We can use composition of linear maps
to define <u>matrix multiplication</u> and give a conceptual proof that it is <u>associative.</u>

Let $A = [a_{ij}] \in F_n^m$ and $B = [b_{jk}] \in F_p^n$ so get diagram

$$\begin{tikzcd} F^p \arrow[r, "L_B"] \arrow[dr, "L_A \circ L_B = L_C"'] & F^n \arrow[d, "L_A"] \\ & F^m \end{tikzcd}$$

Can we find $C = [c_{ik}] \in F_p^m$ such that $L_C = L_A \circ L_B$?

If so, what is C in terms of A and B ?

What would C have to be ?

This would mean that $\forall\, Y = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix} \in F^p$,
$CY = A(BY)$ since
$L_C(Y) = L_A(L_B(Y))$. Let

$$X = BY = \begin{bmatrix} \sum_{k=1}^{p} b_{1k} y_k \\ \vdots \\ \sum_{k=1}^{p} b_{nk} y_k \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in F^n$$ so that

$$AX = A(BY) = \begin{bmatrix} \sum_{j=1}^{n} a_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj} x_j \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} \sum_{k=1}^{p} a_{1j} b_{jk} y_k \\ \vdots \\ \sum_{j=1}^{n} \sum_{k=1}^{p} a_{mj} b_{jk} y_k \end{bmatrix} =$$

$$\begin{bmatrix} \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{1j} b_{jk}\right) y_k \\ \vdots \\ \sum_{k=1}^{p}\left(\sum_{j=1}^{n} a_{mj} b_{jk}\right) y_k \end{bmatrix} = CY \text{ iff } \boxed{c_{ik}} = \sum_{j=1}^{n} a_{ij} b_{jk}$$

for $1 \le i \le m$, $1 \le k \le p$.

<u>Th</u>: For any $A = [a_{ij}] \in F_n^m$, $B = [b_{jk}] \in F_p^n$, the matrix $C = [c_{ik}] \in F_p^m$ such that $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$ satisfies $L_C = L_A \circ L_B$.

<u>Def</u>: Let the matrix C above be called the matrix product of A and B, denoted $C = AB$.

When we defined $L_A : F^n \to F^m$ from a choice of $A \in F^m_n$, we should have considered this question: If $A, B \in F^m_n$ and $L_A = L_B$, does $A = B$ have to be true?

$L_A = L_B$ means $\forall X \in F^n$, $AX = BX$, so in particular, looking back at the definition on p. 10, let
$$X = e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{row } j$$
be the column vector with 1 in row j, 0 in all other rows.

Then $AX = \sum_{j=1}^{n} x_j \, Col_j(A)$ says $Ae_j = Col_j(A)$ and $BX = Col_j(B)$. So $Col_j(A) = Col_j(B)$ for all $1 \le j \le n$, which means $A = B$.

We have proven:

<u>Th</u>: For $A, B \in F^m_n$, if $L_A = L_B$ then $A = B$.

We can now prove that matrix multiplication is associative, and see that it comes from the associativity of composition of functions.

<u>Th</u>: Suppose $A \in F^m_n$, $B \in F^n_p$ and $C \in F^p_q$, so $AB \in F^m_p$, $BC \in F^n_q$, $(AB)C \in F^m_q$ and $A(BC) \in F^m_q$. Then $(AB)C = A(BC)$.

<u>Pf</u>. By definition, $L_{AB} = L_A \circ L_B$, $L_{BC} = L_B \circ L_C$, $L_{(AB)C} = L_{AB} \circ L_C$ and $L_{A(BC)} = L_A \circ L_{BC}$ so

$$L_{(AB)C} = (L_A \circ L_B) \circ L_C = L_A \circ (L_B \circ L_C) = L_{A(BC)}.$$

The middle "=" is from assoc. of composition and the last Theorem tells us that $L_{(AB)C} = L_{A(BC)}$ implies $(AB)C = A(BC)$. $\square$

<u>Note</u>: For $A = [a_{ij}] \in F^m_n$ and $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in F^n_1 = F^n$ our definition of $AX \in F^m$ on page 10 is a special case of the matrix multiplication on page 64. In fact, the direct connection is

$$\underset{(m\times n)(n\times p)}{AB} = \left[A\,Col_1(B) \,\middle|\, A\,Col_2(B) \,\middle|\, \cdots \,\middle|\, A\,Col_p(B) \right] = C$$

that is, $Col_k(AB) = A\,Col_k(B)$ for $1 \le k \le p$.

Example:

$$\underset{\substack{A \\ (2\times 3)}}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}} \underset{B\,(3\times 2)}{\begin{bmatrix} -1 & 1 \\ 5 & 9 \\ 4 & 6 \end{bmatrix}} = \underset{AB \quad (2\times 2)}{\begin{bmatrix} (1)(-1)+(2)(5)+(3)(4) & (1)(1)+(2)(9)+(3)(6) \\ (4)(-1)+(5)(5)+(6)(4) & (4)(1)+(5)(9)+(6)(6) \end{bmatrix}}$$

$$= \begin{bmatrix} (-1+10+12) & (1+18+18) \\ (-4+25+24) & (4+45+36) \end{bmatrix} = \begin{bmatrix} 21 & 37 \\ 45 & 85 \end{bmatrix}$$

while

$$\underset{A}{\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}} \underset{Col_1(B)}{\begin{bmatrix} -1 \\ 5 \\ 4 \end{bmatrix}} = \begin{bmatrix} (1)(-1)+(2)(5)+(3)(4) \\ (4)(-1)+(5)(5)+(6)(4) \end{bmatrix} = Col_1(AB) = \begin{bmatrix} 21 \\ 45 \end{bmatrix}$$

and similarly, $A\, Col_2(B) = Col_2(AB) = \begin{bmatrix} 37 \\ 85 \end{bmatrix}$.

