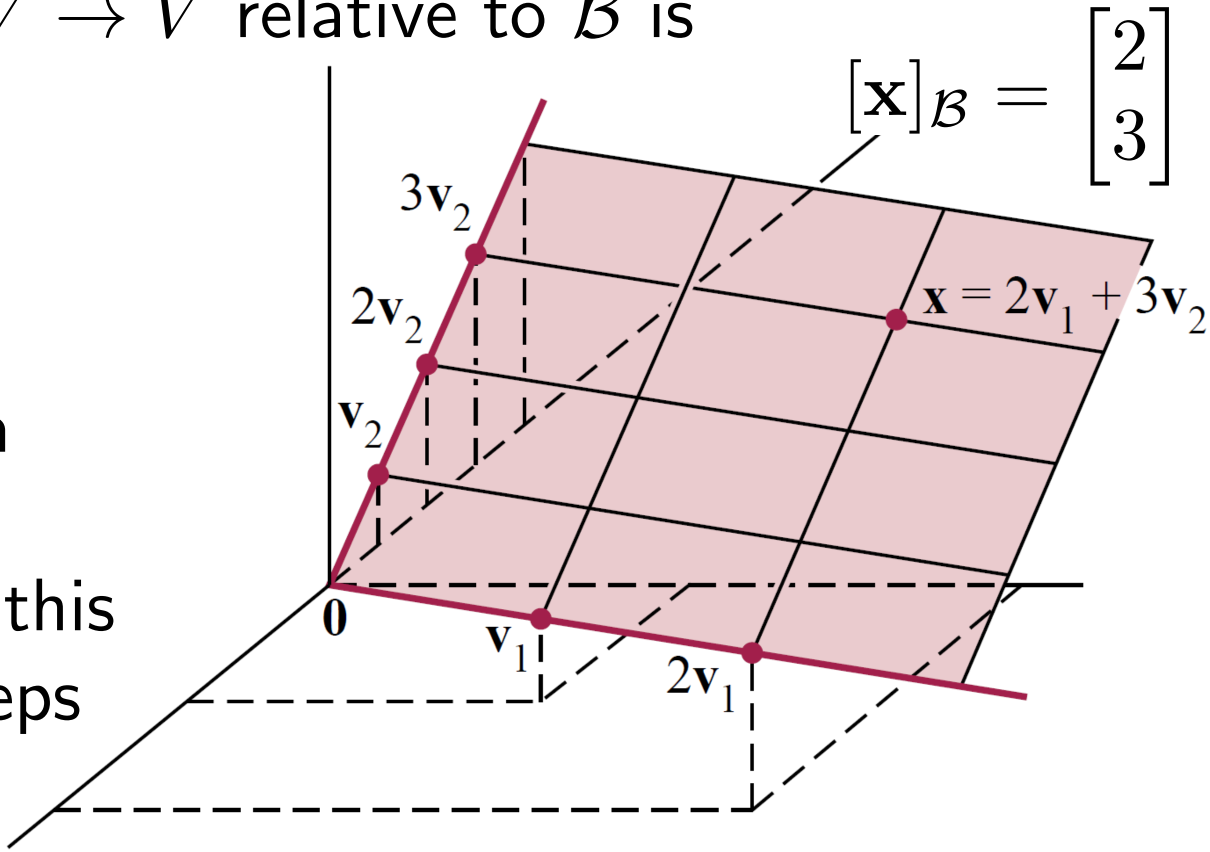
§4.4, 4.7, 5.4: Change of Basis

Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for V. Remember:

- The $\mathcal{B}$-coordinate vector of $\mathbf{x}$ is $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ where $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.
- The matrix for a linear transformation $T : V \to V$ relative to $\mathcal{B}$ is

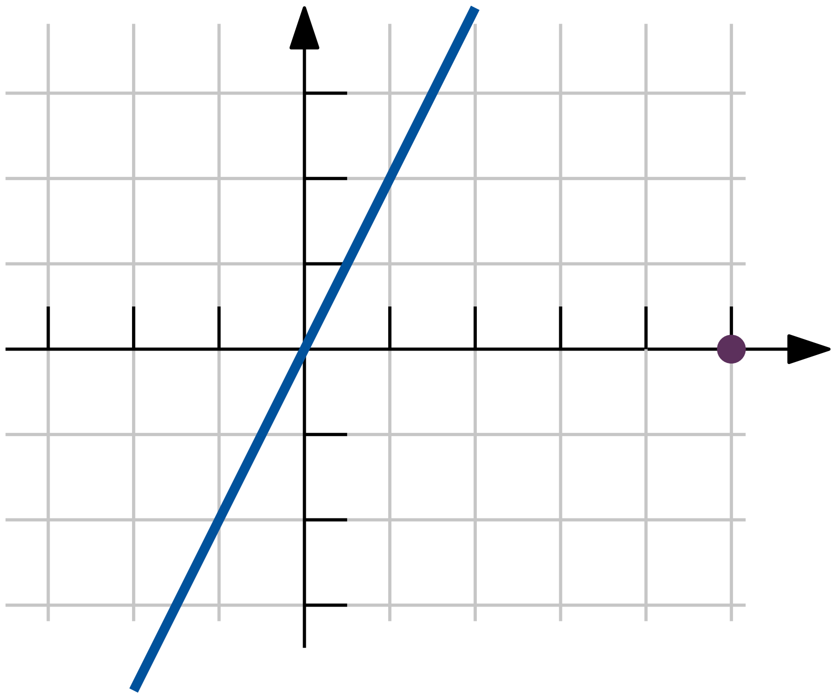
$$[T]_{\mathcal{B}} = \begin{bmatrix} | & & | \\ [T(\mathbf{b}_1)]_{\mathcal{B}} & \ldots & [T(\mathbf{b}_n)]_{\mathcal{B}} \\ | & & | \end{bmatrix}.$$

A basis for this plane in $\mathbb{R}^3$ allows us to draw a coordinate grid on the plane. The coordinate vectors then describe the location of points on this plane relative to this coordinate grid (e.g. 2 steps in $\mathbf{v}_1$ direction, 3 steps in $\mathbf{v}_2$ direction.)



Although we already have the standard coordinate grid on $\mathbb{R}^n$, some computations are much faster and more accurate in a different basis i.e. using a different coordinate grid (see also p18-20):

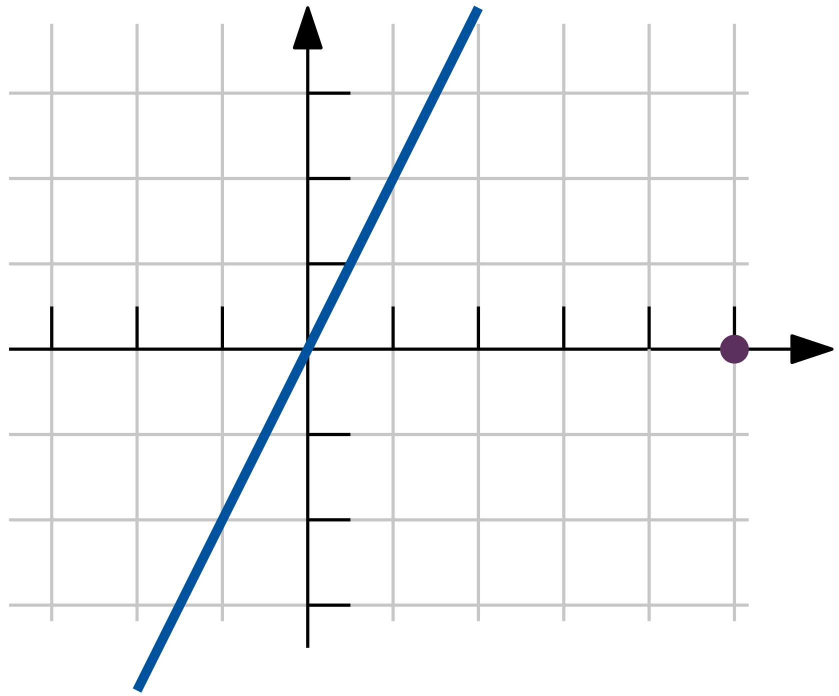
Example: Find the image of the point $(5, 0)$ under reflection about the line $y = 2x$.



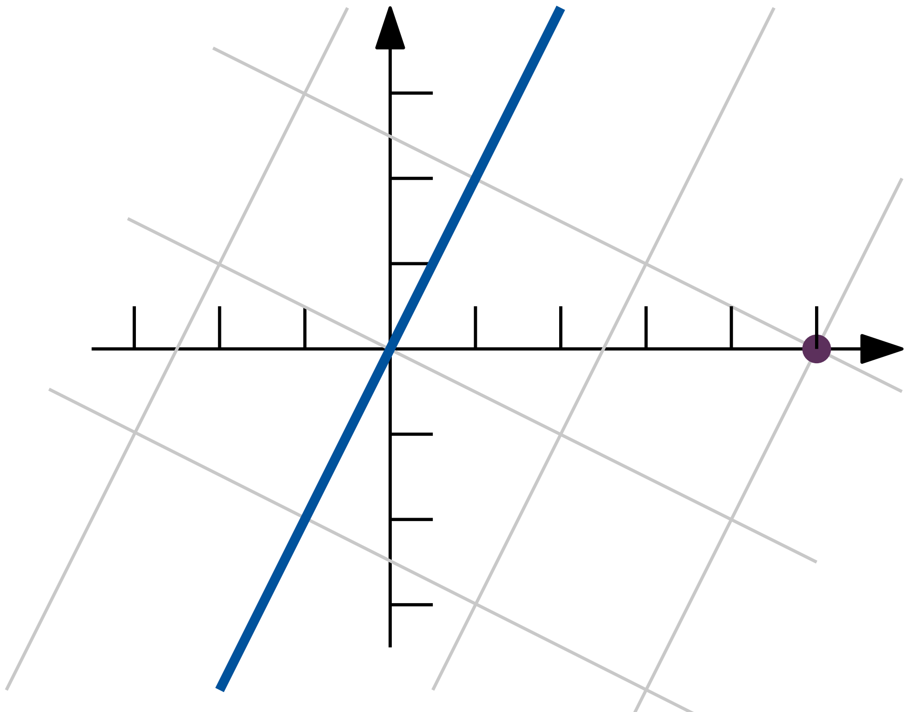
Horizontal and vertical grid lines are not useful for this problem because $y = 2x$ is not horizontal nor vertical.

Although we already have the standard coordinate grid on $\mathbb{R}^n$, some computations are much faster and more accurate in a different basis i.e. using a different coordinate grid (see also p18-20):

Example: Find the image of the point $(5, 0)$ under reflection about the line $y = 2x$.

