_1) + \varphi(e_2) = \left(1 + \frac{1}{2}\right)e_1 + \frac{\sqrt{3}}{2}e_2 = \frac{3}{2}e_1 + \frac{\sqrt{3}}{2}e_2,$$

so

$$\|\varphi(v)\|^2 = \left|\frac{3}{2}\right|^2 + \left|\frac{\sqrt{3}}{2}\right|^2 = \frac{9}{4} + \frac{3}{4} = 3, \quad \text{hence} \quad \|\varphi(v)\| = \sqrt{3} \neq \sqrt{2} = \|v\|.$$

Since φ fails to preserve the norm of v , it cannot preserve distances either (take $v' = 0$). This disproves the claim.

What is the moral you should teach? The hypothesis $\|\varphi(e_i)\| = \|e_i\|$ is “length data on a basis,” but isometries require “full geometric data.” In inner product geometry, the full data is the inner product itself: an isometry is exactly a map that preserves inner products, because inner products determine norms and distances. So the condition

that *would* make the conclusion true is not merely that the images have the correct lengths, but that the images remain orthogonal with the same lengths (equivalently, an orthonormal basis is sent to an orthonormal basis). That missing orthogonality is precisely what the counterexample breaks: $\varphi(e_1)$ and $\varphi(e_2)$ have the right lengths but are no longer perpendicular, and that is why sums like $e_1 + e_2$ change length. \square

5. Fix vector spaces $W \subseteq V$ and a linear transformation $T : V \rightarrow V$ such that $T(W) \subseteq W$. /14

Using the universal property of quotients, prove that T induces a linear transformation $\bar{T} : V/W \rightarrow V/W$ given by $\bar{T}(v + W) = T(v) + W$. Show further that \bar{T} is an isomorphism if T is an isomorphism and V has finite dimension. Must the same hold if V does not have finite dimension? Prove or give a counterexample.

Proof. The right way to see this problem is to remember what a quotient V/W is trying to do. The quotient space is built to “forget the directions in W ”: two vectors $v, v' \in V$ represent the same element of V/W exactly when they differ by something in W , i.e. $v - v' \in W$. The canonical map is the projection

$$\pi : V \rightarrow V/W, \quad \pi(v) = v + W,$$

and the slogan is: π is the universal linear way to force every vector in W to become 0.

This slogan is formalized by the universal property of quotients: if $f : V \rightarrow U$ is linear and $W \subseteq \ker(f)$ (so f already kills W), then there is a unique linear map $\bar{f} : V/W \rightarrow U$ such that $f = \bar{f} \circ \pi$. Intuitively, $f(v)$ depends only on the coset $v + W$ because changing v by something in W does not change $f(v)$, so f “descends” to the quotient.

Now apply this directly to the present situation. We are given a linear map $T : V \rightarrow V$ and a subspace $W \subseteq V$ such that $T(W) \subseteq W$. The condition $T(W) \subseteq W$ is exactly the compatibility you need for T to make sense on cosets: if you change v by an element of W , then $T(v)$ changes by an element of W as well. Concretely, define

$$g := \pi \circ T : V \rightarrow V/W, \quad g(v) = \pi(T(v)) = T(v) + W.$$

If $w \in W$, then $T(w) \in W$ by hypothesis, hence $\pi(T(w)) = W$, i.e. $g(w) = 0$. So $W \subseteq \ker(g)$. By the universal property there exists a unique linear map

$$\bar{T} : V/W \rightarrow V/W$$

with $\bar{T} \circ \pi = g = \pi \circ T$. Unwinding this identity gives the explicit formula the problem asks for:

$$\bar{T}(v + W) = \bar{T}(\pi(v)) = (\pi \circ T)(v) = T(v) + W.$$

This also explains “well-definedness” in the most honest way: if $v + W = v' + W$, then $v' - v \in W$, so $T(v') - T(v) = T(v' - v) \in W$, hence $T(v') + W = T(v) + W$. That is precisely why we needed $T(W) \subseteq W$.