Matrix Algebra.

From the definitions it is easy to prove the basic laws of matrix algebra relating addition, scalar multiplication and matrix mult. Besides associativity of matrix mult. we also have: For appropriate size matrices:

Distributive laws: $A(B+C) = AB + AC$,

$$(A+B)C = AC + BC$$

For $\alpha \in F$, $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

Special matrices: Already defined the $m \times n$ "zero matrix", $O_n^m \in F_n^m$ whose entries are all $0 \in F$.

Def. Matrices in F_n^n are called "square".

Def. In F^n_n the "identity matrix" is $I_n = [\delta_{ij}] = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots 1 \end{bmatrix}$ with 1 on the "main diagonal", 0 elsewhere.

Th. ① $A I_n = A$ for any $A \in F^m_n$

② $I_m A = A$ " " "

③ $A O^n_p = O^m_p$, $\forall A \in F^m_n$

④ $O^m_n A = O^m_p$, $\forall A \in F^n_p$.

EX: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O^2_2$

⑤ $L_{I_n} : F^n \to F^n$ is the identity map I_{F^n}

⑥ $L_{O^m_n} : F^n \to F^m$ is the "zero map" s.t.

$L_{O^m_n}(X) = O^m_1$.

Def. For $A=[a_{ij}]\in F^m_n$ define the <u>transpose of A</u> to be $B=[b_{ji}]\in F^n_m$ s.t. $b_{ji}=a_{ij}$, and denote this $n\times m$ matrix by A^T.

Th: For appropriate size matrices we have

① $(A+B)^T=A^T+B^T$ ② $(\alpha A)^T=\alpha(A^T)$, $\alpha\in F$

③ $(A^T)^T=A$ ④ $(AB)^T=B^TA^T$ (sizes!)

Def. Say A is <u>symmetric</u> when $A^T=A$. so $m=n$ for such a matrix, it must be square.

Say A is <u>anti-symmetric</u> (skew-symmetric) when $A^T=-A$. (Such an A must be square)

Ex: $\begin{bmatrix}0&1\\-1&0\end{bmatrix}$ is anti-sym.

Question: For $A \in F^n_n$ (square), when is $L_A : F^n \to F^n$ invertible? What is the condition on A for this to happen?

Answer: There must be an inverse function $L_A^{-1} : F^n \to F^n$ such that $L_A \circ L_A^{-1} = I_{F^n} = L_A^{-1} \circ L_A$. If $L_A^{-1} = L_B$ for some $B \in F^n_n$ this would mean $L_A \circ L_B = L_{I_n} = L_B \circ L_A$ so $L_{AB} = L_{I_n} = L_{BA}$ so $AB = I_n = BA$.

Def. For $A \in F^n_n$ say A is <u>invertible</u> when $\exists B \in F^n_n$ such that $AB = I_n = BA$.

Problem: Given $A \in F_n^n$ determine whether or not A is invertible, and find B if it is.

Th: If $A \in F_n^n$ is invertible, there is only one (unique) matrix B such that $AB = I_n = BA$ so we can denote it by A^{-1} if it exists.

Pf. Suppose we have two candidates for an inverse of A, say $AB = I_n = BA$ and $AC = I_n = CA$. Then, by assoc. of matrix mult.,

$$C = CI_n = C(AB) = (CA)B = I_nB = B. \quad \square$$

Example: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in F_2^2$. Compute

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} (ad-bc) & 0 \\ 0 & (ad-bc) \end{bmatrix} = (ad-bc) I_2$$

and $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} (ad-bc) & 0 \\ 0 & (ad-bc) \end{bmatrix} = (ad-bc)I_2$

So if $ad-bc \neq 0$ in F, it has a mult. inverse (reciprocal) in F denoted by $(ad-bc)^{-1} = \dfrac{1}{ad-bc}$ and if we multiply these equations by it, we get

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

<u>Exercise</u>: Show that for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in F_2^2$ if $ad-bc = 0$ then A is not invertible.

<u>Def.</u> Let $\det(A) = \det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad-bc$ for $A \in F_2^2$. It "determines" if A is invertible.

How do we find A^{-1}, if it exists, for $A \in F_n^n$?

Since $B = A^{-1}$ has to satisfy $AB = I_n$, we would need the columns of B to satisfy $A\,Col_j(B) = e_j$ since $Col_j(I_n) = e_j \in F_1^n$. For each $1 \leq j \leq n$ we need to solve lin. sys. $AX = e_j$.

But they all have the same coeff. matrix A so it would be efficient to do just one row reduction of $[A \mid e_1 e_2 \cdots e_n] = [A \mid I_n]$.

If $[A \mid I_n]$ row reduces to $[\underset{RREF}{C} \mid B] = [I_n \mid B]$ (iff $\text{rank}(A) = n$) then B is the inverse of A.

If $\text{rank}(A) = r < n$ then C has a zero row so $C \neq I_n$ and A is <u>not</u> invertible.