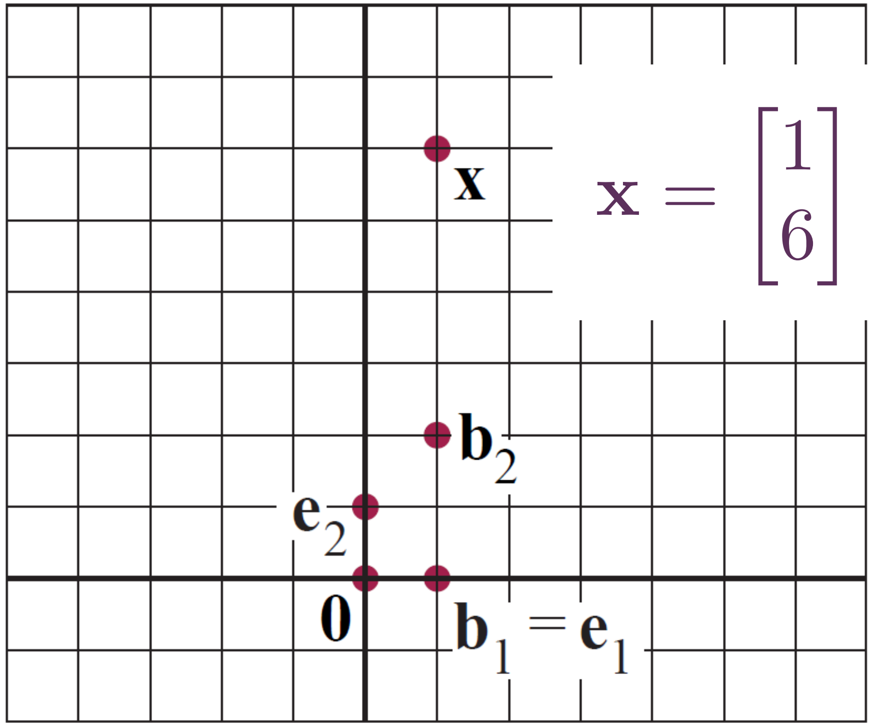


Horizontal and vertical grid lines are not useful for this problem because $y = 2x$ is not horizontal nor vertical.

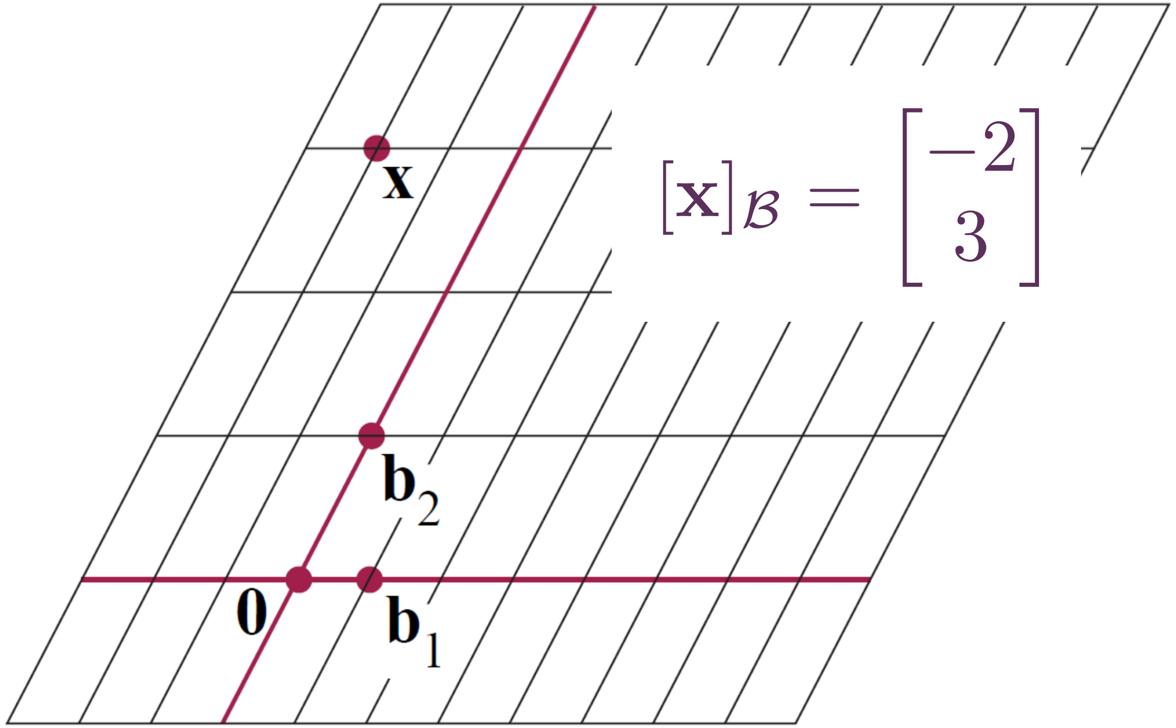


It is more useful to work with lines parallel and perpendicular to $y = 2x$.

Another example of two coordinate grids (note that the lines don't have to be perpendicular):



standard coordinate grid



$\mathcal{B}$-coordinate grid

Important questions:

i how are $\mathbf{x}$ and $[\mathbf{x}]_{\mathcal{B}}$ related (p4-7, §4.4 in textbook);

ii how are $[\mathbf{x}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{F}}$ related for two bases $\mathcal{B}$ and $\mathcal{F}$ (p8-11, §4.7);

iii how are the standard matrix of T and the matrix $[T]_{\mathcal{B}}$ related (p12-16, §5.4).

In general, if $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis for $\mathbb{R}^n$, and $\mathbf{v}$ is any vector in $\mathbb{R}^n$, then

$[\mathbf{v}]_{\mathcal{B}}$ is a solution to $\underbrace{\begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_n \\ | & & | \end{bmatrix}}_{\mathcal{P}_{\mathcal{B}}} \mathbf{x} = \mathbf{v}.$

Because $\mathcal{B}$ is a basis, the columns of $\mathcal{P}_{\mathcal{B}}$ are linearly independent, so by the Invertible Matrix Theorem, $\mathcal{P}_{\mathcal{B}}$ is invertible, and the unique solution to $\mathcal{P}_{\mathcal{B}}\mathbf{x} = \mathbf{v}$ is

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_n \\ | & & | \end{bmatrix}^{-1} \mathbf{v}.$$

In other words, the change-of-coordinates matrix from the standard basis to $\mathcal{B}$ is $\mathcal{P}_{\mathcal{B}}^{-1}$.

Indeed, in the previous example, $\mathcal{P}_{\mathcal{B}}^{-1}\mathbf{x} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 6 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 & -1 \\ -0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}.$

A very common mistake is to get the direction wrong:

Does multiplication by $\mathcal{P}_{\mathcal{B}}$ change from standard coordinates to $\mathcal{B}$-coordinates, or from $\mathcal{B}$-coordinates to standard coordinates?

Don't memorise the formulas. Instead, remember the definition of coordinates:

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \text{ means } \quad \mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \dots & \mathbf{b}_n \\ | & & | \end{bmatrix} [\mathbf{x}]_{\mathcal{B}}$$

and you won't go wrong.

ii: Changing between two non-standard bases:

Example: As before, $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Another basis: $\mathbf{f}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{f}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$.

If $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, then what are its $\mathcal{F}$-coordinates $[\mathbf{x}]_{\mathcal{F}}$?

ii: Changing between two non-standard bases:

Example: As before, $\mathbf{b}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix}1\\2\end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Another basis: $\mathbf{f}_1 = \begin{bmatrix}1\\1\end{bmatrix}, \mathbf{f}_2 = \begin{bmatrix}0\\1\end{bmatrix}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$.

If $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}-2\\3\end{bmatrix}$, then what are its $\mathcal{F}$-coordinates $[\mathbf{x}]_\mathcal{F}$?

Answer 1: $\mathcal{B}$ to standard to $\mathcal{F}$ - works only in $\mathbb{R}^n$, in general easiest to calculate.

$[\mathbf{x}]_\mathcal{B} = \begin{bmatrix}-2\\3\end{bmatrix}$ means $\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2 = -2\begin{bmatrix}1\\0\end{bmatrix} + 3\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}1\\6\end{bmatrix}$.

So if $[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}d_1\\d_2\end{bmatrix}$, then $d_1\begin{bmatrix}1\\1\end{bmatrix} + d_2\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\6\end{bmatrix}$.

Row-reducing $\left[\begin{array}{cc|c}1 & 0 & 1\\1 & 1 & 6\end{array}\right]$ shows $d_1 = 1, d_2 = 5$ so $[\mathbf{x}]_\mathcal{F} = \begin{bmatrix}1\\5\end{bmatrix}$.

In other words, $\mathbf{x} = \mathcal{P}_\mathcal{B}[\mathbf{x}]_\mathcal{B}$ and $[\mathbf{x}]_\mathcal{F} = \mathcal{P}_\mathcal{F}^{-1}\mathbf{x}$, so $[\mathbf{x}]_\mathcal{F} = \mathcal{P}_\mathcal{F}^{-1}\mathcal{P}_\mathcal{B}[\mathbf{x}]_\mathcal{B}$.

Answer 2: A different view that works for abstract vector spaces (without reference to a standard basis) - important theoretically, but may be hard to calculate for general examples in $\mathbb{R}^n$.

$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ means $\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2$.

So $[\mathbf{x}]_{\mathcal{F}} = [-2\mathbf{b}_1 + 3\mathbf{b}_2]_{\mathcal{F}} = -2[\mathbf{b}_1]_{\mathcal{F}} + 3[\mathbf{b}_2]_{\mathcal{F}} = \begin{bmatrix} | & | \\ [\mathbf{b}_1]_{\mathcal{F}} & [\mathbf{b}_2]_{\mathcal{F}} \\ | & | \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

because $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{F}}$ is an isomorphism, so every vector space calculation is accurately reproduced using coordinates.

Answer 2: A different view that works for abstract vector spaces (without reference to a standard basis) - important theoretically, but may be hard to calculate for general examples in $\mathbb{R}^n$.

$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ means $\mathbf{x} = -2\mathbf{b}_1 + 3\mathbf{b}_2$.

