1. (a) $(Df)(x,y,z) = \begin{bmatrix} y & x+z & y \\ z+yz & xz & x+xy \end{bmatrix}$, so $(Df)(2,3,4) = \begin{bmatrix} 3 & 6 & 3 \\ 16 & 8 & 8 \end{bmatrix}$. Hence, for $\mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$ near $\mathbf{0}$ (i.e., all h_j's are near 0) we have

$$\begin{aligned} f(2+h_1, 3+h_2, 4+h_3) \approx f(2,3,4) + ((Df)(2,3,4))\mathbf{h} &= \begin{bmatrix} 18 \\ 32 \end{bmatrix} + \begin{bmatrix} 3 & 6 & 3 \\ 16 & 8 & 8 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \\ &= \begin{bmatrix} 18 \\ 32 \end{bmatrix} + \begin{bmatrix} 3h_1 + 6h_2 + 3h_3 \\ 16h_1 + 8h_2 + 8h_3 \end{bmatrix} \\ &= \begin{bmatrix} 18 + 3h_1 + 6h_2 + 3h_3 \\ 32 + 16h_1 + 8h_2 + 8h_3 \end{bmatrix}. \end{aligned}$$

This final expression is the approximation to the 2-vector $f(\mathbf{a}+\mathbf{h})$ for 3-vectors $\mathbf{h}$ near $\mathbf{0}$.

Likewise, for $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ near $\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$, we have

$$\begin{aligned} f(x,y,z) \approx f(2,3,4) + ((Df)(2,3,4))(\mathbf{x}-\mathbf{a}) &= \begin{bmatrix} 18 \\ 32 \end{bmatrix} + \begin{bmatrix} 3 & 6 & 3 \\ 16 & 8 & 8 \end{bmatrix} \begin{bmatrix} x-2 \\ y-3 \\ z-4 \end{bmatrix} \\ &= \begin{bmatrix} 18 \\ 32 \end{bmatrix} + \begin{bmatrix} 3(x-2) + 6(y-3) + 3(z-4) \\ 16(x-2) + 8(y-3) + 8(z-4) \end{bmatrix} \\ &= \begin{bmatrix} 18 \\ 32 \end{bmatrix} + \begin{bmatrix} 3x + 6y + 3z - 36 \\ 16x + 8y + 8z - 88 \end{bmatrix} \\ &= \begin{bmatrix} 3x + 6y + 3z - 18 \\ 16x + 8y + 8z - 56 \end{bmatrix}. \end{aligned}$$

This final expression is the approximation to the 2-vector $f(\mathbf{x})$ for $\mathbf{x}$ near $\mathbf{a}$.

(b) $(Df)(x,y) = \begin{bmatrix} e^x(x-y)^2 + 2e^x(x-y) & -2e^x(x-y) \\ 3y^2 & 6xy \end{bmatrix}$, so $(Df)(0,1) = \begin{bmatrix} -1 & 2 \\ 3 & 0 \end{bmatrix}$. Hence, for $\mathbf{h} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$ near $\mathbf{0}$ (i.e., all h_j's are near 0) we have

$$\begin{aligned} f(h_1, 1+h_2) \approx f(0,1) + ((Df)(0,1))\mathbf{h} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -h_1 + 2h_2 \\ 3h_1 \end{bmatrix} \\ &= \begin{bmatrix} 1 - h_1 + 2h_2 \\ 3h_1 \end{bmatrix}. \end{aligned}$$

This final expression is the approximation to the 2-vector $f(\mathbf{a}+\mathbf{h})$ for 2-vectors $\mathbf{h}$ near $\mathbf{0}$.

Likewise, for $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ near $\mathbf{a} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have

$$\begin{aligned} f(x,y) \approx f(0,1) + ((Df)(0,1))(\mathbf{x}-\mathbf{a}) &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} x \\ y-1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -x + 2(y-1) \\ 3x \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -x + 2y - 2 \\ 3x \end{bmatrix} \\ &= \begin{bmatrix} -1 - x + 2y \\ 3x \end{bmatrix}. \end{aligned}$$

This final expression is the approximation to the 2-vector $G(\mathbf{x})$ for $\mathbf{x}$ near $\mathbf{a}$.

