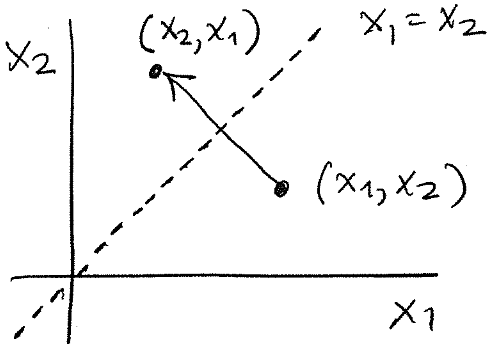
Sec. 3.7 Linear transformations

① Examples and Definition

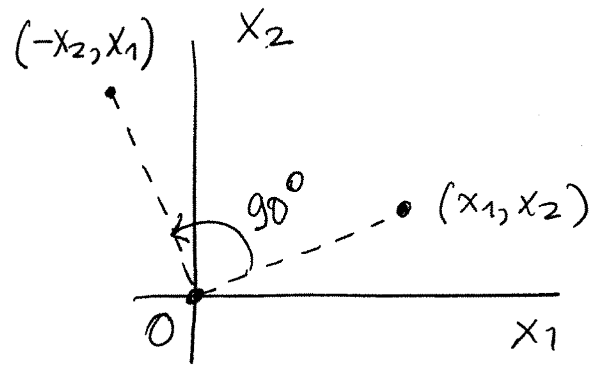
Ex. 1 Reflection about $x_1 = x_2$:



$$T: \mathbb{R}^2 \to \mathbb{R}^2$$

$$T\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

Ex. 2 Rotation about origin by $90°$:

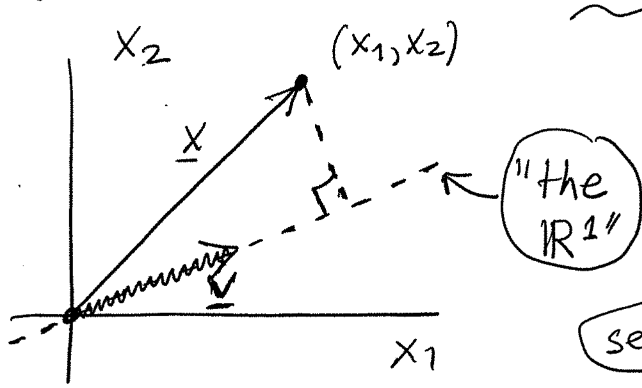


$$T: \mathbb{R}^2 \to \mathbb{R}^2$$

$$T\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$$

(For a rotation by θ, the transformation is found on p. 236 of textbook; see also Project 1.)

Ex. 3 Projection on direction of $\underline{v}$



$$T: \mathbb{R}^2 \to \mathbb{R}^1$$

$$T\Big[\underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\underline{x}}\Big] = P_{\underline{v}}(\underline{x})$$

$$= \frac{\underline{v}^T \underline{x}}{\|\underline{v}\|^2} \cdot \underline{v}.$$

sec. 3.6

Def Let V and W be subspaces of $\mathbb{R}^n$.
$T: V \to W$ is a **linear transformation (l.t.)** if for any $\underline{x}, \underline{y}$ in V it satisfies:

(a) $T(\underline{x}+\underline{y}) = T(\underline{x}) + T(\underline{y})$

(b) $T(c \cdot \underline{x}) = c \cdot T(\underline{x})$ for any scalar c.

Corollary 1

$$(T \text{ is a l.t.}) \iff \left(T(a\underline{x}+b\underline{y}) = aT(\underline{x}) + bT(\underline{y}) \right)$$

Corollary 2 If T is a l.t., then for any vectors $\underline{v}_1, \dots \underline{v}_p$ and scalars $c_1, \dots c_p$:

$$T(c_1\underline{v}_1 + \dots + c_p\underline{v}_p) = c_1 T(\underline{v}_1) + \dots + c_p T(\underline{v}_p)$$

How can we tell if a given T is a l.t.?
We can check properties (a) & (b); see Ex. 1.
However, **a more efficient method will be shown in topic ②**, coming up next.

② Matrix of a l.t.