So $[\mathbf{x}]_{\mathcal{F}} = [-2\mathbf{b}_1 + 3\mathbf{b}_2]_{\mathcal{F}} = -2[\mathbf{b}_1]_{\mathcal{F}} + 3[\mathbf{b}_2]_{\mathcal{F}} = \begin{bmatrix} | & | \\ [\mathbf{b}_1]_{\mathcal{F}} & [\mathbf{b}_2]_{\mathcal{F}} \\ | & | \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

because $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{F}}$ is an isomorphism, so every vector space calculation is accurately reproduced using coordinates.

$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{f}_1 - \mathbf{f}_2$ so $[\mathbf{b}_1]_{\mathcal{F}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

$\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{f}_1 + \mathbf{f}_2$ so $[\mathbf{b}_2]_{\mathcal{F}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

This step can be hard to calculate if the $\mathbf{b}_i$ are not "easy" linear combinations of the $\mathbf{f}_i$. But if you need to change bases in a practical application, the bases are probably "nicely" related.

So $[\mathbf{x}]_{\mathcal{F}} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$.

Theorem 15: Change of Basis: Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be two bases of a vector space V. Then, for all $\mathbf{x}$ in V,

$$[\mathbf{x}]_{\mathcal{F}} = \begin{bmatrix} | & & | \\ [\mathbf{b}_1]_{\mathcal{F}} & \ldots & [\mathbf{b}_n]_{\mathcal{F}} \\ | & & | \end{bmatrix} [\mathbf{x}]_{\mathcal{B}}.$$

Notation: write $\underset{\mathcal{F} \leftarrow \mathcal{B}}{\mathcal{P}}$ for the matrix $\begin{bmatrix} | & & | \\ [\mathbf{b}_1]_{\mathcal{F}} & \ldots & [\mathbf{b}_n]_{\mathcal{F}} \\ | & & | \end{bmatrix}$, the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{F}$.

A tip to get the direction correct:

$$[\mathbf{x}]_{\mathcal{F}} = \underbrace{\underset{\mathcal{F} \leftarrow \mathcal{B}}{\mathcal{P}} [\mathbf{x}]_{\mathcal{B}}}$$

A $\mathcal{F}$-coordinate vector

a linear combination of columns of $\underset{\mathcal{F} \leftarrow \mathcal{B}}{\mathcal{P}}$, so these columns should be $\mathcal{F}$-coordinate vectors

Theorem 15: Change of Basis: Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{F} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be two bases of a vector space V. Then, for all $\mathbf{x}$ in V,

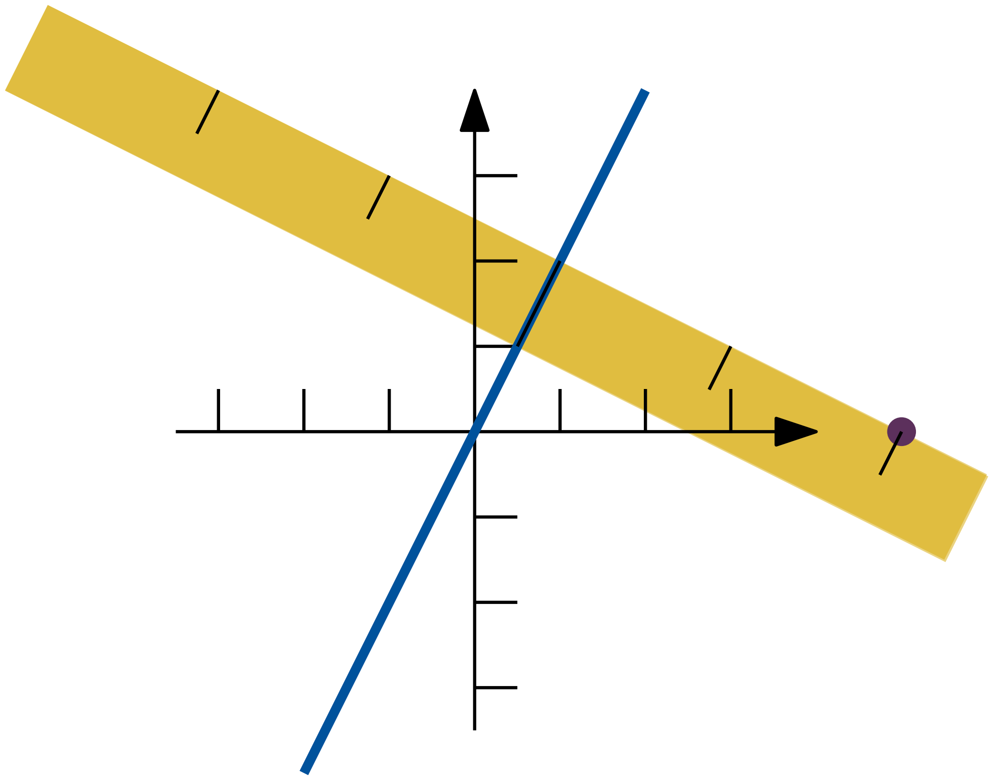
$$[\mathbf{x}]_{\mathcal{F}} = \begin{bmatrix} | & & | \\ [\mathbf{b}_1]_{\mathcal{F}} & \ldots & [\mathbf{b}_n]_{\mathcal{F}} \\ | & & | \end{bmatrix} [\mathbf{x}]_{\mathcal{B}}.$$

Properties of the change-of-coordinates matrix $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}} = \begin{bmatrix} | & & | \\ [\mathbf{b}_1]_{\mathcal{F}} & \ldots & [\mathbf{b}_n]_{\mathcal{F}} \\ | & & | \end{bmatrix}$:

- $\underset{\mathcal{B}\leftarrow\mathcal{F}}{\mathcal{P}} = \underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}$.
- If V is $\mathbb{R}^n$ and $\mathcal{E}$ is the standard basis $\{\mathbf{e}_1, \ldots \mathbf{e}_n\}$, then
  $$\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_{\mathcal{B}} = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \ldots & \mathbf{b}_n \\ | & & | \end{bmatrix}$$
  , because $[\mathbf{b}_i]_{\mathcal{E}} = \mathbf{b}_i$. Also $\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \mathcal{P}_{\mathcal{B}}^{-1}$.
- If V is $\mathbb{R}^n$, then $\underset{\mathcal{F}\leftarrow\mathcal{B}}{\mathcal{P}} = \underset{\mathcal{F}\leftarrow\mathcal{E}}{\mathcal{P}}\,\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_{\mathcal{F}}^{-1}\mathcal{P}_{\mathcal{B}}$ (see p8).

Recall our problem from the start of this week's notes:

Example: Find the image of the point $(5, 0)$ under reflection about the line $y = 2x$.



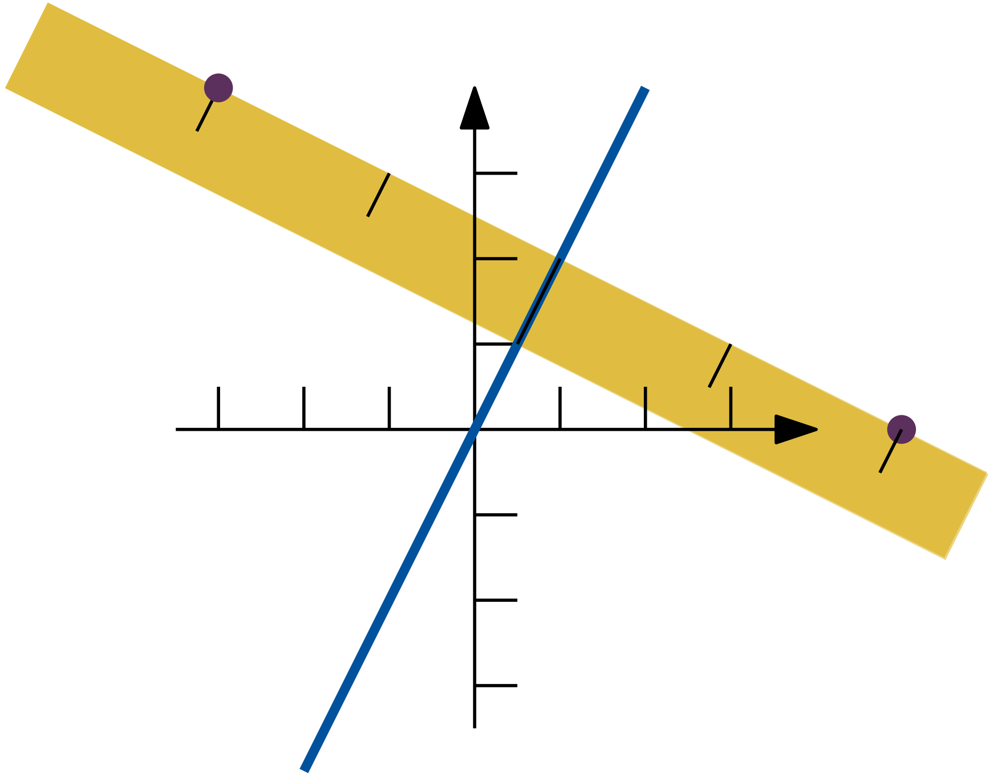
An efficient solution:

1. Measure the perpendicular distance from $(5, 0)$ to the line;
2.
3.

iii: Change of coordinates and linear transformations:

Recall our problem from the start of this week's notes:

Example: Find the image of the point $(5, 0)$ under reflection about the line $y = 2x$.



