1. (a) Setting $g(x,y,z) = x^2 + y^2 - 4z^2$, the constraint is given by the level set $g(x,y,z) = 1$. Thus, using Theorem 12.2.1, the extrema of f subject to $g(x,y,z) = 1$ occur only when $\nabla g = \mathbf{0}$ or $\nabla f = \lambda(\nabla g)$. Now, $(\nabla g)(x,y,z) = \begin{bmatrix} 2x \\ 2y \\ -8z \end{bmatrix}$. This is zero only at the origin, which does not satisfy $g(x,y,z) = 1$. So we can focus on points where the second case happens.

We have $\nabla f = \begin{bmatrix} z \\ z \\ x+y \end{bmatrix}$. Thus, we have the vector equation $\nabla f = \lambda \nabla g$ for some unknown scalar λ; this encodes three scalar equations which together with the constraint equation gives a combined system of 4 equations:

$$z = 2\lambda x, \quad z = 2\lambda y, \quad x + y = -8\lambda z, \quad x^2 + y^2 - 4z^2 = 1.$$

To carry out the method of "solving for λ" with each of the first three equations and setting the resulting expressions equal to each other, as always we have to be careful about division by zero.

If we don't worry about division by zero, the expressions we get are:

$$\lambda = \frac{z}{2x}, \quad \lambda = \frac{z}{2y}, \quad \lambda = \frac{x+y}{-8z},$$

so all three right sides are equal to each other provided we avoid situations in which any of their denominators vanish. So let's first dispose of cases of such vanishing denominators: this means that at least one of x, y, or z vanishes. We treat these one at a time.

If $x = 0$ then the first equation in our combined system forces $z = 0$, and then the third equation in the combined system forces $x + y = 0$ and hence $y = 0$ (since we're assuming $x = 0$), but $(0,0,0)$ violates the constraint equation $g(x,y,z) = 1$. Hence, the possibility $x = 0$ cannot occur.

Likewise, if $y = 0$ then the second equation in our combined system forces $z = 0$, so the third equation in the combined system forces $x + y = 0$ once again, so now $x = 0$ (as we are assuming $y = 0$), but we already ruled out the possibility $x = 0$. So now the possibility $y = 0$ has been ruled out.

Finally, if $z = 0$ then the third equation in the combined system forces $x + y = 0$, so $y = -x$, and the first and second equations in the combined system becomes $0 = 2\lambda x$ and $0 = -2\lambda x$. But we already ruled out the possibility $x = 0$, so necessarily $\lambda = 0$. The constraint equation gives us more information: $1 = x^2 + y^2 - 4(0)^2 = 2x^2$ (since $y = -x$), so $x = \pm 1/\sqrt{2}$ and $y = -x = \mp 1/\sqrt{2}$. So in this way we have arrived at two special points that will require separate consideration: $(1/\sqrt{2}, -1/\sqrt{2}, 0)$ and $(-1/\sqrt{2}, 1/\sqrt{2}, 0)$ (with $\lambda = 0$, but that is not going to matter).

Putting that into the fridge for later, now we can look into equating the preceding various fraction expressions we obtained for λ when all the denominators are nonzero. Equating them all yields the triple equality

$$\frac{z}{2x} = \frac{z}{2y} = \frac{x+y}{-8z}.$$

Since the denominators are all nonzero, so $z \neq 0$, the first equality tells us that $2x = 2y$, so $x = y$. Then the equating the second and third expressions (or equivalently the first and third) gives upon cross-multiplying

$$(-8z)z = 2y(x+y) = 2x(x+x) = 4x^2,$$

so $-2z^2 = x^2$. But $x, z \neq 0$, so the left side is negative and the right side is positive, so this is impossible.

The upshot is that the only candidate points for local extrema, let along global extrema, on the constraint surface $g = 1$ are $(1/\sqrt{2}, -1/\sqrt{2}, 0)$ and $(-1/\sqrt{2}, 1/\sqrt{2}, 0)$.

(b) Evaluating at the two candidates from (a), $f(1/\sqrt{2}, -1/\sqrt{2}, 0) = 0$ and $f(-1/\sqrt{2}, 1/\sqrt{2}, 0) = 0$. On the other hand, we can find points on the hyperboloid such as $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ where f is positive (since $f(1,2,1) = 3$). Similarly, we can find points on the hyperboloid such as $\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ where f is negative (since $f(-2,1,1) = -1$). Thus, the value 0 is not a global maximum nor minimum. But these were the only possibilities, so f has no global extrema on the hyperboloid. $\diamondsuit$

2. (a) We first compute the gradient of f:

$$(\nabla f)(x,y,z) = \begin{bmatrix} 4x+6y \\ 6x+2y \\ 2 \end{bmatrix}.$$

Because the third coordinate doesn't vanish, f has no critical points in $\mathbf{R}^3$. Hence, it has none in the interior of the R.

