MUDDY CARD RESPONSES: LECTURE 24, TUE, DEC 13

TRAVIS SCHEDLER

(1) Just need to understand what $\sigma$ composed with $\tau$ actually means? Example of this? Otherwise the det sign proof is clear. // The proof that $\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \text{sign}(\tau)$ // the pairs $j < k$ such that $\tau(j) > \tau(k)$ are the same as those of $\sigma \circ \tau$ except $(i, i + 1)$ is such a pair for $\tau$ if and only if it is not a pair for $\sigma \circ \tau$

A: First of all, by definition, $\sigma \circ \tau(i) = \sigma(\tau(i))$ for all $i \in \{1, 2, \ldots, n\}$. Note that, if $P_\sigma \in \text{Mat}(n, n, \mathbb{F})$ is the permutation matrix corresponding to $\sigma \in S_n$, then we get

$$P_{\sigma \circ \tau} = P_\sigma P_\tau,$$

so multiplying permutation matrices corresponds to composing permutations.

Here is an example: If $n = 3$ and $\sigma = (2, 3, 1)$, i.e., $\sigma(1) = 2$, $\sigma(2) = 3$, and $\sigma(3) = 1$, and $\tau = (2, 1, 3)$, i.e., $\tau(1) = 2, \tau(2) = 1$, and $\tau(3) = 3$, then $\sigma \circ \tau = (1, 3, 2)$, since $\sigma(\tau(1)) = 1, \sigma(\tau(2)) = 3$, and $\sigma(\tau(3)) = 2$.

To prove that $\text{sign}(\sigma \circ \tau) = \text{sign}(\sigma) \text{sign}(\tau)$, I said in class that it is enough to take $\sigma$ to be a swap of two adjacent numbers, say $i$ and $i + 1$. That is because an arbitrary $\sigma$ is a composition of these (you can get any permutation by a series of swaps of adjacent numbers), so if the sign formula holds when $\sigma$ is this swap, it also holds for a product of them by applying the formula iteratively. E.g., if $\sigma_1$ and $\sigma_2$ are two swaps of adjacent numbers, we would get

$$\text{sign}(\sigma_1 \circ \sigma_2 \circ \tau) = \text{sign}(\sigma_1) \text{sign}(\sigma_2 \circ \tau) = \text{sign}(\sigma_1) \text{sign}(\sigma_2) \text{sign}(\tau) = \text{sign}(\sigma_1 \circ \sigma_2) \text{sign}(\tau),$$

as hence would deduce the formula for $\sigma = \sigma_1 \circ \sigma_2$ as well.

Finally, the proof came down to the final question. Here I admit I made a mistake. First, I should have said $\tau \circ \sigma$ rather then $\tau$. Here is the correct statement:

**Lemma 0.1.** (i) $(i, i + 1)$ is a pair such that $\tau(i) > \tau(i + 1)$ if and only if it is not a pair such that $\tau \circ \sigma(i) > \tau \circ \sigma(i + 1)$.

(ii) If either $j$ or $k$ is not in the set $\{i, i + 1\}$, then $(j, k)$ is a pair such that $j < k$ and $\tau(j) > \tau(k)$ if and only if $(\sigma(j), \sigma(k))$ is a pair such that $\sigma(j) < \sigma(k)$ and $\tau \circ \sigma(j) > \tau \circ \sigma(k)$.

We still get from the lemma a bijection between the pairs that contribute to $o(\tau)$ and those that contribute to $o(\tau \circ \sigma)$, except for a single pair $(i, i + 1)$ which appears in exactly one of these, so we still get $o(\tau \circ \sigma) = o(\tau) \pm 1$. Sorry for the error: I corrected the slides!

Proof of the lemma: First, for (i), First, note that since $\sigma(i) = i + 1$ and $\sigma(i + 1) = i$, we have $$(\tau \circ \sigma(i), \tau \circ \sigma(i + 1)) = (\tau(i + 1), \tau(i)).$$ So evidently $(i, i + 1)$ is a pair of the desired form for $\tau$ if and only if it is not such a pair for $\tau \circ \sigma$.

For part (ii), if $j$ is not in the set $\{i, i + 1\}$, then $j < k$ if and only if $j = \sigma(j) < \sigma(k)$, and similarly if $k$ is not in the set $\{i, i + 1\}$, we get $j < k$ if and only if $\sigma(j) < \sigma(k) = k$. Then, note that $$(\tau(j), \tau(k)) = (\tau \circ \sigma(\sigma(j)), \tau \circ \sigma(\sigma(k)))$$ since $\sigma(\sigma(\ell)) = \ell$ for all $\ell$ (i.e., $\sigma^2$ is the identity). This proves the lemma.

Date: Last revision: Dec 14, 9:30 AM.
(2) Why \( \varphi(x) = x^m \) if continuous.

A: This was the statement that, if \( \varphi : \mathbb{R} \to \mathbb{R} \) satisfies \( \varphi(ab) = \varphi(a)\varphi(b) \) and \( \varphi \) is continuous and nonzero, then \( \varphi(x) = x^m \) for some \( m \geq 0 \) (with \( x^0 \) meaning 1 for all \( x \), even \( x = 0 \)).