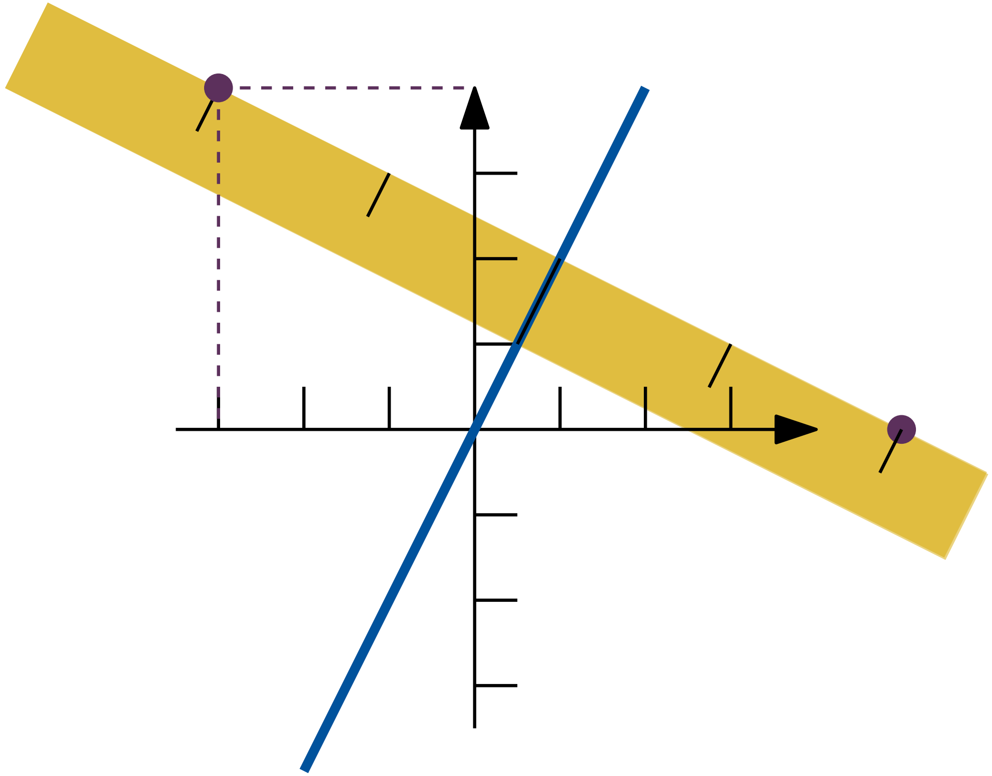
An efficient solution:

1. Measure the perpendicular distance from $(5, 0)$ to the line;
2. The image of $(5, 0)$ is the point that is the same distance away on the other side of the line;
3.

iii: Change of coordinates and linear transformations:

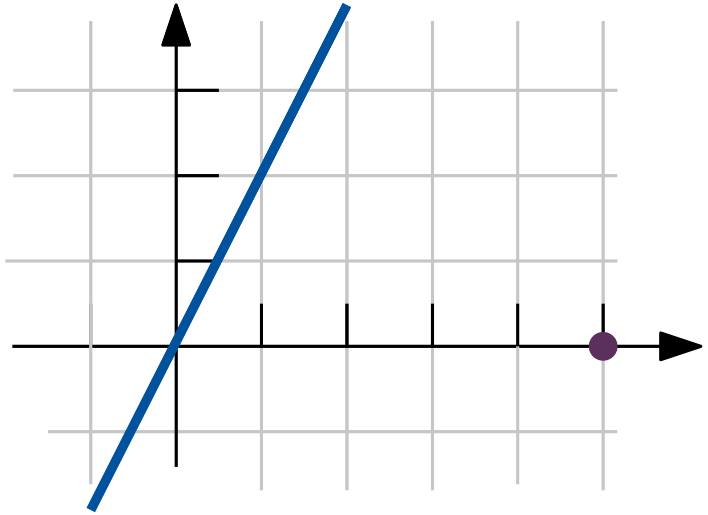
Recall our problem from the start of this week's notes:

Example: Find the image of the point $(5, 0)$ under reflection about the line $y = 2x$.



An efficient solution:

1. Measure the perpendicular distance from $(5, 0)$ to the line;
2. The image of $(5, 0)$ is the point that is the same distance away on the other side of the line;
3. Read off the coordinates of this point: $(-3, 4)$.

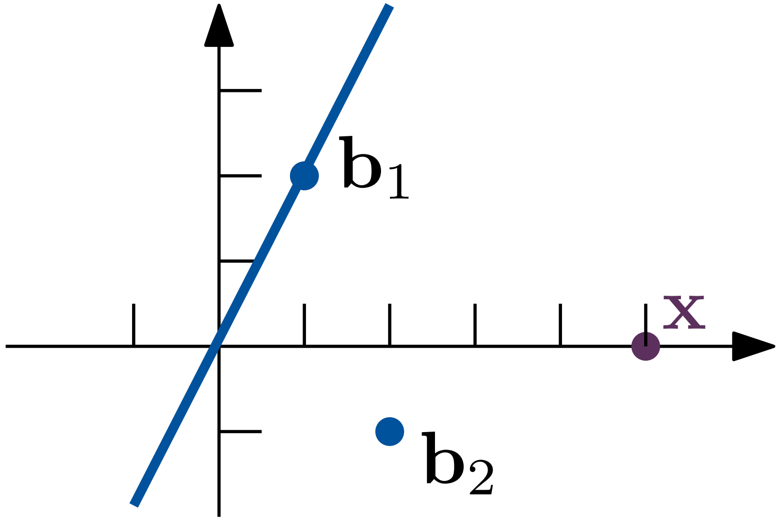


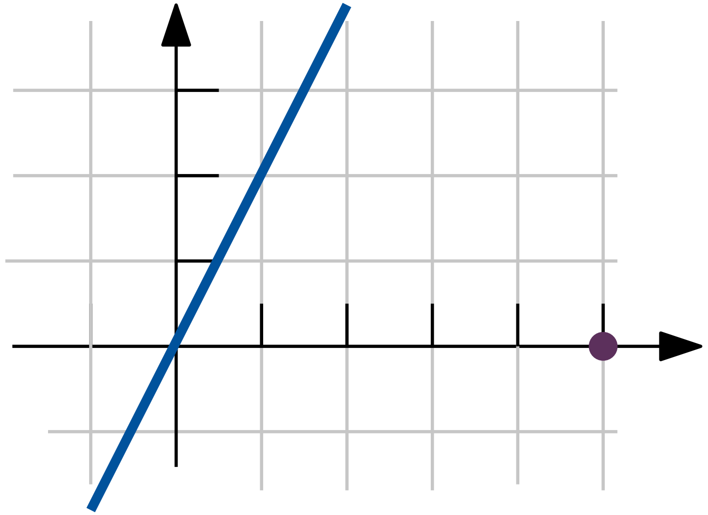
The previous solution in the language of coordinates:

Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

So we want $T(\mathbf{x})$.



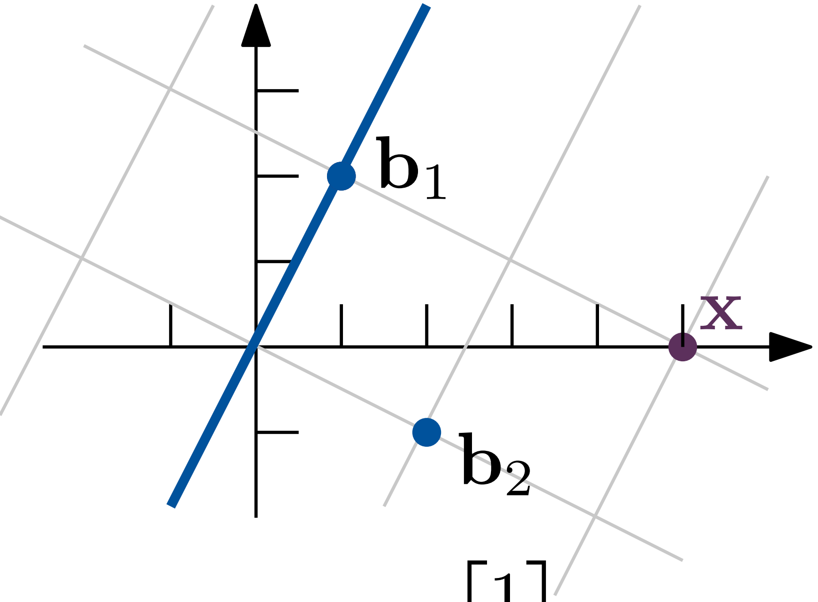


The previous solution in the language of coordinates:

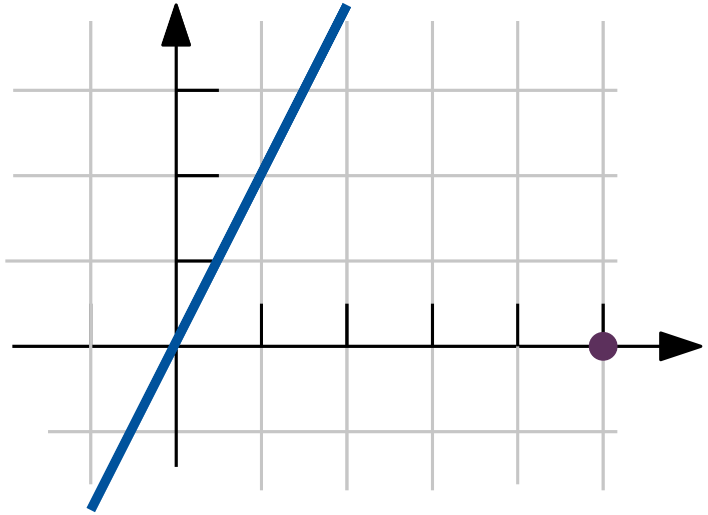
Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

So we want $T(\mathbf{x})$.



1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

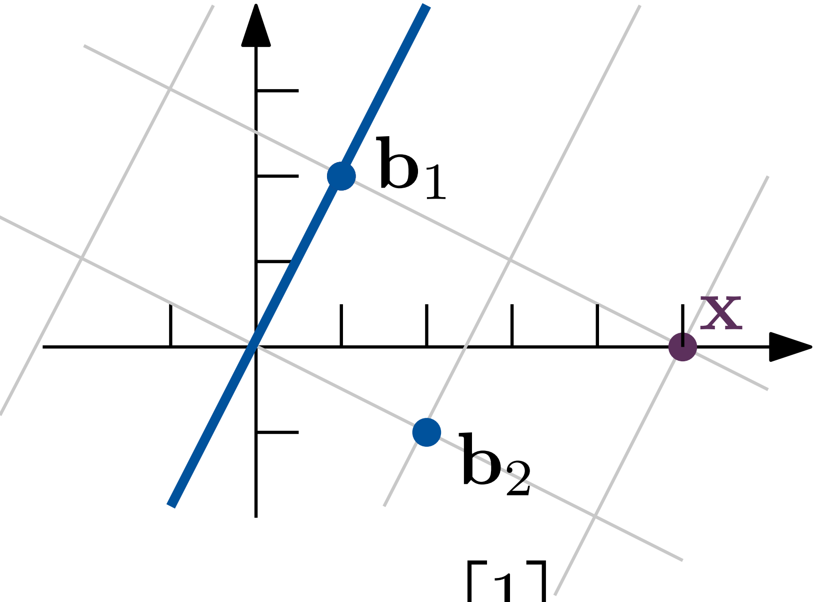


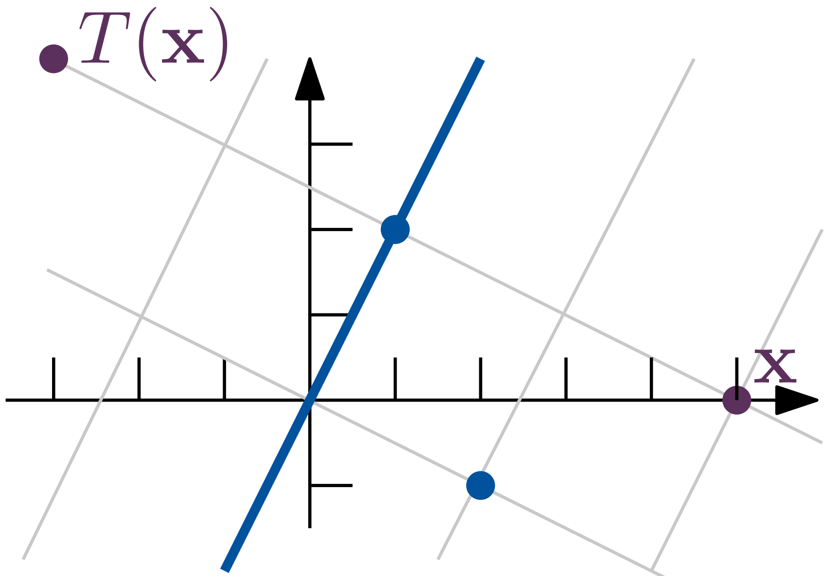
The previous solution in the language of coordinates:

Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

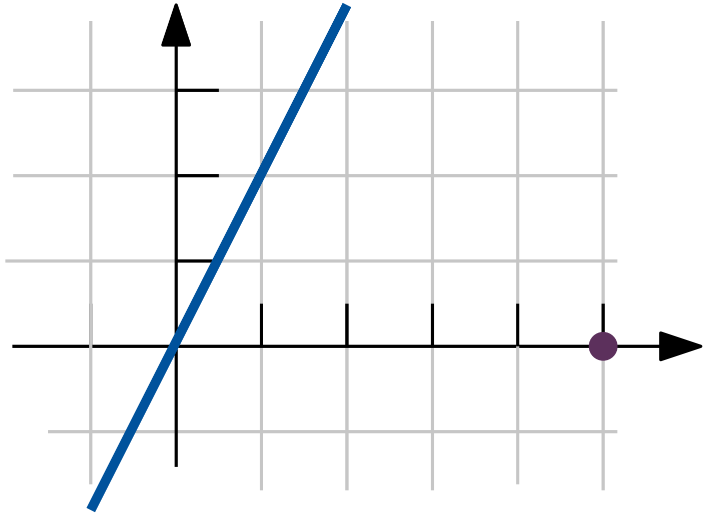
So we want $T(\mathbf{x})$.





1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

2. $[T(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

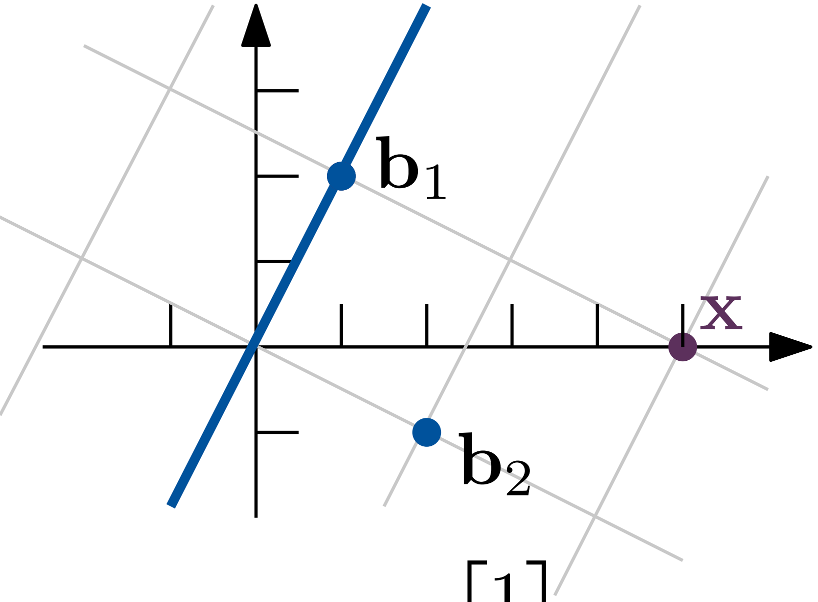


The previous solution in the language of coordinates:

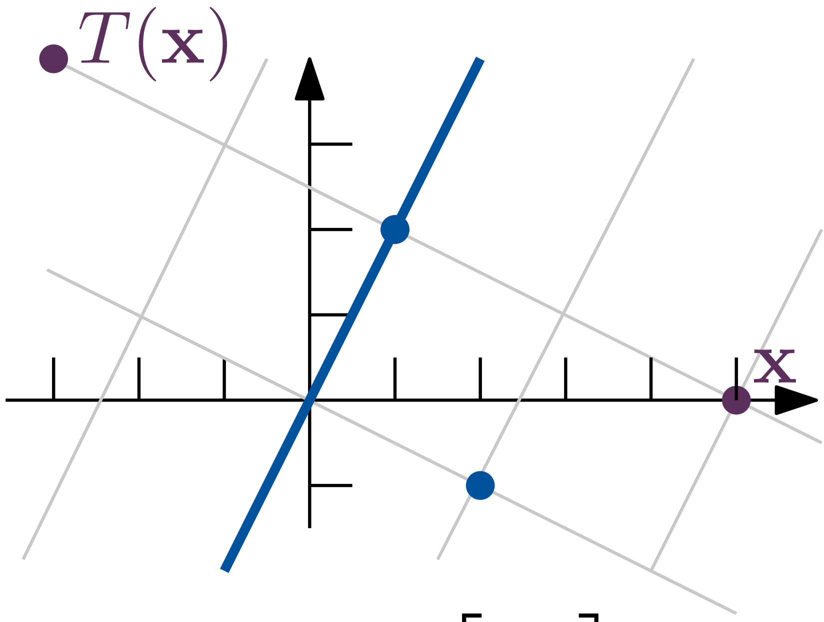
Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

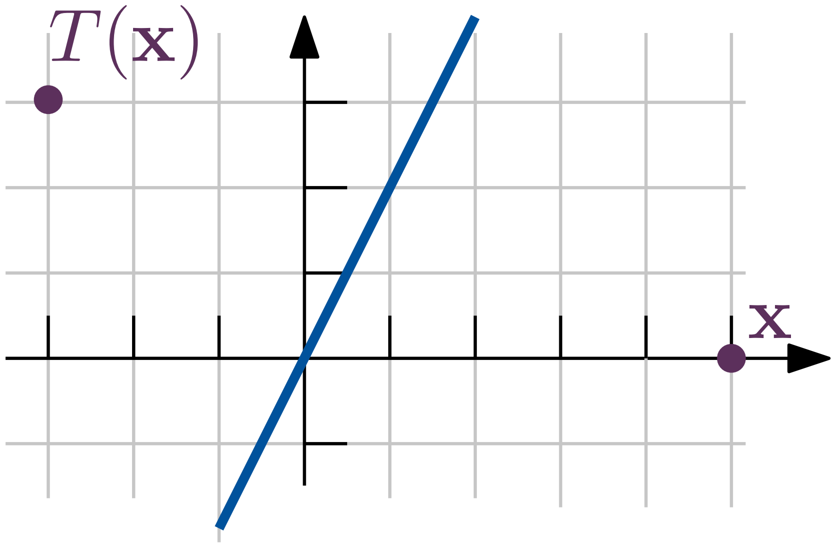
So we want $T(\mathbf{x})$.



1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$



2. $[T(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$



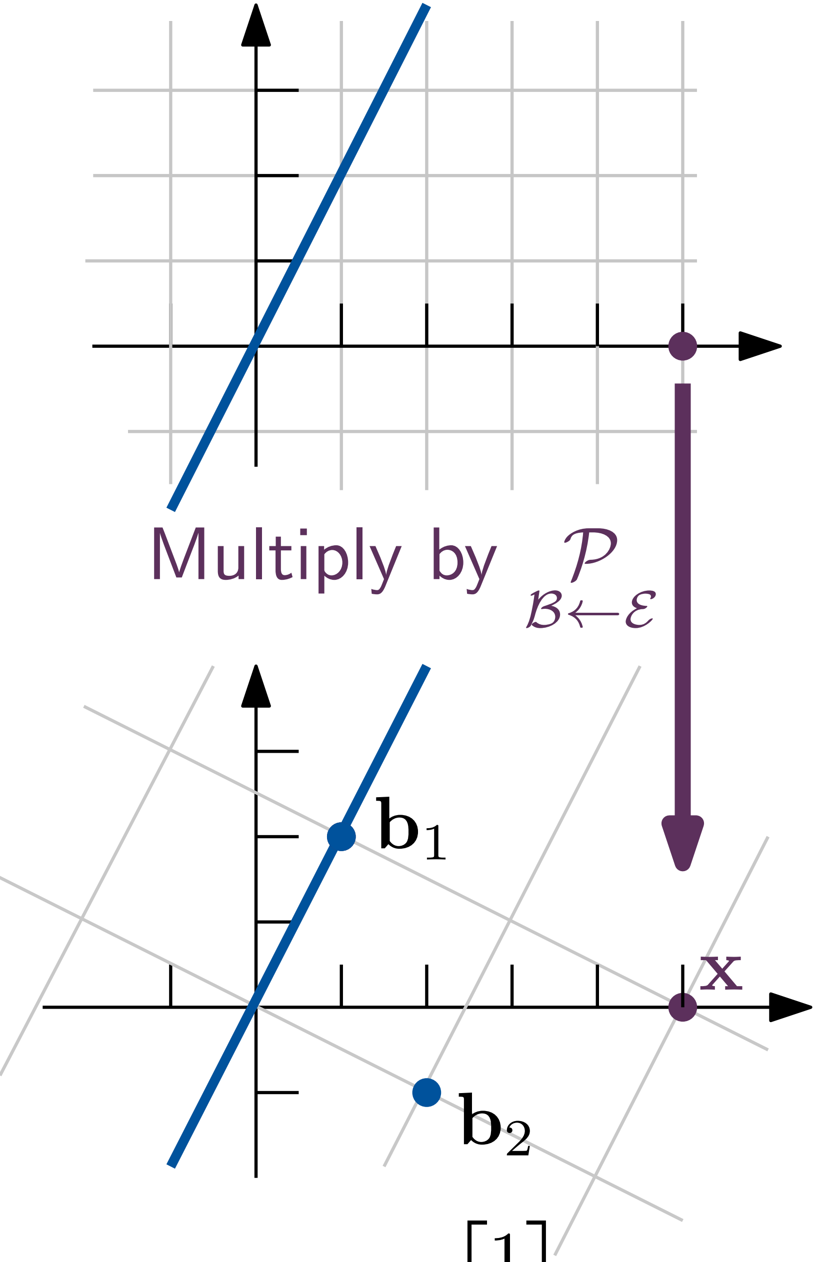
3. $T(\mathbf{x}) = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$