Revisit Ex. 1 and rewrite the result of T as:

$$T\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \equiv \begin{pmatrix} \overset{0}{\cancel{a_{11}}}\, x_1 + \overset{1}{\cancel{a_{12}}}\, x_2 \\ \underset{1}{\cancel{a_{21}}}\, x_1 + \underset{0}{\cancel{a_{22}}}\, x_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

some matrix, **independent of $\underline{x}$**

Similarly, in Ex. 2: $T\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Thus, in both cases, for all $\underline{x}$:

$$T(\underline{x}) = A\,\underline{x}$$

some matrix, independent of $\underline{x}$

Claim:

$$(T \text{ is l.t.}) \Leftrightarrow (T(\underline{x}) = A\,\underline{x})$$

Proof (only of "<=")

All we need to do is check that $T(\underline{x}) = A\underline{x}$ satisfies properties (a) and (b).

(a) $A(\underline{x}+\underline{y}) \overset{?}{=} A\underline{x} + A\underline{y}$

This is true by the distributive property.

(b) $A(c\cdot\underline{x}) \overset{?}{=} c\,A\,\underline{x}$

This is true because one can move a scalar before or after A.

q.e.d.

Thus, by this Claim, the transformations in Ex. 1 & 2 are both l.t.

Ex. 4 Let us show that the projection, considered in Ex. 3, is a l.t.. According to the Claim, we need to exhibit matrix A (independent of $\underline{x}$!) such that $P_{\underline{v}}(\underline{x}) = A\,\underline{x}$.

Sol'n: 1) We first simplify the formula for $P_{\underline{v}}(\underline{x})$. The motivation for this simplification is as follows. When we say "projection **on** $\underline{v}$", we actually mean "projection on the **direction** of $\underline{v}$". So, the length of $\underline{v}$ (but not the length of $\underline{x}$!) does not matter, and so what we do below is eliminate the length of $\underline{v}$ from the formula.

$$P_{\underline{v}}(\underline{x}) = \frac{\underline{v}^T \underline{x}}{\|\underline{v}\|^2} \cdot \underline{v} = \frac{\underline{v}^T}{\|\underline{v}\|} \cdot \underline{x} \cdot \frac{\underline{v}}{\|\underline{v}\|} = (\underline{w}^T \underline{x}) \cdot \underline{w},$$

where $\boxed{\underline{\mathbf{w}} = \underline{\mathbf{v}} / \|\underline{\mathbf{v}}\|}$ is the unit vector along $\underline{v}$.

So,

$$\boxed{P_{\underline{\mathbf{v}}}(\underline{\mathbf{x}}) \equiv P_{\underline{\mathbf{w}}}(\underline{\mathbf{x}}) = (\underline{\mathbf{w}}^T \mathbf{x}) \cdot \underline{\mathbf{w}}}$$

2) We now transform this into the form $A \cdot \underline{x}$.

$$\underbrace{(\underline{w}^T \underline{x})}_{\text{scalar}} \cdot \underline{w} = \underline{w} \cdot (\underline{w}^T \underline{x}) \overset{\text{associative rule}}{=} \underbrace{(\underline{w}\,\underline{w}^T)}_{\text{matrix independent of } \underline{x}} \underline{x}$$

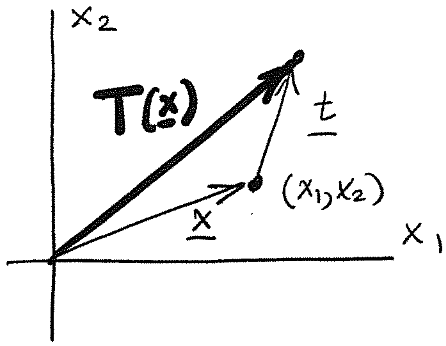
Thus,

$$\left(\begin{array}{l}\textbf{the matrix of projection} \\ \textbf{on the direction of } \underline{\mathbf{v}}\end{array}\right) = \underline{\mathbf{w}}\,\underline{\mathbf{w}}^T$$

where $\underline{w} = \underline{v} / \|\underline{v}\|$, the unit vector along $\underline{v}$.

And, according to the Claim on p. 15-3, projection is a l.t..

Ex. 5 Translation is not a l.t.



