1. Suppose that $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a function, and we know that this function obeys these two rules:

A: $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ for any two vectors $\vec{v}$, $\vec{w}$ in $\mathbb{R}^n$.

B: $T(c\vec{v}) = c\,T(\vec{v})$ for any vector $\vec{v}$ in $\mathbb{R}^n$, and any number c.

We'd like to know if this means that T also has to obey this rule:

C: $T(c_1\vec{v_1} + c_2\vec{v_2} + \cdots + c_r\vec{v_r}) = c_1\,T(\vec{v_1}) + c_2\,T(\vec{v_2}) + \cdots + c_r\,T(\vec{v_r})$, for any vectors $\vec{v}_1, \ldots, \vec{v}_r$ in $\mathbb{R}^n$, and any numbers $c_1, \ldots, c_r$.

First, let's establish that rule A tells us that we can move more than just the sum of two vectors through the brackets. More formally, let's establish that if T obeys rule A, then it also obeys

A$_r$: $T(\vec{v_1} + \cdots + \vec{v}_r) = T(\vec{v_1}) + \cdots + T(\vec{v}_r)$ for any r vectors $\vec{v}_1, \ldots, \vec{v}_r$ in $\mathbb{R}^n$.

I guess it's pretty clear that A should imply A$_r$ for all $r \geq 2$; presumably we just need to repeatedly use A somehow. One nice way to organize that argument is to set it up as an induction. (Reminder: induction is just an organizational tool we can try and use in an argument, and not some magic phrase that automatically confers validity on the conclusion. The correctness of an argument is more important than any particular form used to express or organize it).

Proof: (that if A is true, then A$_r$ is true for all $r \geq 2$) by induction on r:

Base case: $r = 2$ (you don't always have to start with 1...): That's exactly statement A, which we're assuming to be true, so the base case is true.

Inductive step: Assume that A$_r$ is true for r, then given $r+1$ vectors $\vec{v}_1$, $\ldots$, $\vec{v}_r$, $\vec{v}_{r+1}$, group their sum as $(\vec{v}_1 + \cdots + \vec{v}_r) + \vec{v}_{r+1}$, then

$$\begin{aligned} T(\vec{v_1} + \cdots + \vec{v}_{r+1}) &= T((\vec{v}_1 + \cdots + \vec{v}_r) + \vec{v}_{r+1}) \\ &= T(\vec{v}_1 + \cdots + \vec{v}_r) + T(\vec{v}_{r+1}) \text{ [by rule A]} \\ &= T(\vec{v}_1) + \cdots + T(\vec{v}_r) + T(\vec{v}_{r+1}) \text{ [using A}_r \text{ on the first part]} \end{aligned}$$

which is exactly the induction step.

Okay, now we know that if a function T obeys rule A, it also obeys A_r for all $r \geq 2$. Let's use this and B to show that T also obeys rule C for any choices of $c_1, \ldots, c_r$, and $\vec{v}_1, \ldots, \vec{v}_r$:

For any such choices

$$\begin{aligned} T(c_1\vec{v}_1 + \cdots + c_r\vec{v}_r) &= T(c_1\vec{v}_1) + \cdots + T(c_r\vec{v}_r) \ \text{[using rule } \mathrm{A}_r\text{]} \\ &= c_1\,T(\vec{v}) + \cdots + c_r\,T_r(\vec{v}_r) \ \text{[using rule B on each of the } T(c_i\vec{v}_i)\text{]} \end{aligned}$$

which is exactly rule C.

If T obeys rule C, does it have to obey A and B? Yes – they're both special cases of the more general rule C. If we apply rule C with $r = 2$ and $c_1 = c_2 = 1$, it says

$$T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2),$$

which is exactly rule A. If we apply rule C with $r = 1$ and c_1 any number, then we get

$$T(c_1\vec{v}_1) = c_1\,T(\vec{v}_1)$$

which is exactly rule B. Therefore a function $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ satisfies rules A and B if and only if it satisfies rule C.