Next, assume T is an isomorphism and $\dim V < \infty$. The clean strategy here is: show \bar{T} is surjective, then use finite-dimensional linear algebra to upgrade surjective to bijective. To see surjectivity, take any coset $x + W \in V/W$. Since T is onto, choose $v \in V$ with $T(v) = x$. Then

$$\bar{T}(v + W) = T(v) + W = x + W,$$

so every coset has a preimage and \bar{T} is surjective. Because V/W is finite-dimensional (as a quotient of a finite-dimensional space), any surjective linear map $V/W \rightarrow V/W$ is automatically injective (rank–nullity). Therefore \bar{T} is bijective, hence an isomorphism.

Finally, the same conclusion need not hold when V is infinite-dimensional, because surjective linear maps can fail to be injective in infinite dimension. Here is a concrete counterexample that isolates exactly what breaks. Let V be the vector space with basis $\{e_i : i \in \mathbb{Z}\}$ (finite linear combinations), and define the shift

$$T(e_i) = e_{i+1}.$$

This T is an isomorphism with inverse $T^{-1}(e_i) = e_{i-1}$. Let

$$W := \text{span}\{e_i : i \geq 1\}.$$

Then $T(W) \subseteq W$ since $T(e_i) = e_{i+1}$ stays in the span of $\{e_j : j \geq 1\}$ whenever $i \geq 1$, so \bar{T} is well-defined. But \bar{T} is not injective: the coset $e_0 + W$ is nonzero (because $e_0 \notin W$), yet

$$\bar{T}(e_0 + W) = T(e_0) + W = e_1 + W = W,$$

which is 0 in the quotient. So $\ker(\bar{T}) \neq 0$, hence \bar{T} is not an isomorphism. The intuition is that, in infinite dimension, T can “push” new directions into W in a way that creates collapse modulo W even though T itself is perfectly invertible on V .

In summary, $T(W) \subseteq W$ is exactly the condition that lets T descend to a well-defined linear map on the quotient; in finite dimension, invertibility passes down to the quotient because surjectivity forces injectivity; in infinite dimension, this fails and the shift example shows the induced map may have a nontrivial kernel. \square

6. Fix integers $0 < d < e < n$. Express $\{\text{chains } V_d \subseteq V_e\}$ of subspaces of F^n of dimensions d and e as a quotient of $\text{GL}_n(F)$ modulo a subgroup. How many such chains are there if $F = \mathbb{F}_q$ is a field with q elements? /14

Proof. A good first move is to decode the word “chain.” In linear algebra, a *chain of subspaces* just means a nested inclusion, i.e. a picture where one subspace sits inside another. Here the problem fixes integers $0 < d < e < n$ and asks about all pairs

$$V_d \subseteq V_e \subseteq F^n$$

where $\dim(V_d) = d$ and $\dim(V_e) = e$. So you should literally imagine: first choose an e -dimensional “plane” inside F^n , and then inside that plane choose a smaller d -dimensional “plane.” For instance, when $n = 4$, $e = 2$, $d = 1$, a chain is exactly “a line contained in a plane” inside F^4 . If you like coordinates, one concrete example is

$$\langle e_1 \rangle \subseteq \langle e_1, e_2 \rangle \subseteq F^4,$$

and another is

$$\langle e_1 + e_2 \rangle \subseteq \langle e_1 + e_2, e_3 \rangle \subseteq F^4.$$

So the set in the problem is the set of all such nested pairs; in geometry language this is a *partial flag* of type (d, e) .

This is exactly related to Grassmannians (your intuition is right). The Grassmannian $\text{Gr}(k, n)$ is the set of k -dimensional subspaces of F^n . A chain $V_d \subseteq V_e$ can be thought of as: pick $V_e \in \text{Gr}(e, n)$, then pick $V_d \in \text{Gr}(d, V_e)$ (and $\text{Gr}(d, V_e)$ is naturally the same as $\text{Gr}(d, e)$ once you identify $V_e \cong F^e$). So morally, “chains of type (d, e) ” are like a bundle over $\text{Gr}(e, n)$ whose fiber looks like $\text{Gr}(d, e)$.