Translation moves (translates) all points on the plane by the same vector (say, $\underline{t}$).
So: $T(\underline{x}) = \underline{x} + \underline{t}$
(see the figure; recall Calculus).

There is no way to write $\underline{x} + \underline{t} = A \cdot \underline{x}$ that would work for all $\underline{x}$ with an A that is independent of $\underline{x}$ (if you are not convinced, try to find $a = \text{const}$ such that $x + 1 = a \cdot x$ for all x).

3 l.t. and expansion over a basis

Let $\{\underline{v}_1, \dots \underline{v}_p\}$ be a basis for some subspace W of $\mathbb{R}^n$, and let $\underline{x}$ be any vector in W. We know that $\underline{x}$ can be expanded over a basis:

$$\underline{x} = c_1 \underline{v}_1 + \dots + c_p \underline{v}_p,$$

where $c_1, \dots c_p$ are **coordinates** of $\underline{x}$ **in basis** $\{\underline{v}_1, \dots \underline{v}_p\}$.
Then for some l.t. T:

$$T(\underline{x}) = T(c_1 \underline{v}_1 + \dots + c_p \underline{v}_p) \overset{\text{Corollary 2 on p. 15-2}}{=}$$

$$= c_1 T(\underline{v}_1) + \dots + c_p T(\underline{v}_p).$$

This implies that we now have
two ways to define the action of a l.t.:

$\boxed{1}$ As per the Claim on p. 15-3, we can define $T(\underline{x}) = A\underline{x}$ (which is equivalent to what we did in Ex. 1 & 2).

$\boxed{2}$ If we know a basis $\{\underline{v_1}, \dots \underline{v_p}\}$ in W and know how a l.t. transforms it, i.e. know $\{T(\underline{v_1}), \dots T(\underline{v_p})\}$, we can find **for any $\underline{x}$**:

$$\mathbf{T(\underline{x}) = \underbrace{c_1 T(\underline{v_1}) + \dots + c_p}_{\text{coordinates of } \underline{x} \text{ in basis } \{\underline{v_1}, \dots \underline{v_p}\}} T(\underline{v_p})}$$

<u>Ex. 6</u> (see also Ex. 6 in book)

Let $T\left[\begin{pmatrix}-1\\0\end{pmatrix}\right] = \begin{pmatrix}3\\4\end{pmatrix}$, $T\left[\begin{pmatrix}2\\1\end{pmatrix}\right] = \begin{pmatrix}5\\6\end{pmatrix}$.

Find $T\left[\begin{pmatrix}8\\9\end{pmatrix}\right]$.

<u>Sol'n:</u> 1) $\underbrace{\begin{pmatrix}-1\\0\end{pmatrix}}_{\underline{v_1}}, \underbrace{\begin{pmatrix}2\\1\end{pmatrix}}_{\underline{v_2}}$ are not proportional,

$\Rightarrow$ not $\parallel$ $\Rightarrow$ they form a basis in $\mathbb{R}^2$.

Expand $\underline{x} = \begin{pmatrix}8\\9\end{pmatrix}$ over this basis:

$$\begin{pmatrix}8\\9\end{pmatrix} = c_1\begin{pmatrix}-1\\0\end{pmatrix} + c_2\begin{pmatrix}2\\1\end{pmatrix} \Rightarrow$$

$$\left(\begin{array}{cc|c} -1 & 2 & 8 \\ 0 & 1 & 9 \end{array}\right) \Rightarrow \begin{array}{l} -c_1 + \overset{18}{2c_2} = 8 \Rightarrow c_1 = 10 \\ c_2 = 9 \leftarrow \text{by inspection} \end{array}$$

<u>Note:</u> When a <u>problem like this</u> is given on a quiz or test, you <u>are allowed</u> to find c_1, c_2 not by REF but <u>by inspection</u>. <u>In all other problems</u> you <u>must</u> still <u>use REF</u>.

2) Now we can use the formula:

$$T(\underline{x}) = c_1 T(\underline{v_1}) + c_2 T(\underline{v_2}) = 10\begin{pmatrix}3\\4\end{pmatrix} + 9\begin{pmatrix}5\\6\end{pmatrix} = \begin{pmatrix}75\\94\end{pmatrix}.$$

The two ways, $\boxed{1}$ and $\boxed{2}$, to define T must be equivalent. We can easily "go from $\boxed{1}$ to $\boxed{2}$": if we know the A, we can find

$\{T(\underline{v_1}), \ldots T(\underline{v_p})\} = \{A\underline{v_1}, \ldots A\underline{v_p}\}$.