(b) Our constraint is $g(x,y,z) = 1$, where $g(x,y,z) = x+y+z$. Thus, $(\nabla g)(x,y,z) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

This doesn't vanish anywhere, so the method of Lagrange multipliers says that our candidate extrema on the given region all satisfy

$$\begin{bmatrix} 4x+6y \\ 6x+2y \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

From equality in the third coordinate, we see that $\lambda = 2$. We can then solve for x, y based on equality in the first two coordinates:

$$\begin{cases} 4x+6y = 2 \\ 6x+2y = 2 \end{cases}.$$

Multiply the first equation by 2 and the second by 6 to obtain $8x+12y = 4$ and $36x+12y = 12$. Subtracting these, we obtain $28x = 8$, so $x = 2/7$. Plugging this back in to the second equation, we have $y = 1/7$. For our point to satisfy the constraint $x+y+z = 1$, we must therefore have $z = 4/7$. So we have a single candidate extremum on this region: $(2/7, 1/7, 4/7)$.

(c) We have

$$f(t, 1-t, 0) = 2t^2 + 6t(1-t) + (1-t)^2 = 2t^2 + 6t - 6t^2 + 1 - 2t + t^2 = -3t^2 + 4t + 1$$

with $0 \leq t \leq 1$. This function has a graph that is an upside-down parabola with vertex at $t = 2/3$ where it attains its maximum $-3(2/3)^2 + 4(2/3) + 1 = 7/3$ (as could also be deduced using calculus, but that is overkill for parabolas). So on this edge the maximal value $7/3$ is attained at $(2/3, 1/3, 0)$.

(d) Since there are no critical points in the interior of the R, we just have to compare the two values $f(2/3, 1/3, 0) = 7/3$ (slightly bigger than 2) and

$$\begin{aligned} f\left(\frac{2}{7}, \frac{1}{7}, \frac{4}{7}\right) &= 2\left(\frac{2}{7}\right)^2 + 6\left(\frac{2}{7}\right)\left(\frac{1}{7}\right) + \left(\frac{1}{7}\right)^2 + 2 \cdot \frac{4}{7} \\ &= \frac{2 \cdot 2^2 + 6 \cdot 2 \cdot 1 + 1^2 + 8 \cdot 7}{7^2} \\ &= \frac{77}{49}, \end{aligned}$$

which happens to equal 11/7, but we don't expect you to reduce the fraction; the key point is to recognize that this value is less than 2. So the maximal value for f on R is $7/3$, occurring at $(2/3, 1/3, 0)$. $\diamond$

3. (a) Here is a picture of the region R.

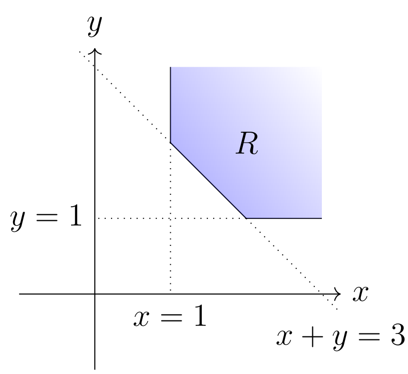

FIGURE 2. The region simultaneously "on or above" each of three lines

The function $4x + 3y$ becomes as big as we want by taking x or y as big as we want (since the coefficients 4 and 3 are both positive), so there is no maximum value. Since this function has partial derivatives equal to nonzero constants (4 for the x-partial, 3 for the y-partial), it has no local extremum on the interior. Likewise, if we view it as a function on any boundary segment it is a linear function on a parameter for that boundary segment and hence its minimal value on any boundary segment is at an endpoint of that segment (just like for a linear function $at + b$ on any interval of t-values). Thus, the minimal value must be attained at one of the corners, and the picture shows that there are two corners: $(2, 1)$ and $(1, 2)$. Evaluating $4x + 3y$ at the corner points yields $8 + 3 = 11$ and $4 + 6 = 10$ respectively. So the minimum is attained at $(1, 2)$.

(b) Here is a picture of the parallelogram formed with vertices $\mathbf{0}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\mathbf{v} + \mathbf{w}$.