2. The statement that

$$\mathrm{RREF}(AB) = \mathrm{RREF}(A) \cdot \mathrm{RREF}(B)$$

for matrices A and B which can be multiplied is *not* true in general. Here's a simple counterexample:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Then $\mathrm{RREF}(A) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathrm{RREF}(B) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and

$$\mathrm{RREF}(A) \cdot \mathrm{RREF}(B) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

But $AB = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$, so $\mathrm{RREF}(AB) = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ which is not what we got above!

So, the statement is simply not true.

There is a similar sounding statment which *is* true though: if you multiply two matrices in RREF, then the result will still be in RREF. One of the arguments for this is computational – you think about what a matrix in RREF looks like, and then try and deal with what the product looks like. Another is to give a slightly different characterization of what it means for a matrix to be in RREF, one better suited to thinking about composition of functions.

That statment doesn't contradict the fact that the previous one is false. What it tells us is that $\mathrm{RREF}(A) \cdot \mathrm{RREF}(B)$ will always be a matrix in RREF, there's just simply no guarantee that this will be the same matrix in RREF as $\mathrm{RREF}(AB)$.

3.

(a) Here's a proof of the formula which doesn't use induction at all: If

$$S_m = a + ar + ar^2 + \cdots ar^{m-1} + ar^m,$$

then

$$r \cdot S_m = ar + ar^2 + ar^3 + \cdots ar^m + ar^{m+1},$$

so subtracting gives

$$(r-1)S_m = -a + (ar - ar) + (ar^2 - ar^2) + \cdots + (ar^m - ar^m) + ar^{m+1} = a(r^{m+1} - 1).$$

Since $r \neq 1$, we can divide both sides by $(r-1)$ to get the formula.

Here's a proof which does use induction:

Base case: $m = 1$: The sum is $a + am$, and the predicted formula is $\frac{a(r^2-1)}{r-1}$. Since $r^2 - 1 = (r-1)(r+1)$, this is the same as $a(r+1) = ar + a$ which verifies the base case. We could also start the base case with $m = 0$, in which case the sum would be a, and the predicted answer $\frac{a(r-1)}{r-1} = a$, which also works.

Inductive step: Suppose that the formula is true for m, then

$$\begin{aligned} a + ar + \cdots + ar^m + ar^{m+1} &= \frac{a(r^{m+1} - 1)}{r - 1} + ar^{m+1} \\ &= \frac{a(r^{m+1} - 1) + (ar^{m+2} - ar^{m+1})}{r - 1} \\ &= \frac{a(r^{m+2} - 1)}{r - 1} \end{aligned}$$

which is the formula for $m + 1$.

(b) Suppose that a_n, b_n, c_n, and d_n are the entries of A^n, i.e., that

$$A^n = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}.$$

(NOTE: If you haven't seen this kind of language before, this isn't really an assumption. It's a polite way of saying that we're establishing some notation, that we're defining the numbers a_n, b_n, c_n, and d_n to be those numbers that show up when we compute A^n.)

Here are some small powers of A:

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 4 & 5 \\ 0 & 9 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 8 & 19 \\ 0 & 27 \end{bmatrix}, \quad \text{and,} \quad A^4 = \begin{bmatrix} 16 & 65 \\ 0 & 81 \end{bmatrix}.$$

According to the notation we've established, that means that $a_1 = 2$, $a_2 = 4$, $a_3 = 8$, and $a_4 = 16$. (and also that $b_1 = 1$, $b_2 = 5$, $b_3 = 19$, and $b_4 = 65$, that $d_1 = 3$, $d_2 = 9$, $d_3 = 27$, and $d_4 = 81$, and that $c_1 = c_2 = c_3 = c_4 = 0$.)

From the examples, it's probably pretty clear that the correct formulas should be $a_n = 2^n$, $d_n = 3^n$, and $c_n = 0$ for all $n \geq 0$. It's only a slightly larger stretch to notice that $b_n + a_n$ is always d_n, which if true means that we must have $b_n = 3^n - 2^n$.