Now, why does $\text{GL}_n(F)$ appear? Because $\text{GL}_n(F)$ is the group of invertible linear changes of coordinates of F^n , and it acts on subspaces by $g \cdot U := g(U)$. In particular it acts on chains by

$$g \cdot (V_d \subseteq V_e) := (gV_d \subseteq gV_e).$$

The key teaching idea is: *if a group acts transitively on a set, then that set is the same thing as a quotient G/H where H is the stabilizer of one chosen point.* Here “quotient modulo a subgroup” means the set of left cosets $G/H = \{gH : g \in G\}$; it is a clean way of parametrizing “all ways to move a fixed configuration around,” where two moves are considered the same if they differ by something that does not change the configuration.

So we choose a standard chain. Let

$$U_d := \text{span}(e_1, \dots, e_d), \quad U_e := \text{span}(e_1, \dots, e_e),$$

so $U_d \subseteq U_e \subseteq F^n$. Any chain $V_d \subseteq V_e$ can be obtained from this one by some $g \in \text{GL}_n(F)$: choose a basis of V_d , extend it to a basis of V_e , extend it to a basis of F^n , and send the standard basis to that basis. This shows the action is transitive, hence

$$\{\text{chains } V_d \subseteq V_e\} \cong \text{GL}_n(F) / \text{Stab}(U_d \subseteq U_e).$$

The stabilizer is the subgroup

$$P_{d,e} := \{g \in \mathrm{GL}_n(F) : g(U_d) \subseteq U_d \text{ and } g(U_e) \subseteq U_e\}.$$

If you want a matrix picture (which is often the most teachable), $P_{d,e}$ is the set of invertible matrices that, with respect to the decomposition

$$F^n = \underbrace{U_d}_d \oplus \underbrace{(U_e/U_d)}_{e-d} \oplus \underbrace{(F^n/U_e)}_{n-e},$$

have block form

$$\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}$$

with the diagonal blocks invertible of sizes d , $e - d$, and $n - e$. The “zeros below the diagonal blocks” are exactly the linear-algebra way of saying: you are not allowed to send a vector from U_d down into the complementary directions, and you are not allowed to send a vector from U_e down into directions outside U_e . In other words, $P_{d,e}$ is the subgroup that preserves the standard chain, and the quotient $\mathrm{GL}_n(F)/P_{d,e}$ is the space of all chains. This is the conceptual answer to the first part:

$$\boxed{\{\text{chains } V_d \subseteq V_e\} \cong \mathrm{GL}_n(F)/P_{d,e}}.$$

Now assume $F = \mathbb{F}_q$. The clean counting philosophy here is the same “two-stage choice” suggested by the definition of a chain: first choose the larger subspace V_e , then choose the smaller one V_d inside it. Over \mathbb{F}_q , the number of k -dimensional subspaces of \mathbb{F}_q^m is given by the Gaussian binomial coefficient

$$\binom{m}{k}_q := \prod_{i=0}^{k-1} \frac{q^m - q^i}{q^k - q^i}.$$

This formula is worth knowing conceptually: it comes from counting ordered bases. An ordered basis of a k -dimensional subspace of \mathbb{F}_q^m is the same as choosing k linearly independent vectors in \mathbb{F}_q^m , which can be done in

$$(q^m - 1)(q^m - q) \cdots (q^m - q^{k-1})$$

ways; but each fixed k -dimensional subspace has

$$(q^k - 1)(q^k - q) \cdots (q^k - q^{k-1})$$

ordered bases. Dividing gives exactly $\binom{m}{k}_q$.

Using that philosophy, the number of choices for $V_e \subseteq \mathbb{F}_q^n$ is $\binom{n}{e}_q$, and once V_e is chosen, the number of choices for $V_d \subseteq V_e$ is $\binom{e}{d}_q$ (because $V_e \cong \mathbb{F}_q^e$ as a vector space). Therefore the total number of chains is

$$\boxed{\#\{V_d \subseteq V_e \subseteq \mathbb{F}_q^n\} = \binom{n}{e}_q \binom{e}{d}_q}.$$