The previous solution in the language of coordinates:

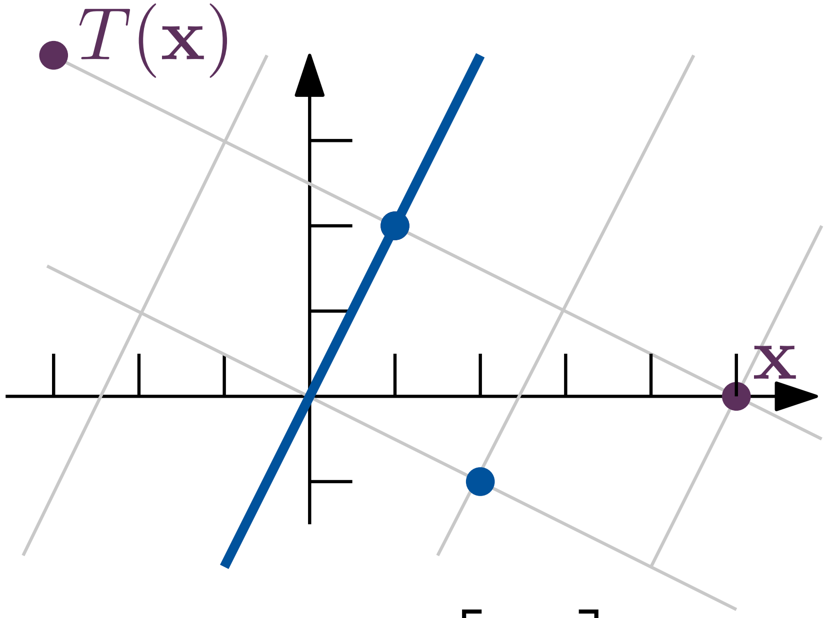
Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

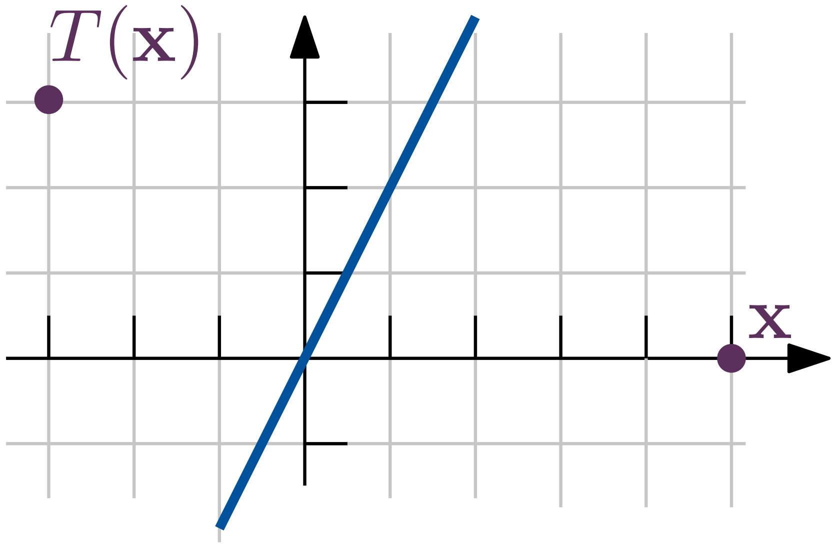
Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

So we want $T(\mathbf{x})$.



In terms of matrix multiplication:





1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

2. $[T(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

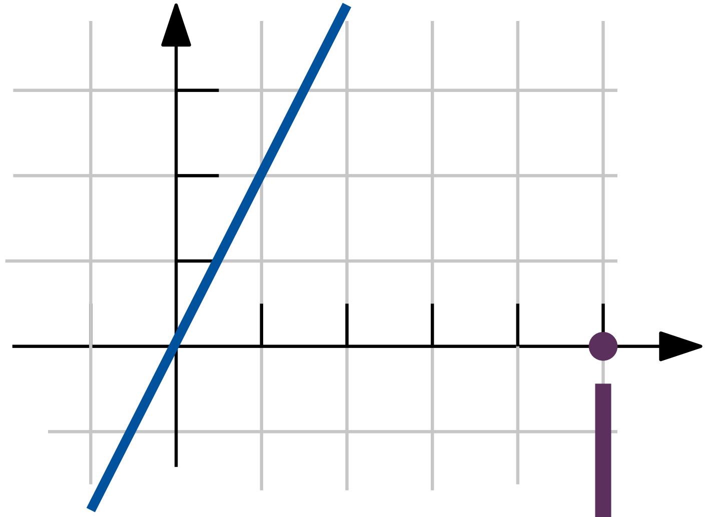
3. $T(\mathbf{x}) = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$

The previous solution in the language of coordinates:

Let $\mathbf{b}_1 = \begin{bmatrix}1\\2\end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix}2\\-1\end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

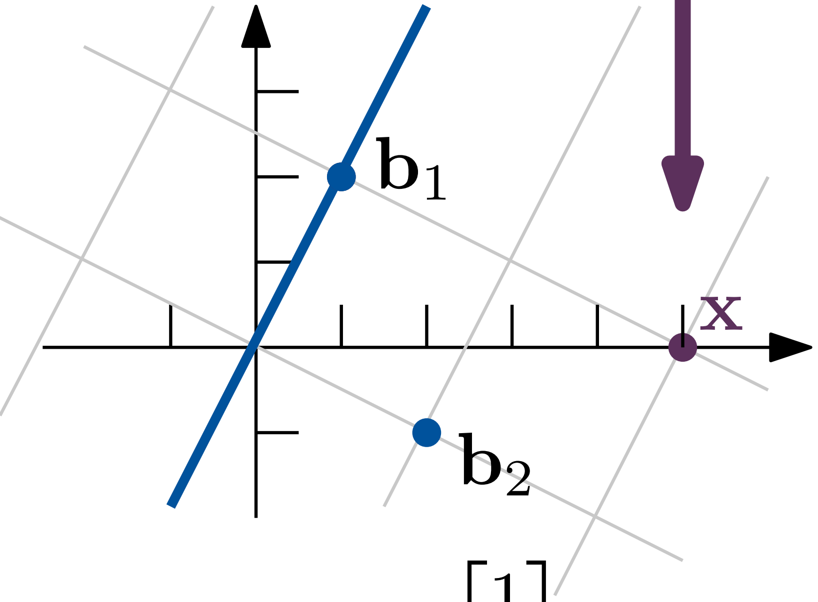
Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix}5\\0\end{bmatrix}$.

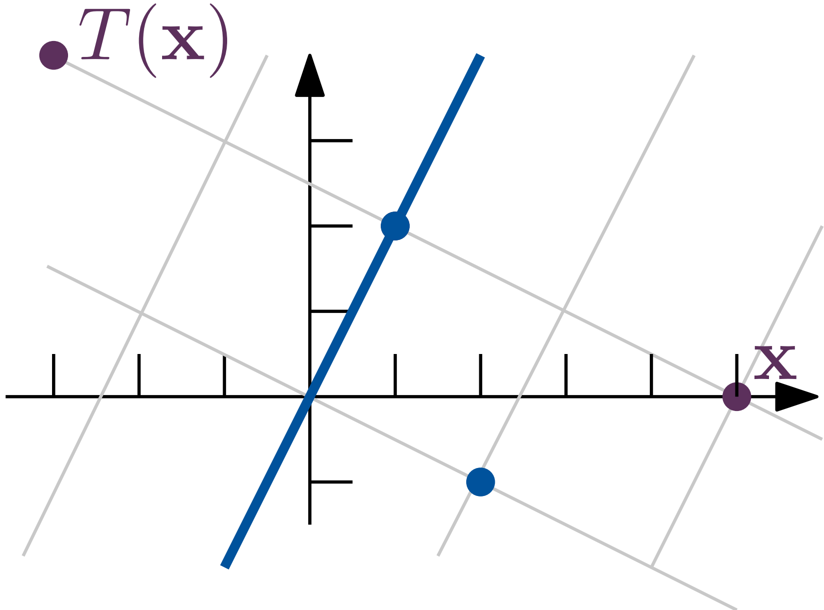
So we want $T(\mathbf{x})$.

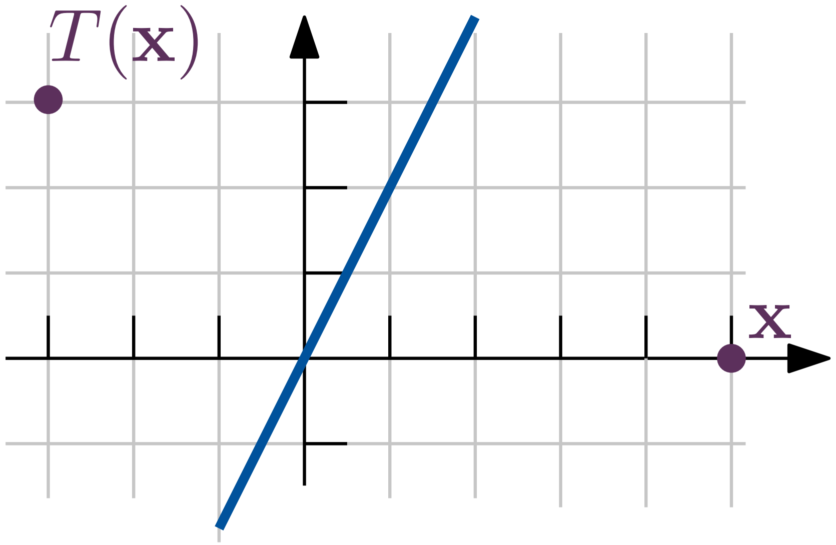


Multiply by $\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}$

In terms of matrix multiplication:







1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}1\\2\end{bmatrix}$

Multiply by $[T]_{\mathcal{B}}$:

2. $[T(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix}1\\-2\end{bmatrix}$