◇

2. (a) We have
$$f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -2 & 3 & 1 \\ -4 & 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 5 \\ -7 \end{bmatrix} = \begin{bmatrix} -2x+3y+z+5 \\ -4x+2z-7 \end{bmatrix},$$
so the component functions are $f_1(x,y,z) = -2x+3y+z+5$ and $f_2(x,y,z) = -4x+2z-7$. The partial derivatives of f_1 and f_2 with respect to any of x, y, z are all *constant* (i.e., independent of the point at which we evaluate them). More specifically,
$$(Df)\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -2 & 3 & 1 \\ -4 & 0 & 2 \end{bmatrix}$$
is *independent* of x, y, z. Hence, for any $\mathbf{c} \in \mathbf{R}^3$ at all, we have
$$(Df)(\mathbf{c}) = \begin{bmatrix} -2 & 3 & 1 \\ -4 & 0 & 2 \end{bmatrix},$$
and this is A by inspection.

(b) We explicitly compute
$$f\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + b_1 \\ a_{21}x + a_{22}y + b_2 \end{bmatrix},$$
so the component functions are $f_1(x,y) = a_{11}x + a_{12}y + b_1$ and $f_2(x,y) = a_{21}x + a_{22}y + b_2$. As in (a), the partial derivatives of f_1 and f_2 are constant functions:
$$(Df)\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
regardless of the point $\begin{bmatrix} x \\ y \end{bmatrix}$ at which we work. Hence, for any $\mathbf{c} \in \mathbf{R}^2$ we have
$$(Df)(\mathbf{c}) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$
and this is A by inspection.

(c) Writing a_{ij} for the ij-entry of A, we explicitly compute
$$\begin{aligned} f\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{mn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + b_m \end{bmatrix} \end{aligned}$$
so the ith component function of f is
$$f_i(x_1, \dots, x_n) = a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n + b_i.$$
This has x_j-coefficient a_{ij}, so $\partial f_i/\partial x_j = a_{ij}$ is a constant, independent of at what n-vector $\mathbf{c}$ one evaluates this partial derivative. Hence, assembling these constants into a matrix, we see that for any $\mathbf{c} \in \mathbf{R}^n$ at all,
$$(Df)(\mathbf{c}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{mn} \end{bmatrix} = A.$$

◇

3. (a) By Proposition 13.4.5 , the matrix A of this linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ is given by

$$A = \begin{bmatrix} | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) \\ | & | \end{bmatrix}.$$

The next effect of T is

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \xrightarrow{\text{rotated clockwise by } 45^\circ} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \xrightarrow{\text{horizontally doubled}} \begin{bmatrix} \sqrt{2} \\ -1/\sqrt{2} \end{bmatrix};$$

i.e., the result $T(\mathbf{e}_1)$ of our operation on $\mathbf{e}_1$ is

$$T(\mathbf{e}_1) = \begin{bmatrix} \sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

Likewise,

$$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \xrightarrow{\text{rotated clockwise by } 45^\circ} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \xrightarrow{\text{horizontally doubled}} \begin{bmatrix} \sqrt{2} \\ 1/\sqrt{2} \end{bmatrix},$$

so the result $T(\mathbf{e}_2)$ of our operation on $\mathbf{e}_2$ is

$$T(\mathbf{e}_2) = \begin{bmatrix} \sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Therefore, the matrix A we're looking for is

$$A = \begin{bmatrix} | & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

(b) Using Proposition 13.4.5 , $R = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and $M = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$, and multiplying in the order of composition (with first step on the right!) gives that T has matrix

$$MR = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix},$$

which indeed agrees with matrix obtained in (a).

(c) We are stretching along only one direction, and whether doing that before or after the rotation has a huge effect because the direction along which the stretching occurs will be different (either alone the x-direction or along the line $y = x$). In terms of matrices, if we multiply in the other order we get

$$RM = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ -\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

This is not the same as MR computed in (b).