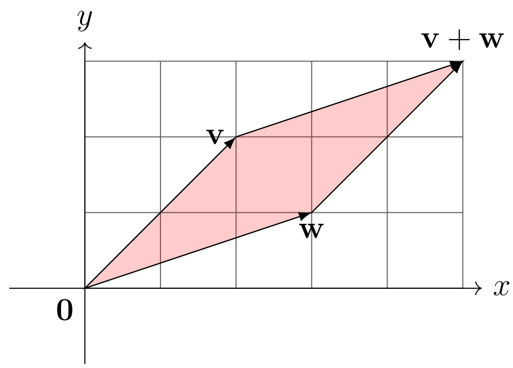

FIGURE 3. The parallelogram arising from two specific vectors in $\mathbf{R}^2$

By the same reasoning as in part (a), one of the corners must be the maximizing point. So we evaluate $4x+3y$ at the corners $(0,0)$, $(2,2)$, $(3,1)$, $(5,3)$. The respective values are 0, 14, 15, 29, so the maximum is attained at $(5,3)$ with value 29.

(c) Once again, one of the corners must be the minimizing point, so we evaluate $-x+2y$ at the corners $(0,0)$, $(2,2)$, $(3,1)$, $(5,3)$. The respective values are 0, 2, -1, 1, so the minimum is attained at $(3,1)$ with value -1. $\Diamond$

4. We want to minimize $f(x,y) = x^2+y^2$ subject to the constraint $g(x,y) = 3x+2y = 6$. The gradients are

$$\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix}, \qquad \nabla g = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Constrained extrema of f can only occur where $\nabla g = \mathbf{0}$ or $\nabla f = \lambda \nabla g$ for some λ. Since $\nabla g \neq \mathbf{0}$, we only need to study $\nabla f = \lambda \nabla g$, which amounts to the scalar equations

$$2x = 3\lambda, \qquad 2y = 2\lambda.$$

Solving for λ in each gives $(2/3)x = \lambda = y$. The constraint then gives $6 = g(x,y) = g(x, 2x/3) = 3x + 4x/3 = 13x/3$, so $x = 18/13$. Since $y = (2/3)x = 12/13$, the desired point is $(18/13, 12/13)$. (Its distance to the origin, which you weren't asked to compute, is the square root of $f(18/13, 12/13) = 468/169 = 36/13$, which is $6/\sqrt{13}$.) $\Diamond$

5. Suppose the rectangular box has side lengths $x, y, z > 0$, so we want to minimize the surface area $f(x,y,z) = 2(xy+xz+yz)$ subject to the constraint that the volume $g(x,y,z) = xyz$ is equal to a given V. The gradients are

$$\nabla f = \begin{bmatrix} 2y+2z \\ 2x+2z \\ 2x+2y \end{bmatrix}, \qquad \nabla g = \begin{bmatrix} yz \\ xz \\ xy \end{bmatrix}.$$

Constrained extrema of f can only occur where $\nabla g = \mathbf{0}$ or $\nabla f = \lambda \nabla g$ for some λ. The gradient of g can't vanish since its entries are all positive (as $x, y, z > 0$). So we only need to study $\nabla f = \lambda \nabla g$, which amounts to the triple of scalar equations $2(y+z) = \lambda yz$, $2(x+z) = \lambda xz$, $2(x+y) = \lambda xy$. Since $x, y, z > 0$, we can solve each for λ via division. Setting those expressions all equal to each other, and cancelling their common factor of 2, gives

$$\frac{y+z}{yz} = \frac{x+z}{xz} = \frac{x+y}{xy},$$

or in other words

$$\frac{1}{z}+\frac{1}{y}=\frac{1}{z}+\frac{1}{x}=\frac{1}{y}+\frac{1}{x}.$$

The first equality gives $1/y = 1/x$ and the second gives $1/z = 1/y$, so all three reciprocals are equal and hence $x = y = z$. So the box must be a cube. ◊

6. (a) The effect of $T_{2,1/3}$ is to scale by a factor of 2 in the x-direction and scale by a factor of 1/3 in the y-direction. Applying this to $\begin{bmatrix}-1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\2\end{bmatrix}$ yields $\begin{bmatrix}-2\\1/3\end{bmatrix}$ and $\begin{bmatrix}2\\2/3\end{bmatrix}$ respectively. When these four vectors are drawn, the relation between $\begin{bmatrix}-1\\1\end{bmatrix}$ and $\begin{bmatrix}-2\\1/3\end{bmatrix}$ and between $\begin{bmatrix}1\\2\end{bmatrix}$ and $\begin{bmatrix}2\\2/3\end{bmatrix}$ fits this description. In Figure 4 we show the effect of $T_{2,1/3}$ on a 2×2 grid centered on the origin.