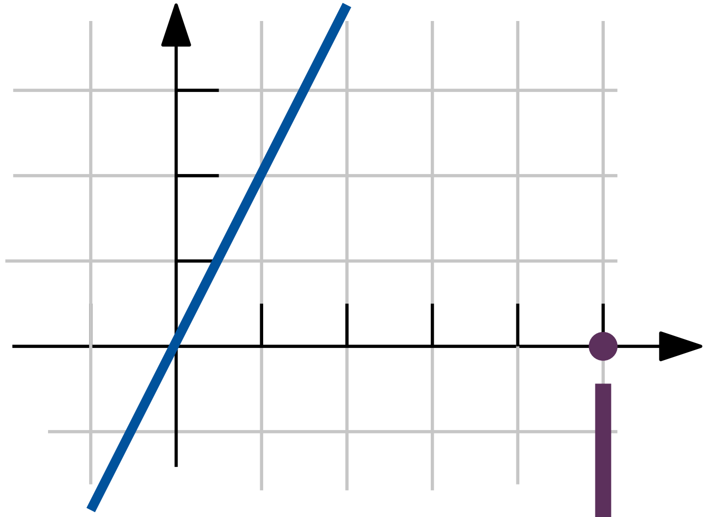
3. $T(\mathbf{x}) = \begin{bmatrix}-3\\4\end{bmatrix}$

The previous solution in the language of coordinates:

Let $\mathbf{b}_1 = \begin{bmatrix}1\\2\end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix}2\\-1\end{bmatrix}$ and work in the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

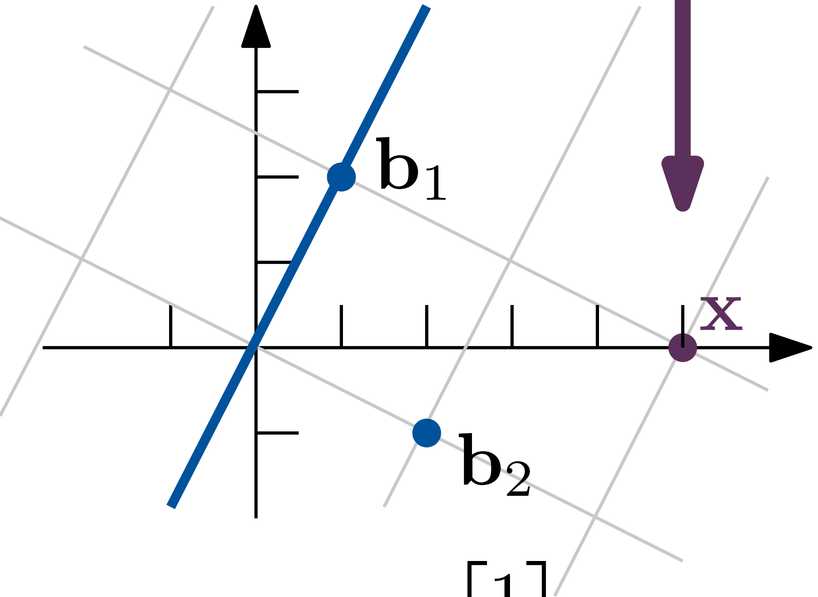
Let T be reflection about the line $y = 2x$, and $\mathbf{x} = \begin{bmatrix}5\\0\end{bmatrix}$.

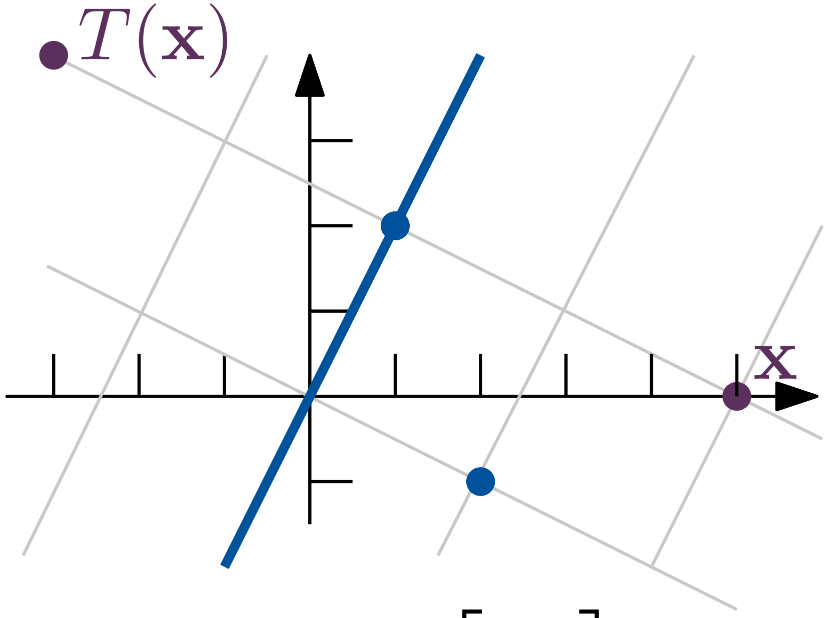
So we want $T(\mathbf{x})$.

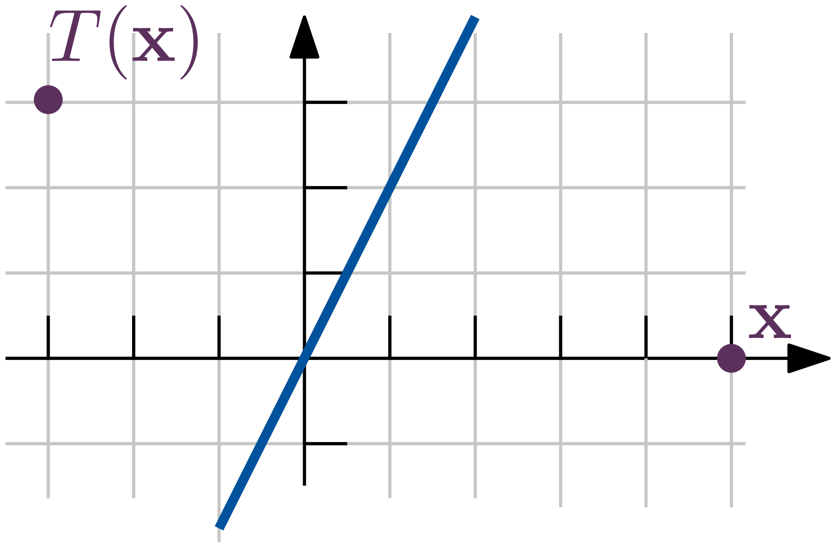


Multiply by $\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}$

In terms of matrix multiplication:







1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}1\\2\end{bmatrix}$ $\longrightarrow$ 2. $[T(\mathbf{x})]_{\mathcal{B}} = \begin{bmatrix}1\\-2\end{bmatrix}$ $\longrightarrow$ 3. $T(\mathbf{x}) = \begin{bmatrix}-3\\4\end{bmatrix}$

Multiply by $[T]_{\mathcal{B}}$:

Multiply by $\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}$

The 3-step solution above shows that $T(\mathbf{x}) = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}\mathbf{x}$.

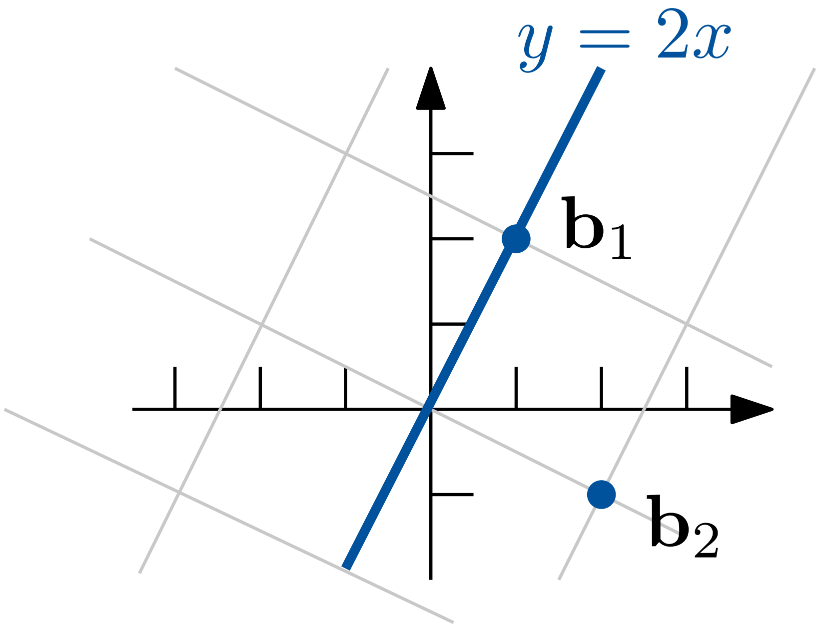
Write $[T]_{\mathcal{E}}$ for the standard matrix of T. Then $T(\mathbf{x}) = [T]_{\mathcal{E}}\mathbf{x}$, so the equation $[T]_{\mathcal{E}}\mathbf{x} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}\mathbf{x}$ is true for all $\mathbf{x}$. So the matrices on the two sides must be equal (e.g. letting $\mathbf{x} = \mathbf{e}_i$ shows that each column of the matrices must be equal)

$$\boxed{[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}.}$$

The 3-step solution above shows that $T(\mathbf{x}) = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}\mathbf{x}$.

Write $[T]_{\mathcal{E}}$ for the standard matrix of T. Then $T(\mathbf{x}) = [T]_{\mathcal{E}}\mathbf{x}$, so the equation $[T]_{\mathcal{E}}\mathbf{x} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}\mathbf{x}$ is true for all $\mathbf{x}$. So the matrices on the two sides must be equal (e.g. letting $\mathbf{x} = \mathbf{e}_i$ shows that each column of the matrices must be equal)

$$\boxed{[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}.}$$



This equation is useful because, for geometric linear transformations T, it is often easier to find $[T]_{\mathcal{B}}$ for some "natural" basis $\mathcal{B}$ than to find the standard matrix $[T]_{\mathcal{E}}$.