In order to prove these formulas, we can use the recursive relation:

$$\begin{bmatrix} a_{n+1} & b_{n+1} \\ c_{n+1} & d_{n+1} \end{bmatrix} = A^{n+1} = A^n A = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2a_n & 3b_n + a_n \\ 2c_n & 3d_n + c_n \end{bmatrix}.$$

In other words, that

$$\begin{aligned} a_{n+1} &= 2a_n \\ b_{n+1} &= 3b_n + a_n \\ c_{n+1} &= 2c_n, \quad \text{and} \\ d_{n+1} &= 3d_n + c_n \end{aligned}$$

For all $n \geq 1$.

To organize the proof that our formulas hold for all n, we can again use induction; let's use it to prove all the formulas simultaneously:

Base case: $n = 1$. Here we just have to check that $a_1 = 2^1 = 2$, $b_1 = 3^1 - 2^1 = 1$, $c_1 = 0$, and $d_1 = 3^1 = 3$, which we did when we wrote out the entries of A. We could also check for $n = 0$, in which case A^0 is just the identity matrix.

Inductive step: Supposing the formulas true for n, and using the recursive relations, we have

$$\begin{aligned}
a_{n+1} &= 2a_n = 2 \cdot 2^n = 2^{n+1}, \\
b_{n+1} &= 3b_n + a_n = 3(3^n - 2^n) + 2^n = 3^{n+1} - (3-1)2^n = 3^{n+1} - 2^{n+1}, \\
c_{n+1} &= 2c_n = 2 \cdot 0 = 0, \quad \text{and} \\
d_{n+1} &= 3d_n + c_n = 3 \cdot 3^n + 0 = 3^{n+1}
\end{aligned}$$

proving the inductive step.

So, the formulas are true for all $n \geq 0$.

What about for negative n? If we compute a few values:

$$A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{6} \\ 0 & \frac{1}{3} \end{bmatrix}, \quad A^{-2} = \begin{bmatrix} \frac{1}{4} & -\frac{5}{36} \\ 0 & \frac{1}{9} \end{bmatrix}, \quad \text{and} \quad A^{-3} = \begin{bmatrix} \frac{1}{8} & -\frac{19}{216} \\ 0 & \frac{1}{27} \end{bmatrix}$$

we can see that the entries still match the formulas. Certainly for these values $a_n = 2^n$, $c_n = 0$, $d_n = 3^n$, and (after a little computation) $b_n = 3^n - 2^n$ for $n = -1$, -2, and -3.

So, the formulas seem to be true for negative n as well. How could we prove this? I guess one way might be to use induction again, this time starting at $n = -1$ and going down. There is a faster way though: We want to see that the formulas give the matrix A^n for negative n. We know that the formulas are correct for A^{-n} (since now $-n$ is positive), and if we use these formulas and multiply the matrices together we get:

$$\begin{bmatrix} 2^n & 3^n - 2^n \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} 2^{-n} & 3^{-n} - 2^{-n} \\ 0 & 3^{-n} \end{bmatrix} = \begin{bmatrix} 1 & 2^n(3^{-n} - 2^{-n}) + 3^{-n}(3^n - 2^n) \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2^n 3^{-n} - 1 + 1 - 2^n 3^{-n} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

In other words, the formulas for negative n give a matrix which is the inverse of the matrix for A^{-n}. Since the inverse of a matrix (if it exists) is unique, this means that the formulas for negative n must give A^n, and the formulas are true for all n.

Next semester we'll see that something like this almost always happens: for almost every square matrix A, there are formulas for the entries of the powers of A, and each entry will be a linear combination of the n-th powers of some numbers. These numbers are special numbers attached to the matrix A, ones which govern its behaviour and the behaviour of its powers.

4. Suppose that A is an $n \times n$ matrix, and that A^m is the zero matrix, for some number $m \geq 1$. We want to show that A^n is the zero matrix.

Here are the details of the two strategies suggested:

(I) To start with, the sequence of vector spaces $\ker(A^r)$ for $r = 1, 2, \ldots$ are each contained in the next:

$$\ker(A) \subseteq \ker(A^2) \subseteq \ker(A^3) \subseteq \cdots$$

We've seen this kind of argument before (for example, on homework 6, 2(a), or homework 5 2(c)): If $\vec{v}$ is in $\ker(A^r)$, then $A^{r+1}\vec{v} = AA^r\vec{v} = A\vec{0} = \vec{0}$. Therefore we must have

$$\dim(\ker(A)) \leq \dim(\ker(A^2)) \leq \dim(\ker(A^3)) \leq \cdots$$

Suppose that at one point we have $\dim(\ker(A^r)) = \dim(\ker(A^{r+1})$. Then we actually have $\ker(A^r) = \ker(A^{r+1})$ since one is contained in the other, and they both have the same dimension.