◇

4. (a) The effect of T on $\mathbf{e}_3$ is to do nothing, whereas $T(\mathbf{e}_1) = \mathbf{e}_2$ and $T(\mathbf{e}_2) = -\mathbf{e}_1$, so

$$A = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Likewise, the effect of T' on $\mathbf{e}_1$ is to do nothing, whereas $T'(\mathbf{e}_2) = \mathbf{e}_3$ and $T'(\mathbf{e}_3) = -\mathbf{e}_2$, so

$$A' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}.$$

(b) There are many entries equal to 0, so we readily compute the matrix product to be

$$A'A = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}.$$

By thinking in terms of rotations, alternatively we get the columns as follows:

$$T'(T(\mathbf{e}_1) = T'(\mathbf{e}_2) = \mathbf{e}_3,\;\; T'(T(\mathbf{e}_2)) = T'(-\mathbf{e}_1) = -\mathbf{e}_1,\;\; T'(T(\mathbf{e}_3)) = T'(\mathbf{e}_3) = -\mathbf{e}_2.$$

(c) There are many entries equal to 0, so we readily compute the matrix product to be

$$AA' = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

By thinking in terms of rotations, alternatively we get the columns as follows:

$$T(T'(\mathbf{e}_1) = T(\mathbf{e}_1) = \mathbf{e}_2,\;\; T(T'(\mathbf{e}_2)) = T(\mathbf{e}_3) = \mathbf{e}_3,\;\; T(T'(\mathbf{e}_3)) = T(-\mathbf{e}_2) = \mathbf{e}_1.$$

◇

5. (a) Since the jth column is $p(\mathbf{e}_j)$ for A, and similarly with B for i, we calculate the matrices as

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

(b) The product AB corresponds to first including into the first two components and then projecting on the last two. Consequently $p \circ i$ is the zero function. On the other hand, one calculates directly that the 2×2 matrix

$$AB = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

has all entries equal to a sum of 0's, so it is the zero matrix of size 2×2, which indeed computes the zero function $\mathbf{R}^2 \to \mathbf{R}^2$.

(c) The product BA projects onto the last two components and then includes in the first two. This kills $\mathbf{e}_1$ and $\mathbf{e}_2$, and carries $\mathbf{e}_3$ to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ to $\mathbf{e}_1$ and carries $\mathbf{e}_4$ to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ to $\mathbf{e}_2$. Consequently, by chasing the effect on each $\mathbf{e}_j$, this has matrix

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This agrees with the matrix product

$$BA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

since for the entries in this product we get 0's everywhere apart from contributions of $1 \cdot 1 = 1$ as appear in the 4×4 output in the $(1, 3)$ and $(2, 4)$ positions.

◇

6. (a) Since $R\mathbf{v} = \begin{bmatrix} (x+y)/\sqrt{2} \\ (-x+y)/\sqrt{2} \end{bmatrix}$, to say $R\mathbf{v}$ lies in $H_{\pm}$ is exactly to say that

$$\pm 1 = \left(\frac{x+y}{\sqrt{2}}\right)^2 - \left(\frac{x-y}{\sqrt{2}}\right)^2.$$

By high school algebra, the right side of this equation is always equal to $(x+y)^2/2 - (x-y)^2/2 = xy - (-xy) = 2xy$, so it is the same to say that $xy = \pm 1/2$, as desired.

(b) From the interpretation of R as a $45°$ clockwise rotation, it follows from (a) that $H_{\pm}$ is the $45°$ clockwise rotation of the graph of $\pm 1/(2x)$. This rotation carries asymptotes to asymptotes, and this rotation carries the coordinate axes to the lines $y = \pm x$, so these latter lines are the asymptotes. The curves $H_{\pm}$ and their asymptotes are shown in Figure 1 .