But how can we "go from $\boxed{2}$ to $\boxed{1}$"? I.e.:

Given: $\{\underline{v_1}, \ldots \underline{v_p}\}$ and $\{T(\underline{v_1}), \ldots T(\underline{v_p})\}$, Find: A.

Solution: we seek A s.t.

$$\underbrace{T(\underline{v_1})}_{\underline{u_1}} = A\underline{v_1}, \ldots \underbrace{T(\underline{v_p})}_{\underline{u_p}} = A\underline{v_p}, \Rightarrow$$

$\underline{u_1} \leftarrow$ denote $\rightarrow \underline{u_p}$

$$\underbrace{[\underline{u_1}, \ldots \underline{u_p}]}_{U} = [A\underline{v_1}, \ldots A\underline{v_p}] = A\underbrace{[\underline{v_1}, \ldots \underline{v_p}]}_{V}$$

$$\Rightarrow (U = AV)V^{-1} \Rightarrow UV^{-1} = A(\cancel{VV^{-1}}) \rightarrow I$$

$$\Rightarrow \boxed{A = UV^{-1}}$$

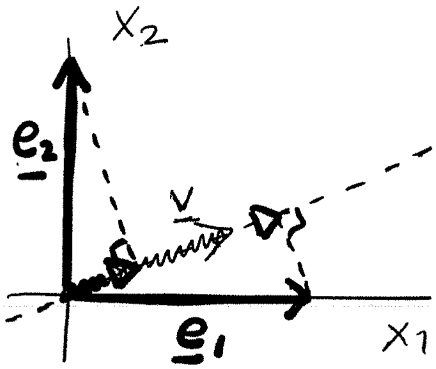
Note 1: V^{-1} exists because $\{\underline{v_1}, \ldots \underline{v_p}\}$ are lin. indep., $\Rightarrow$ V is nonsingular.

Note 2: When $\{\underline{v_1}, \underline{v_2}\} = \{\underline{e_1}, \underline{e_2}\}$ (the natural basis (pp. 195-196 of book), then $V = I$. Then $A = U \equiv [T(\underline{e_1}), T(\underline{e_2})]$ (Thm. 15 in book)

Note 3: While $\{\underline{v}_1, \underline{v}_2\}$ are lin. indep., $\{\underline{u}_1, \underline{u}_2\} = A \cdot \{\underline{v}_1, \underline{v}_2\}$ may be lin. dep. This occurs when A is singular.

Ex. 7 The matrix of the projection (Ex. 4) is singular. ← Show this.

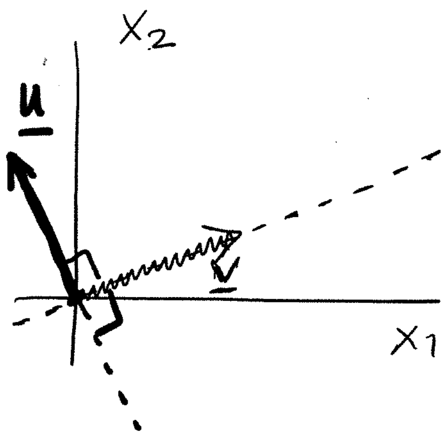
Sol'n: 1) First, it's clear that for any basis (e.g., for $\{\underline{e}_1, \underline{e}_2\}$), its projections on the direction of $\underline{v}$ are parallel (since they are on the same line).



2) To demonstrate that the matrix A of the projection is singular, we need to show that for some $\underline{u} \neq \underline{\theta}$,

$$A\underline{u} \equiv P_{\underline{v}}(\underline{u}) = \underline{\theta}.$$

This is true for any $\underline{u} \perp \underline{v}$ (see figure).



One can write: $\mathcal{N}(A) = \{\underline{x} : \underline{x} \perp \underline{v}\}$.

Note: The matrices of l.t. in Ex. 1, 2 were nonsingular.