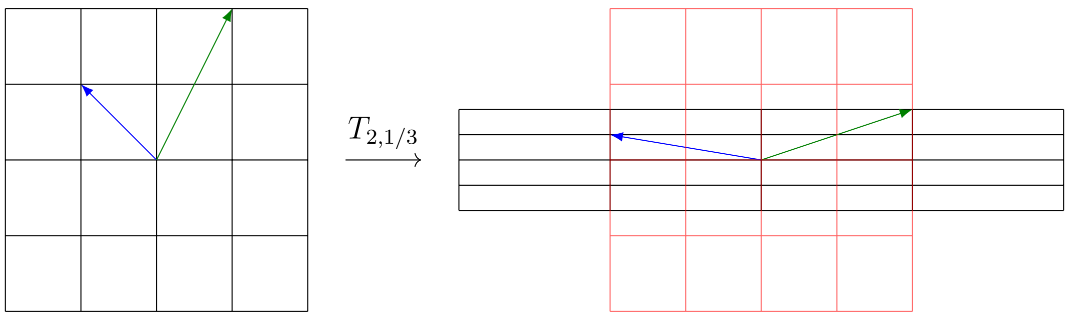

FIGURE 4. The effect of $T_{2,1/3}$ on a 2×2 grid (the pink grid on the right is just the original 2×2 grid overlaid on top of the new grid, for comparison purposes).

In general, the effect of $T_{a,b}$ is scaling by a horizontally and by b vertically.

(b) We calculate

$$T_{1/a,1/b}\left(T_{a,b}\left(\begin{bmatrix}x\\y\end{bmatrix}\right)\right) = T_{1/a,1/b}\left(\begin{bmatrix}ax\\by\end{bmatrix}\right) = \begin{bmatrix}(1/a)(ax)\\(1/b)(by)\end{bmatrix} = \begin{bmatrix}x\\y\end{bmatrix},$$

$$T_{a,b}\left(T_{1/a,1/b}\left(\begin{bmatrix}x\\y\end{bmatrix}\right)\right) = T_{a,b}\left(\begin{bmatrix}x/a\\y/b\end{bmatrix}\right) = \begin{bmatrix}a(x/a)\\b(y/b)\end{bmatrix} = \begin{bmatrix}x\\y\end{bmatrix}.$$

That these undo each other is seen also via (a) since the scaling effects by a and $1/a$ in the x-direction cancel each other out, and likewise for scaling by b and $1/b$ in the y-direction.

(c) If $v_1^2 + v_2^2 = 1$ then (av_1, bv_2) lies in $E_{a,b}$ since $(av_1)^2/a^2 + (bv_2)^2/b^2 = v_1^2 + v_2^2 = 1$. Likewise, if (av_1, bv_2) lies on the curve $E_{a,b}$ then this means $(av_1)^2/a^2 + (bv_2)^2/b^2 = 1$, but the left side is $v_1^2 + v_2^2$, so we have $v_1^2 + v_2^2 = 1$, which says that (v_1, v_2) lies on the circle C.

In other words, if $\mathbf{v} \in C$ then $T_{a,b}(\mathbf{v}) \in E_{a,b}$ and likewise if $T_{a,b}(\mathbf{v}) \in E_{a,b}$ then $\mathbf{v} \in C$. Hence, by the description of $T_{a,b}$ in (a), $E_{a,b}$ is obtained from C by applying horizontal scaling by a and vertical scaling by b (even though the equation $(x/a)^2 + (y/b)^2 = 1$ of $E_{a,b}$ involves *dividing* by a and *dividing* by b; cf. Exercise 13.1).

(d) We need to sketch $E_{2,3}$ and $E_{2,1/2}$; these are respectively the output of applying $T_{2,3}$ and $T_{2,1/2}$ to C. For the first, we double lengths horizontally and triple lengths vertically; this carries the x-intercepts $(\pm 1, 0)$ of C to $(\pm 2, 0)$ and carries the y-intercepts $(0, \pm 1)$ of C to $(0, \pm 3)$. For the second, we double lengths horizontally but *halve* lengths vertically; this carries the x-intercepts

$(\pm 1, 0)$ of C to $(\pm 2, 0)$ and carries the y-intercepts $(0, \pm 1)$ of C to $(0, \pm 1/2)$. The outcome of these is sketched in Figure 5.