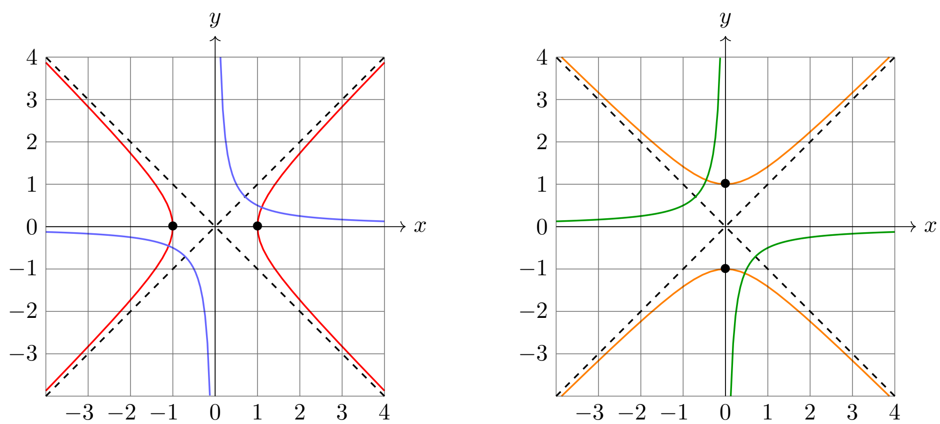


Figure 1: The curves H_+ (red) and H_- (orange), with dotted asymptotes $y = \pm x$ and black dots where each crosses the coordinate axes. The graphs of $1/(2x)$ (blue) and $-1/(2x)$ (green) with coordinate axes as their asymptotes are overlaid on the curves H_+ and H_- respectively, for comparison purposes.

(c) The reasoning as in Exercise 13.2 shows that the curve $\frac{x^2}{4} - \frac{y^2}{9} = 1$ is $T_{2,3}(H_+)$, and the curve $\frac{x^2}{4} - 4y^2 = -1$ is $T_{2,1/2}(H_-)$. The corresponding asymptotes are obtained by respectively applying $T_{2,3}$ and $T_{2,1/2}$ to the asymptotes $y = \pm x$ of H_+ and H_-. By thinking about the scaling effects in the horizontal and vertical directions, these respective pairs of asymptotes are the "steeper" lines $y = \pm(3/2)x$ (tripling vertically but only doubling horizontally, so multiplying the original slopes ± 1 by $3/2$) and the "flatter" lines $y = \pm x/4$ (halving vertically and doubling horizontally, so dividing the original slopes ± 1 by 4). The resulting curves along with their asymptotes and intercepts with the coordinate axes are shown in Figure 2 .