Actually I have to make another correction to this (sorry! I am updating the slides): it should be either the constant function \( \varphi = 1 \), or else one of the functions:

\[
\varphi(x) = \text{sign}(x)|x|^c, \quad \text{or} \quad \varphi(x) = |x|^c,
\]

for \( c > 0 \). These include the above, since for \( c = m \geq 1 \) an odd integer, \( \text{sign}(x)|x|^m = x^m \), and for \( c = m \geq 2 \) an even integer, \( |x|^m = x^m \).

To prove this, first note that, since \( \varphi \) is nonzero, as we saw before, \( \varphi(1) = 1 \), since \( \varphi(1) = \varphi(1)^2 \), and \( \varphi(x) = \varphi(x)\varphi(1) \) for all \( x \). Next, \( \varphi(-1)^2 = 1 \) so \( \varphi(-1) = \pm 1 \) as well, and then \( \varphi(-x) = \pm \varphi(x) \) for all \( x \). So up to a choice of sign, it is enough to figure out what possible functions \( \varphi|_{\mathbb{R}_+} : \mathbb{R}_+ \to \mathbb{R}_+ \) we have, where \( \mathbb{R}_+ \) is the set of positive real numbers (note that \( \varphi \) takes positive numbers to positive numbers since it is continuous and it does so at 1, and \( \varphi \) cannot take any positive number \( a > 0 \) to zero since then \( 1 = \varphi(1) = \varphi(a)\varphi(1/a) = 0 \)).

Now, let us actually suppose that \( \varphi \) is differentiable at 1 (proving this follows from continuity is a bit more difficult: see the end of the answer). Now, letting \( x \) be a variable, we get \( ax \varphi'(ax) = \varphi(a)\varphi'(x) \), so plugging in \( x = 1 \), we get \( a\varphi'(a) = \varphi(a)\varphi'(1) \). Plugging in \( x = a \) now, we get \( \varphi'(x) = \varphi'(1) \frac{x^2}{2} \), which is a differential equation we can solve. Write \( c = \varphi'(1) \). Then \( \frac{dx}{\varphi} = \frac{dx}{x} \), so integrating near \( x = 1 \), \( \log \varphi = c \log x \). Exponentiating, \( \varphi(x) = x^c \), near \( x = 1 \).

If \( c < 0 \), then this cannot extend continuously to zero, since \( \varphi(x) \) will go to infinity there. So we can assume \( c \geq 0 \). Then we get a continuous extension to \( \mathbb{R}_+ \), \( \varphi(x) = x^c \) for all \( x > 0 \), and this is actually unique by existence and uniqueness of differential equations. We can always extend this continuously to all of \( \mathbb{R} \) by placing \( \varphi(x) = |x|^c \). This will have \( \varphi(-1) = 1 \). To have \( \varphi(-1) = -1 \), we would need the function \( \varphi(x) = \text{sign}(x)|x|^c \). Since everything is determined by the value of \( \varphi \) on positive numbers and on \(-1 \), these are exactly the two possibilities.

Now, let me remove the assumption that \( \varphi'(1) \) exists. Indeed, in general, pick any \( a > 1 \) and consider the value of \( \varphi(a) \in \mathbb{R}_+ \). Then this determines \( \varphi(a^m) \) for all integers \( m \), and in fact determines \( \varphi(a^q) \) for all rationals \( q \), since positive numbers have unique positive \( k \)-th roots for all \( k \geq 1 \). Since the numbers \( a^q \) are dense in \( \mathbb{R}_+ \) (for any positive real number \( x \), there is a sequence \( q_i \) of rationals such that \( a^{q_i} \) converges to \( x \), i.e., let \( q_i \) converge to \( \log_a(x) = \log(x)/\log(a) \)), this means that \( \varphi \) is uniquely determined by \( \varphi(a) \), so we get that \( \varphi(x) = x^{\log_a(\varphi(a))} = x^{\log(\varphi(a))/\log(a)} \) for \( x > 0 \), which fills in the necessary step (recall \( \varphi(-1) = \pm 1 \)).

(3) \( \int f(T(v)) \, d\text{vol} = |\det T|^{-1} \int f(T(v)) \, d(\text{vol} \circ T) = |\det T|^{-1} \int f(v) \, d\text{vol} \)

A: The point is that, if \( T \) is linear, then \( \text{vol}(T(R)) = |\det T| \text{vol}(R) \). That is, \( \text{vol} \circ T = |\det T| \text{vol} \). So that explains the first inequality: we can replace \( d\text{vol} \) by \( |\det T|^{-1} d(\text{vol} \circ T) \), since \( \det T \) is just a number which we can pull out of the integral. The second equality is just changing the variable from \( T(v) \) to \( v \), which is fine since \( T \) is invertible and we are replacing also the \( d(\text{vol} \circ T) \) by \( d\text{vol} \).

(4) Why we lost Jeopardy

A: I am going to answer why the Final Jeopardy! answer was five: If the characteristic polynomial is \( x^3 \), that means that the matrix is \( 5 \times 5 \), and that it is nilpotent, i.e., if it is in Jordan form (or generally in upper-triangular form), it has zeros on the diagonal. Now if the minimal polynomial is \( x^3 \), that means that the largest size of Jordan block that occurs
is size 3. This means that we can either have two Jordan blocks: one of size 3 and one of size 2, or three Jordan blocks: one of size 3 and two of size 1. There are two Jordan form matrices of the first type (one can order the two blocks in either order) and there are three Jordan form matrices of the second type (one can order the three blocks in any of three possible orderings, corresponding to whether the large block occurs first, second, or third).