- **MUST READ pp. 233-234 about null space and range of a l.t.**
  - **range of T** = all possible outcomes of T;
  - **null space of T** = all $\underline{x}$ s.t. $T(\underline{x}) = \underline{\theta}$ (Ex. 7).

4) Composition of l.t.

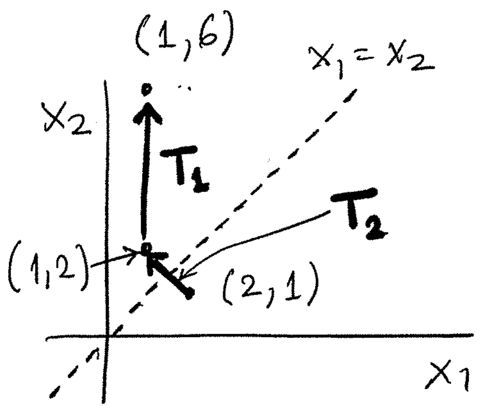
In sec. 1.5 we learned that $AB \neq BA$ for matrices in general. We can now illustrate this.

Ex. 8 Let T_1 be the vertical stretch by a factor of 3 in $\mathbb{R}^2$, and T_2 be the reflection about $x_1 = x_2$. Find the composition of T_1 & T_2 and then of T_2 & T_1, and verify that they yield different results.

Sol'n: 1) $T_2(\underline{x}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \underline{x}, \Rightarrow A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (p. 15-2)

$$T_1\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right] = \begin{pmatrix} x_1 \\ 3x_2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \Rightarrow A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

2) $T_1 \circ \mathbf{T_2} \leftarrow$ composition of T_1 & T_2, where $\mathbf{T_2}$ **is applied first:**



$$(T_1 \circ T_2)(\underline{x}) \equiv T_1(T_2(\underline{x}))$$

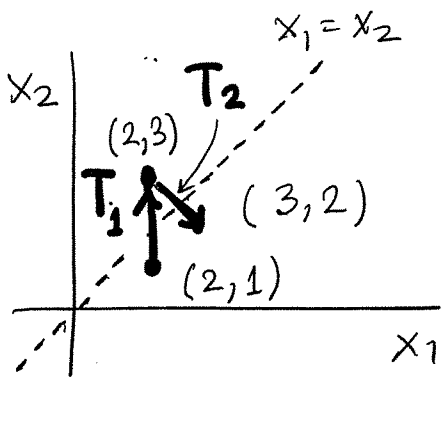
$$T_2\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$T_1\left(T_2\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right]\right) = T_1\left[\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$$

Via matrix multiplication:

$$T_1 \circ T_2(\underline{x}) = A_1(A_2\underline{x}) = A_1A_2\,\underline{x} \qquad \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}$$

3) $T_2 \circ \mathbf{T_1} \leftarrow$ composition of T_2 & T_1, where $\mathbf{T_1}$ **is applied first:**



$$T_1\left[\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$T_2 \circ T_1\begin{pmatrix} 2 \\ 1 \end{pmatrix} = T_2\left[\begin{pmatrix} 2 \\ 3 \end{pmatrix}\right] = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Via matrix multiplication:

$$T_2 \circ T_1(\underline{x}) = A_2A_1\,\underline{x} \qquad \begin{pmatrix} 0 & 3 \\ 1 & 0 \end{pmatrix}$$

This example explains visually why
$A_1 A_2 \neq A_2 A_1$.
In words, **the result of a composition of two l.t. depends on the order in which the l.t. are performed**.
Note: This visual example also allows us to see when two transformations (and hence their matrices) will commute. For example, if we perform two consecutive rotations, by angles θ_1 and θ_2, about the origin, the result will not depend in which order we perform these rotations. So, if R_θ is the matrix of rotation by angle θ (see p. 236), then $R_{\theta_1} \cdot R_{\theta_2} = R_{\theta_2} \cdot R_{\theta_1}$.

5 Orthogonal l.t.

Def: A l.t. $T(\underline{x})$ s.t.

$$\| T(\underline{x}) \| = \| \underline{x} \| \quad \text{for all } \underline{x}$$

is called an orthogonal l.t.