But now, following the same argument as homework 5, 2(c), we can show that $\ker(A^{r+1}) = \ker(A^{r+2})$: Take any vector $\vec{v}$ in $\ker(A^{r+2})$, and let $\vec{w} = A\vec{v}$. Then $A^{r+1}\vec{w} = \vec{0}$, so $\vec{w}$ is in $\ker(A^{r+1})$ and hence (since $\ker(A^{r+1}) = \ker(A^r)$) $\vec{w}$ is in $\ker(A^r)$ too. But then $A^r\vec{w} = \vec{0}$, so $A^rA\vec{v} = \vec{0}$, i.e., $\vec{v}$ is in $\ker(A^{r+1}$. This argument shows that $\ker(A^{r+1}) \subset \ker(A^{r+1})$ and hence they are equal, since we already know that $\ker(A^{r+1}) \subseteq \ker(A^{r+2})$.

Continuing by induction, we can then show that $\ker(A^{r+2}) = \ker(A^{r+3})$, $\ker(A^{r+3}) = \ker(A^{r+4})$, etc. In other words, once we have $\dim(\ker(A^r) = \dim(\ker(A^{r+1})$ for some r, they're equal for all higher powers of A.

So, at each step, $\dim(\ker(A^r))$ either increases, or stays the same, and once it stays the same once it remains the same for ever. Since eventually $\ker(A^r) = n$, it can't stop until the dimension of the kernel is n, ie., all of $\mathbb{R}^n$.

Since $\ker(A)$ is at least one dimensional (if it were zero dimensional, A would be invertible, and so it would be impossible for some power of A to be the zero matrix) $\ker(A^2)$ is at least two dimensional, $\ker(A^3)$ at least three dimensional, etc, up to $\ker(A^n)$ is at least n dimensional. Therefore, A^n is the zero matrix.

(II) Let r be the smallest positive integer so that A^r is the zero matrix. Since A^{r-1} is not the zero matrix, it must have a nonzero column. Suppose that the i-th column is nonzero, then since $A^{r-1}\vec{e}_i$ is the i-th column of A^{r-1}, it is a nonzero vector. In other words, there is a vector $\vec{v}$ with $A^{r-1}\vec{v} \neq \vec{0}$

Look at the r vectors $\vec{v}$, $A\vec{v}$, $A^2\vec{v}$, $\ldots$, $A^{r-1}\vec{v}$. We want to see that they are linearly independent. Suppose that there is a linear relation

$$c_0\vec{v} + c_1A\vec{v} + c_2A^2\vec{v} + \cdots + c_{r-1}A^{r-1}\vec{v} = \vec{0}$$

Multiplying through by A^{r-1} we get

$$c_0A^{r-1}\vec{v} + c_1A^r\vec{v} + c_2A^{r+1}\vec{v} + \cdots + c_{r-1}A^{2r-2}\vec{v} = \vec{0}$$

But since A^r is the zero matrix, the above sum is just $c_0A^{r-1}\vec{v} = \vec{0}$. Since $A^{r-1}\vec{v} \neq \vec{0}$, this means that $c_0 = 0$, so the supposed linear relation is now of the form

$$c_1A\vec{v} + c_2A^2\vec{v} + \cdots + c_{r-1}A^{r-1}\vec{v} = \vec{0}$$

Multiplying through by A^{r-2}, we get $c_1A^{r-1}\vec{v} = \vec{0}$, and so $c_1 = 0$ too. Continuing by induction, we can show that of the c_i's are zero, and hence the vectors are linearly independent.

But, since $\mathbb{R}^n$ is n-dimensional, it can have at most n linearly independent vectors. Above we constructed r linearly independent vectors, so we must have $r \leq n$, i.e., A^n must be the zero matrix, since the smaller power A^r already is.