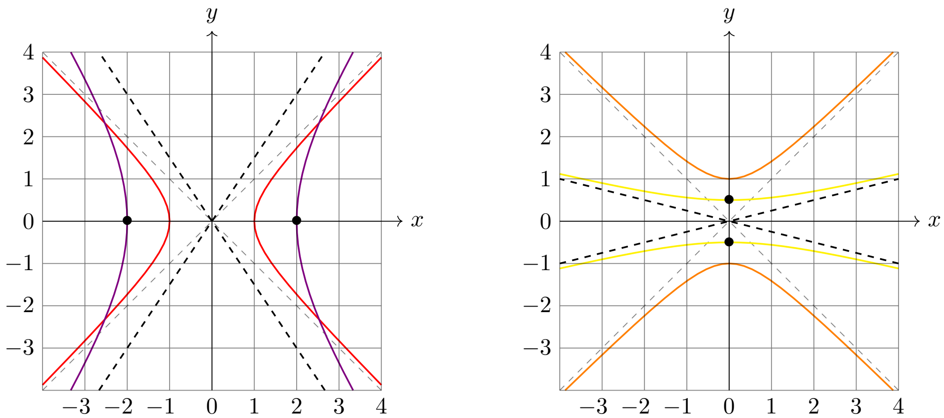
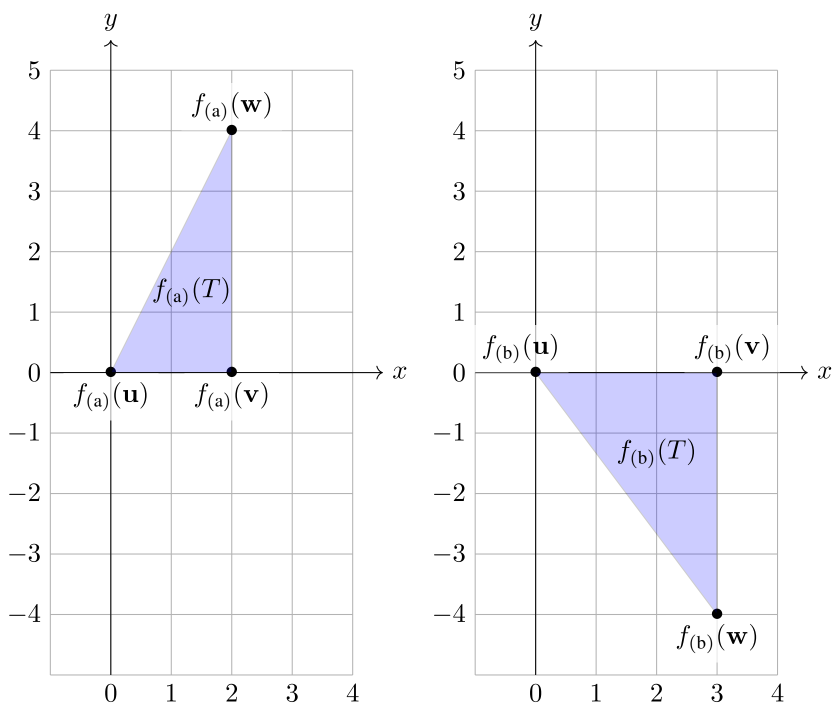
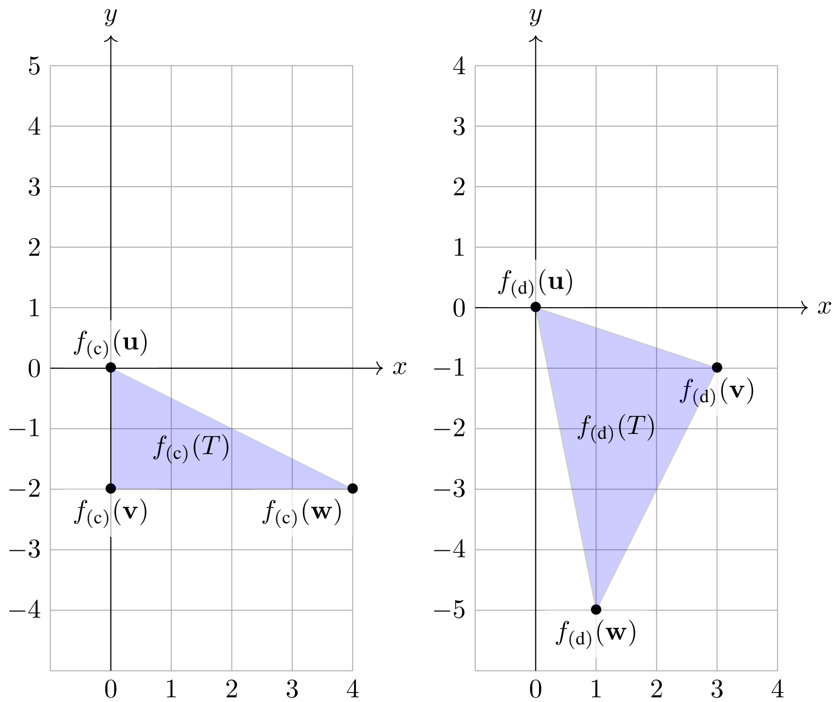


Figure 2: The curves $x^2/4 - y^2/9 = 1$ (purple) and $x^2/4 - 4y^2 = -1$ (yellow) and their respective dotted asymptotes $y = \pm(3/2)x$ and $y = \pm x/4$, and black dots where each graph crosses the coordinate axes. The original curves H_+ (red) and H_- (orange) and their asymptotes (in a lighter tone) are included solely for comparison purposes (not needed).