E.g. in our example of reflection in $y = 2x$:

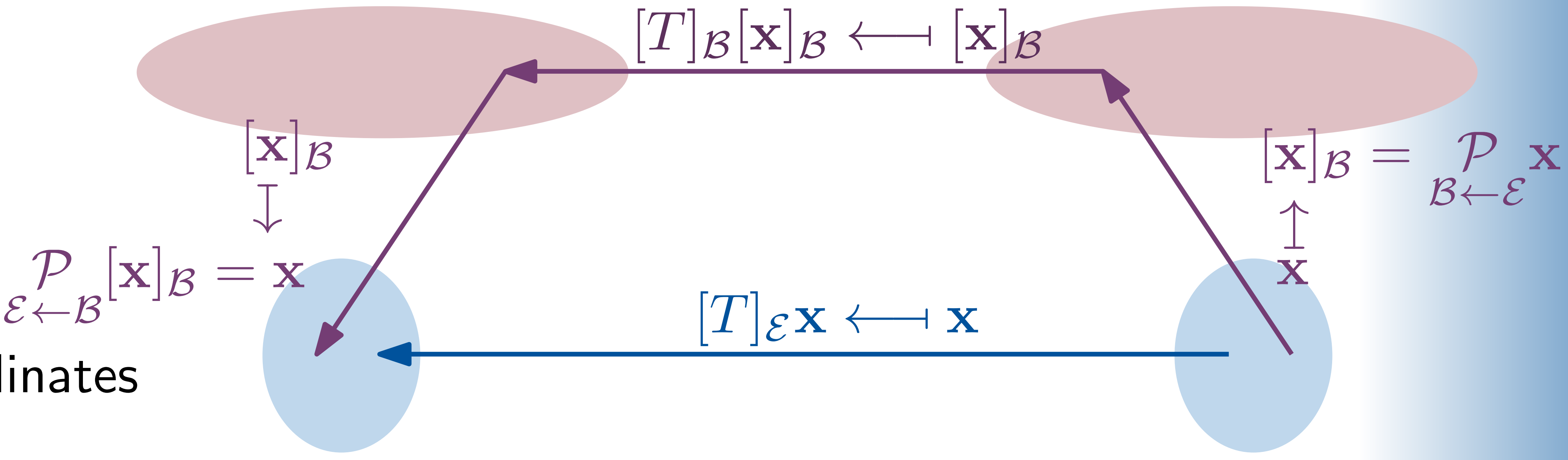
$\begin{bmatrix}1\\2\end{bmatrix}$ is on the line $y = 2x$, so it is unchanged by the reflection: $T\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}1\\2\end{bmatrix}$.

$\begin{bmatrix}2\\-1\end{bmatrix}$ is perpendicular to $y = 2x$, so its image is its negative: $T\left(\begin{bmatrix}2\\-1\end{bmatrix}\right) = \begin{bmatrix}-2\\1\end{bmatrix}$.

A different picture to understand $[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}$:

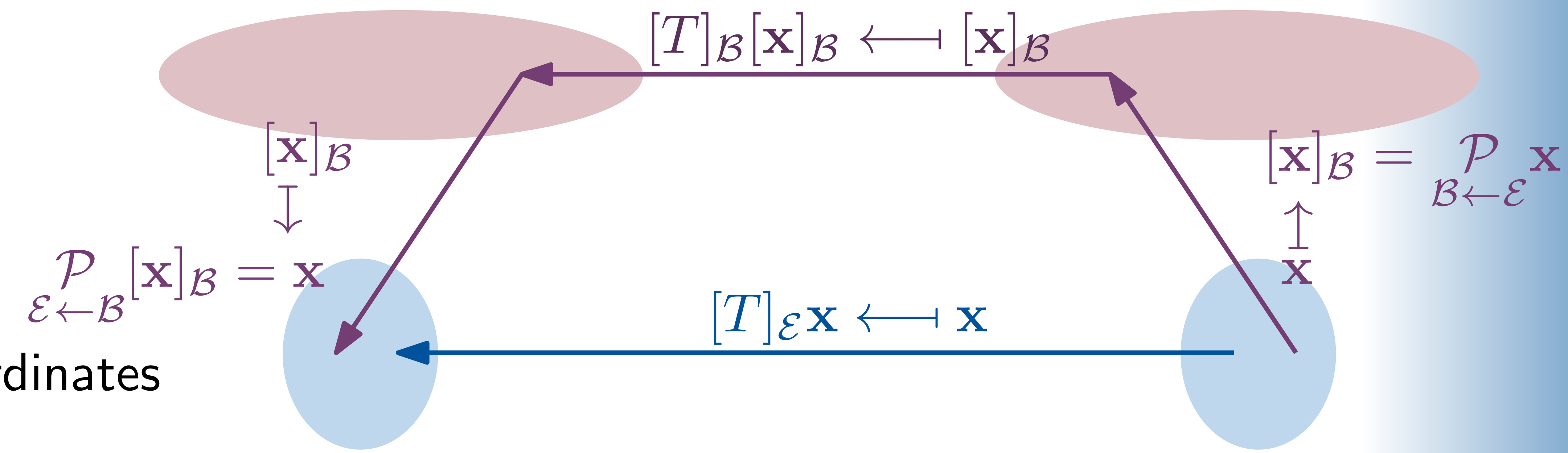
$\mathcal{B}$-coordinates

Standard coordinates



A different picture to understand $[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}$:

$\mathcal{B}$-coordinates



Standard coordinates

Because $\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_{\mathcal{B}}$ and $\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \mathcal{P}_{\mathcal{B}}^{-1}$:

$$[T]_{\mathcal{E}} = \mathcal{P}_{\mathcal{B}}[T]_{\mathcal{B}}\mathcal{P}_{\mathcal{B}}^{-1}.$$

Multiply both sides by $\mathcal{P}_{\mathcal{B}}^{-1}$ on the left and by $\mathcal{P}_{\mathcal{B}}$ on the right:

$$\mathcal{P}_{\mathcal{B}}^{-1}[T]_{\mathcal{E}}\mathcal{P}_{\mathcal{B}} = [T]_{\mathcal{B}}$$

These two equations are hard to remember ("where does the inverse go?"). Instead, remember $\boxed{[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}}}$ (which works for all vector spaces, not just $\mathbb{R}^n$).

Remember

$$[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}[T]_{\mathcal{B}}\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}.$$

This motivates the following definition:

Definition: Two square matrices A and D are *similar* if there is an invertible matrix P such that $A = PDP^{-1}$.

Similar matrices represent the same linear transformation in different bases.

Similar matrices have the same determinant and the same rank, because the signed volume scaling factor and the dimension of the image are coordinate-independent properties of the linear transformation. (Exercise: prove that $\det D = \det(PDP^{-1})$ using the multiplicative property of determinants.)

Why is change of basis important?

Example: If x, y are the prices of two stocks on a particular day, then their prices the next day are respectively $\frac{1}{2}y$ and $-x + \frac{3}{2}y$. How are the prices after many days related to the prices today?

Why is change of basis important?

Example: If x, y are the prices of two stocks on a particular day, then their prices the next day are respectively $\frac{1}{2}y$ and $-x + \frac{3}{2}y$. How are the prices after many days related to the prices today?

Answer: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the function representing the changes in stock prices from one day to the next, i.e. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}y \\ -x + \frac{3}{2}y \end{bmatrix}$. We are interested in T^k for large k. (You will NOT be required to do this step.)

T is a linear transformation; its standard matrix is $[T]_{\mathcal{E}} = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix}$. Calculating

$\begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix}^k$ by direct matrix multiplication will take a long time.

Answer: (continued) Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

$$T(\mathbf{b}_1) = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \frac{1}{2}\mathbf{b}_1, \quad T(\mathbf{b}_2) = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{b}_2,$$

so $[T]_{\mathcal{B}} = \begin{bmatrix} | & | \\ [T(\mathbf{b}_1)]_{\mathcal{B}} & [T(\mathbf{b}_2)]_{\mathcal{B}} \\ | & | \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$.

Answer: (continued) Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

$$T(\mathbf{b}_1) = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \tfrac{1}{2}\mathbf{b}_1, \quad T(\mathbf{b}_2) = \begin{bmatrix} 0 & \frac{1}{2} \\ -1 & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{b}_2,$$

so $[T]_{\mathcal{B}} = \begin{bmatrix} | & | \\ [T(\mathbf{b}_1)]_{\mathcal{B}} & [T(\mathbf{b}_2)]_{\mathcal{B}} \\ | & | \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$. Use $[T]_{\mathcal{E}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} [T]_{\mathcal{B}} \underset{\mathcal{B}\leftarrow\mathcal{E}}{\mathcal{P}} = \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} [T]_{\mathcal{B}} \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}$:

$$\begin{aligned}
[T]_{\mathcal{E}}^k &= \left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} [T]_{\mathcal{B}} \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1} \right)^k \\
&= \left(\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} [T]_{\mathcal{B}} \cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}} \right) \left(\cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}} [T]_{\mathcal{B}} \cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}} \right) \cancel{\cdots\cdots\cdots} \left(\cancel{\underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}} [T]_{\mathcal{B}} \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1} \right) \\
&= \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}} [T]_{\mathcal{B}}^k \underset{\mathcal{E}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1} \\
&= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}^k \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2}^k & 0 \\ 0 & 1^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -1 + \frac{1}{2}^{k-1} & 1 - \frac{1}{2}^k \\ -2 + \frac{1}{2}^{k-1} & 2 - \frac{1}{2}^k \end{bmatrix}.
\end{aligned}$$

So $[T]_{\mathcal{E}}^k = \begin{bmatrix} -1 + \frac{1}{2}^{k-1} & 1 - \frac{1}{2}^{k} \\ -2 + \frac{1}{2}^{k-1} & 2 - \frac{1}{2}^{k} \end{bmatrix}$. When k is very large, this is very close to $\begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}$.
So essentially the stock prices after many days is $-x + y$ and $-2x + 2y$, where x, y are the prices today. (In particular, the prices stabilise, which was not clear from $[T]_{\mathcal{E}}$.)

So $[T]_{\mathcal{E}}^k = \begin{bmatrix} -1+\frac{1}{2}^{k-1} & 1-\frac{1}{2}^k \\ -2+\frac{1}{2}^{k-1} & 2-\frac{1}{2}^k \end{bmatrix}$. When k is very large, this is very close to $\begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}$.
So essentially the stock prices after many days is $-x+y$ and $-2x+2y$, where x, y are the prices today. (In particular, the prices stabilise, which was not clear from $[T]_{\mathcal{E}}$.)

The important points in this example:

- We have a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ and we want to find T^k for large k.
- We find a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ where $T(\mathbf{b}_1) = \lambda_1 \mathbf{b}_1$ and $T(\mathbf{b}_2) = \lambda_2 \mathbf{b}_2$ for some scalars λ_1, λ_2. (In the example, $\lambda_1 = \frac{1}{2}, \lambda_2 = 1$.)
- Relative to the basis $\mathcal{B}$, the matrix for T is a diagonal matrix $[T]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$.
- It is easy to compute with $[T]_{\mathcal{B}}$, and we can then use change of coordinates to transfer the result to the standard matrix $[T]_{\mathcal{E}}$.

Next week (§5): does a "magic" basis like this always exist, and how to find it?
(Don't worry: you can do many of the computations in §5 without fully understanding change of coordinates.)