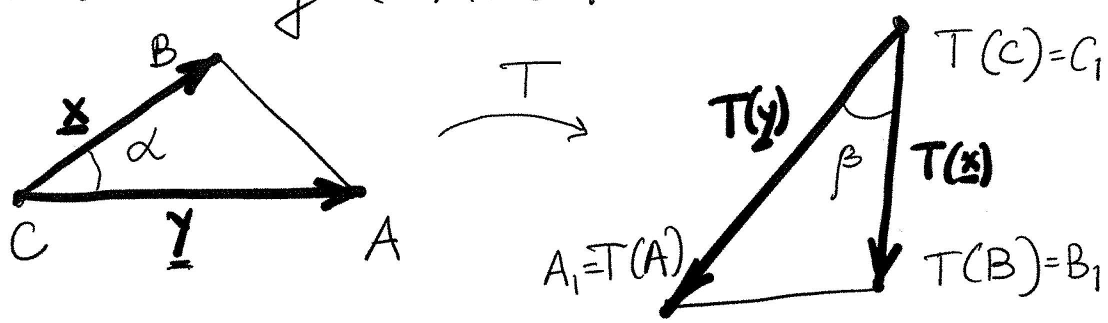
Ex. 9 A reflection (about any line through the origin) and a rotation about the origin are orthogonal l.t. See pp. 236-237 in book. Moreover, they are the only orthogonal l.t. in $\mathbb{R}^2$. In $\mathbb{R}^3$, orthogonal l.t. also include only rotations and reflections, but there are more possibilities there.

Claim: **An orthogonal l.t. preserves the angle between any two vectors.** I.e.,

$$\begin{pmatrix} T = \\ \text{orthogonal l.t.} \end{pmatrix} \Rightarrow \left(\angle(T(\underline{x}), T(\underline{y})) = \angle(\underline{x}, \underline{y}) \right)$$

for any $\underline{x}$, $\underline{y}$.

- One can prove this using the concept of the dot product, similar to the proof of Thm. 16. However, there is a much easier visual proof: Consider any ΔABC.



T transforms it into $\Delta T(A)T(B)T(C) \equiv \Delta A_1B_1C_1$. Since T is orthogonal, the sides of $\Delta A_1B_1C_1$ have the same lengths as the respective sides of ΔABC. Then by comparing the congruent $\Delta A_1B_1C_1 \cong \Delta ABC$, we have $\angle A_1C_1B_1 = \angle ACB$ (i.e., $\angle(T(\underline{x}), T(\underline{y})) = \angle(\underline{x}, \underline{y})$), and similarly for the other angles.

q.e.d.

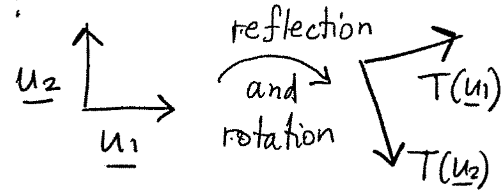
Corollary (= Thm. 16): An **orthogonal l.t.** transforms **an orthonormal basis into another orthonormal basis:**

$$\underbrace{\begin{pmatrix} T = \\ \text{orthogonal l.t.} \end{pmatrix} \text{ and } \begin{pmatrix} \{\underline{u}_1, \underline{u}_2\} = \text{orthonormal}, \\ \text{i.e. } \|\underline{u}_1\| = \|\underline{u}_2\| = 1, \\ \underline{u}_1 \perp \underline{u}_2 \end{pmatrix}}$$

$$\Downarrow$$

$$\begin{pmatrix} \{T(\underline{u}_1), T(\underline{u}_2)\} = \text{orthonormal, i.e.} \\ \|T(\underline{u}_1)\| = \|T(\underline{u}_2)\| = 1 \text{ and } T(\underline{u}_1) \perp T(\underline{u}_2) \end{pmatrix}$$

Note 1 : You can easily visualize this in $\mathbb{R}^2$, where reflections and rotations (which are the only orthogonal l.t.; see p. 15-10) preserve the orthonormality of a basis:



Note 2 : The same Claim also holds in any $\mathbb{R}^n$.

Note 3 The matrix of an orthogonal l.t. is called the orthogonal matrix. We will consider them (defined differently, but called the same name) in Sec. 4.7.
One can also show, using a manipulation of orthogonal matrices, that a composition of two orthogonal l.t. is again an orthogonal l.t. (the the figure above on this page).