◇

7. Since linear transformations carry line segments to line segments (determined by the effect on the endpoints), due to a line segment being given by convex combinations of its endpoints with parameter $0 \leq t \leq 1$ (and a linear transformation preserves the formation of convex combinations, as a special case of linear combinations), we just have to compute the effect on the vertices: the edges will get carried onto the corresponding edges, and the interior goes to the interior (the interior of a triangle is the collection of convex combinations $r\mathbf{u} + s\mathbf{v} + t\mathbf{w}$ with $0 \leq r, s, t \leq 1$ and $r + s + t = 1$). If we let $f_{(\mathrm{a})}$, $f_{(\mathrm{c})}$, $f_{(\mathrm{c})}$ and $f_{(\mathrm{d})}$ denote the linear transformations defined by the matrices in (a), (b), (c), (d), then the images of T are the triangles shown in the following figures.





◇

8. We take the approach of multiplying everything out. (There are ways to proceed that involve less brute force.)

   (a) We directly multiply matrices:

$$AM = \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} = \begin{bmatrix} m_{11} + am_{21} & m_{12} + am_{22} & m_{13} + am_{23} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix},$$

   which is exactly as desired: we have added a times the second row of M to the first row of M.
   For MA, the effect is to add a times the first column of M to the second column of M. This can be seen by multiplying the matrices:

$$MA = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} m_{11} & am_{11} + m_{12} & m_{13} \\ m_{21} & am_{21} + m_{22} & m_{23} \\ m_{31} & am_{31} + m_{32} & m_{33} \end{bmatrix}.$$

   (b) We direct multiply matrices:

$$BM = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} = \begin{bmatrix} m_{21} & m_{22} & m_{23} \\ m_{11} & m_{12} & m_{13} \\ m_{31} & m_{32} & m_{33} \end{bmatrix},$$

   which is exactly as desired: we have swapped the first and second rows of M.
   For MB, the effect is to swap the first and second columns of M. This can be seen by multiplying the matrices:

$$MB = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} m_{12} & m_{11} & m_{13} \\ m_{22} & m_{21} & m_{23} \\ m_{32} & m_{31} & m_{33}. \end{bmatrix}$$

◇

9. (a) We first compute

$$A^2 = \begin{bmatrix} 1 & 2 \\ -5 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} 1 - 10 & 2 + 4 \\ -5 - 10 & -10 + 4 \end{bmatrix} = \begin{bmatrix} -9 & 6 \\ -15 & -6 \end{bmatrix}$$

and

$$B^2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 2-1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

Thus,

$$f(A) = 2A^2 + 3A - \mathrm{I}_2 = \begin{bmatrix} -18 & 12 \\ -30 & -12 \end{bmatrix} + \begin{bmatrix} 3 & 6 \\ -15 & 6 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -16 & 18 \\ -45 & -7 \end{bmatrix}$$

and

$$f(B) = 2B^2 + 3B - \mathrm{I}_3 = \begin{bmatrix} 8 & 2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 18 \end{bmatrix} + \begin{bmatrix} 6 & 3 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 9 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 13 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 26 \end{bmatrix}.$$

(b) Since we already computed A^2 for part (a), we can just add the three relevant terms together:

$$g(A) = A^2 - 3A + 12\mathrm{I}_2 = \begin{bmatrix} -9 & 6 \\ -15 & -6 \end{bmatrix} - \begin{bmatrix} 3 & 6 \\ -15 & 6 \end{bmatrix} + \begin{bmatrix} 12 & 0 \\ 0 & 12 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

(c) On the left side, we have

$$\begin{aligned} h(C) &= C^2 + 2C + \mathrm{I}_2 \\ &= \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 6 & -2 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 9 & -4 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 6 & -2 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 16 & -6 \\ 0 & 4 \end{bmatrix}. \end{aligned}$$

On the right side, we have

$$(C + \mathrm{I}_2)^2 = \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 4 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 16 & -6 \\ 0 & 4 \end{bmatrix},$$

so the two sides are indeed equal.

◇

10. (a) Based on the hint, we can take M for which the first two columns are given by $B - C$ and the third column is all zeros. We should expect this to work, because the first two columns of AM will then be given by $A(B - C)$, while the third column is given by $A \cdot \mathbf{0}$. Indeed, if $M = \begin{bmatrix} -1 & 2 & 0 \\ 2 & -4 & 0 \\ 1 & -2 & 0 \end{bmatrix}$, then

$$AM = \begin{bmatrix} -2+6-4 & 4-12+8 & 0+0+0 \\ -4-2+6 & 8+4-12 & 0+0+0 \\ -5+4+1 & 10-8-2 & 0+0+0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

as we wanted.