This answer is also consistent with the quotient viewpoint: since the action is transitive, the number of chains equals the number of cosets, i.e. $|\mathrm{GL}_n(\mathbb{F}_q)|/|P_{d,e}|$, and the product $\binom{n}{e}_q \binom{e}{d}_q$ is exactly what you get when you simplify that ratio. The important takeaway is not the simplification, but the idea: *chains are partial flags; GL_n moves one standard flag to all flags; the subgroup that fixes the standard flag is block upper triangular; and over \mathbb{F}_q you count by choosing the big subspace and then the small one inside it.* \square

7. Is the ∞ -norm $\|\cdot\|_\infty$ on \mathbb{R}^n induced by an inner product? Justify your response. /14

Proof. Recall first what it means for a norm to be “induced by an inner product.” A norm $\|\cdot\|$ on a real vector space is induced by an inner product $\langle \cdot, \cdot \rangle$ if for every vector v we have $\|v\|^2 = \langle v, v \rangle$. The point of this definition is geometric: in an inner product space, squared length behaves like “length squared = sum of squared components + an angle–interaction term.” Concretely, if one writes the squared length of a sum, the inner product forces an expansion of the form

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle + 2\langle v, w \rangle,$$

and similarly

$$\|v - w\|^2 = \langle v - w, v - w \rangle = \langle v, v \rangle + \langle w, w \rangle - 2\langle v, w \rangle.$$

The important intuition here is that the cross term $2\langle v, w \rangle$ measures the “angle information” between v and w ; it is positive if they point roughly the same way, negative if they point opposite ways, and 0 if they are orthogonal. Now, if we add the two displayed identities, that angle information cancels perfectly. What remains is a purely length-based identity that must hold in any inner product norm:

$$\|v + w\|^2 + \|v - w\|^2 = 2\langle v, v \rangle + 2\langle w, w \rangle = 2\|v\|^2 + 2\|w\|^2.$$

So, without needing any extra theorem, we have derived a necessary “cancellation identity”: whenever a norm comes from an inner product, the squared lengths of the two diagonals of the parallelogram spanned by v and w add up to twice the sum of the squared side lengths. (This is exactly the geometric statement that, in Euclidean geometry, the diagonals of a parallelogram encode the side lengths once the angle terms cancel.)

We now test this cancellation identity for the ∞ -norm on \mathbb{R}^n , defined by

$$\|x\|_\infty = \max_i |x_i|.$$

It is helpful to keep the geometry in mind: in \mathbb{R}^2 the unit ball $\{x : \|x\|_\infty \leq 1\}$ is a square with corners $(\pm 1, \pm 1)$. A square has flat sides and sharp corners, and that already suggests trouble: inner-product unit balls are always ellipses (round in every direction), because “length” coming from an inner product is determined by a quadratic form, and quadratic level sets have curved boundaries. The sharp corners are exactly where the ∞ -norm stops behaving like a quadratic length.

To see the failure concretely, it suffices to work in \mathbb{R}^2 (since any inner product on \mathbb{R}^n restricts to an inner product on a coordinate plane). Let

$$v = (1, 0), \quad w = (0, 1).$$

Then $\|v\|_\infty = \|w\|_\infty = 1$. Moreover,

$$v + w = (1, 1) \quad \text{and} \quad v - w = (1, -1),$$

so $\|v + w\|_\infty = \|v - w\|_\infty = 1$ as well, because the largest coordinate in absolute value is still 1 in each case. Therefore

$$\|v + w\|_\infty^2 + \|v - w\|_\infty^2 = 1^2 + 1^2 = 2,$$

while

$$2\|v\|_\infty^2 + 2\|w\|_\infty^2 = 2 \cdot 1^2 + 2 \cdot 1^2 = 4.$$

These quantities are not equal, so the cancellation identity forced by inner products fails for $\|\cdot\|_\infty$.

Geometrically, this is exactly the “square corner” phenomenon: moving from v and w to $v + w$ pushes you toward the corner of the ℓ_∞ ball, where the norm is controlled by a maximum coordinate rather than by a smooth quadratic measurement of direction. Because the maximum operation discards the kind of angle/cross-term information that inner products encode, there is no way for the squared lengths of the diagonals to consistently reflect the squared lengths of the sides as they must in an inner-product geometry.

Hence the ∞ -norm on \mathbb{R}^n cannot be induced by any inner product. □