(b) We compute

$$MA = \begin{bmatrix} -2+8 & -3-2 & 4+12 \\ 4-16 & 6+4 & -8-24 \\ 2-8 & 3+2 & -4-12 \end{bmatrix} = \begin{bmatrix} 6 & -5 & 16 \\ -12 & 10 & -32 \\ -6 & 5 & -16 \end{bmatrix},$$

which is not the zero matrix, so it isn't equal to AM.

◇

11. We note that T is a pointwise sum of two linear transformations: if $R_{\pi/4}$ is the rotation by $45°$ and S is the scaling-by-2 transformation, then $T(\mathbf{x}) = R_{\pi/4}(\mathbf{x}) + S(\mathbf{x})$. Thus, the matrix for T is given by the matrix sum of the respective matrices

for $R_{\pi/4}$ and S. Now, the matrix $A_{\pi/4}$ for $R_{\pi/4}$ was computed in Example 14.4.1 , and equals $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. The matrix of S is $2\mathrm{I}_2$ (twice the 2×2 identity matrix). Therefore, the matrix A for T is given by

$$A = A_{\pi/4} + 2\mathrm{I}_2 = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 + (1/\sqrt{2}) & -1/\sqrt{2} \\ 1/\sqrt{2} & 2 + (1/\sqrt{2}) \end{bmatrix}.$$

◇

12. (a) The upper triangular condition says that the ith column of A has vanishing entries beyond the ith for each i. which is to say that it lies in the span of $\mathbf{e}'_1, \ldots, \mathbf{e}'_i$. But the ith column is $L(\mathbf{e}_i)$, so A is upper triangular precisely when $L(\mathbf{e}_i) \in V'_i$ for every i. If this happens then for every $j \leq i$ we have $L(\mathbf{e}_j) \in V'_j$, but V'_j is contained in V'_i since $j \leq i$, so

$$L(\mathbf{e}_1), \ldots, L(\mathbf{e}_i) \in V'_i$$

for every i. But the span of those i vectors then lies inside V'_i, and such spans are exactly the vectors

$$c_1 L(\mathbf{e}_1) + \cdots + c_i L(\mathbf{e}_i) = L(c_1\mathbf{e}_1 + \cdots + c_i\mathbf{e}_i),$$

which are exactly the vectors $L(\mathbf{v})$ for $\mathbf{v} \in V_i$. In other words, when A is upper triangular we have $L(V_i)$ is contained in V'_i for every i.

This goes in reverse: if $L(V_i)$ is contained in V'_i for every i then in particular $L(\mathbf{e}_i) \in V'_i$ for every i. This latter condition has already been seen to encode exactly that A is upper triangular.

(b) Let $L_1 : \mathbf{R}^n \to \mathbf{R}^m$ be the linear transformation corresponding to U_1, and $L_2 : \mathbf{R}^p \to \mathbf{R}^n$ be the linear transformation corresponding to U_2. By part (a), the upper triangular condition says $L_1(V_i)$ is contained in V'_i for every i and $L_2(V''_i)$ is contained in V_i for every i (where V''_i denotes the span of the first i standard basis vectors of $\mathbf{R}^p$, with $V''_i = \mathbf{R}^p$ when $i > p$). Since U_1U_2 corresponds to the composition $L_1 \circ L_2$, by part (a) to check that the product U_1U_2 is upper triangular we just need to check that $L_1 \circ L_2$ carries V''_i into V'_i itself for each i.

For a choice of i, if $\mathbf{v} \in V''_i$ then $(L_1 \circ L_2)(\mathbf{v}) = L_1(L_2(\mathbf{v}))$ with $L_2(\mathbf{v}) \in V_i$ as noted above. But then $L_1(L_2(\mathbf{v}))$ belongs to V'_i since L_1 carries *everything* in V_i into V'_i, and so, as we mentioned above, we conclude that U_1U_2 is upper triangular.

◇