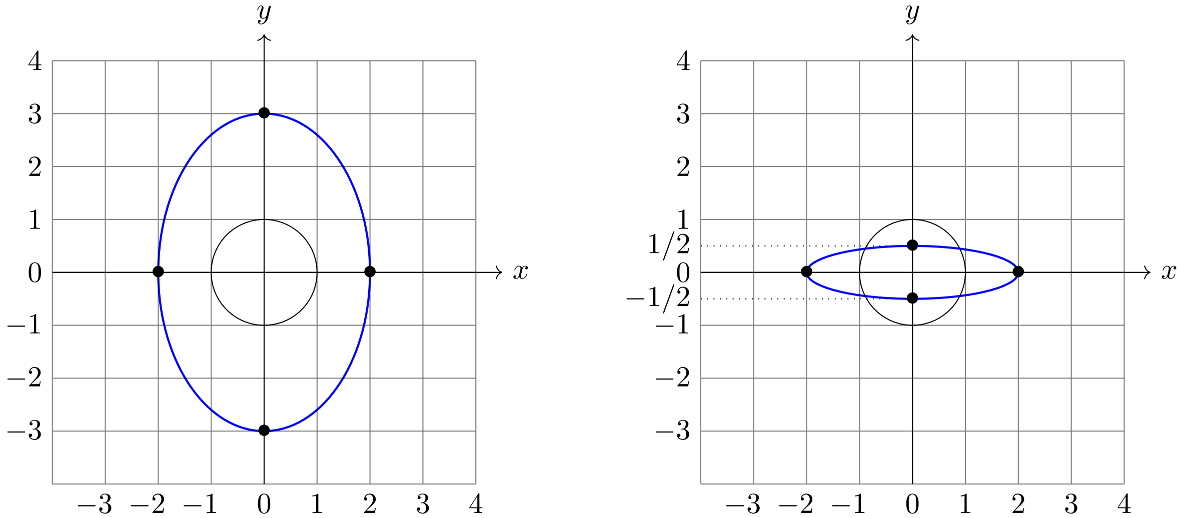

FIGURE 5. The curves $E_{2,3}$ (on the left) and $E_{2,1/2}$ (on the right), each along with the unit circle C for comparison purposes).

$\Diamond$

7. As a general observation (which isn't expected in your solution, and we mention here just for context), if $A, B, C > 0$ then the equation $Ax^2 + By^2 = C$ is an ellipse because dividing the equation by C turns it into $(A/C)x^2 + (B/C)y^2 = 1$ that is the same as $x^2/a^2 + y^2/b^2 = 1$ for $a = \sqrt{C/A}$ and $b = \sqrt{C/B}$. This is *not* used in the solutions below.

   (a) We treat each ellipse in turn.

      (i) The x-intercept corresponds to setting $y = 0$, so it is $x^2 = 10$, or $x = \pm\sqrt{10}$. The y-intercept corresponds to setting $x = 0$, so it is $6y^2 = 10$, or equivalently $y^2 = 5/3$, so $y = \pm\sqrt{5/3}$. Since $9 < 10 < 16$ and $1 < 5/3 < 4$ (as $5/3$ is between 1 and 2), we have $\pm\sqrt{10}$ lies between ± 3 and ± 4 and likewise $\pm\sqrt{5/3}$ lies between ± 1 and ± 2.

      (ii) The x-intercept corresponds to setting $y = 0$, so it is $3x^2 = 13$, or equivalently $x^2 = 13/3$, so $x = \pm\sqrt{13/3}$. The y-intercept corresponds to setting $x = 0$, so it is $5y^2 = 13$, or equivalently $y^2 = 13/5$, so $y = \pm\sqrt{13/5}$. Since $4 < 13/3 < 9$ (as $13/3$ lies between 4 and 5) and $1 < 13/5 < 4$ (as $13/5$ lies between 2 and 3), we have $\pm\sqrt{13/3}$ lies between ± 2 and ± 3, and likewise $\pm\sqrt{13/5}$ lies between ± 1 and ± 2.

      (iii) The x-intercept corresponds to setting $y = 0$, so it is $7x^2 = 18$, or equivalently $x^2 = 18/7$, so $x = \pm\sqrt{18/7}$. The y-intercept corresponds to setting $x = 0$, so it is $2y^2 = 18$, or equivalently $y^2 = 9$, so $y = \pm 3$. Since $1 < 18/7 < 4$ (as $18/7$ lies between 2 and 3), we have $\pm\sqrt{18/7}$ lies between ± 1 and ± 2.

      (iv) The x-intercept corresponds to setting $y = 0$, so it is $5x^2 = 21$, or equivalently $x^2 = 21/5$, so $x = \pm\sqrt{21/5}$. The y-intercept corresponds to setting $x = 0$, so it is $y^2 = 21$, so $y = \pm\sqrt{21}$. Since $4 < 21/5 < 9$ (as $21/5$ lies between 4 and 5), we have $\pm\sqrt{21/5}$ lies between ± 2 and ± 3. Since $16 < 21 < 25$, we have $\pm\sqrt{21}$ lies between ± 4 and ± 5.

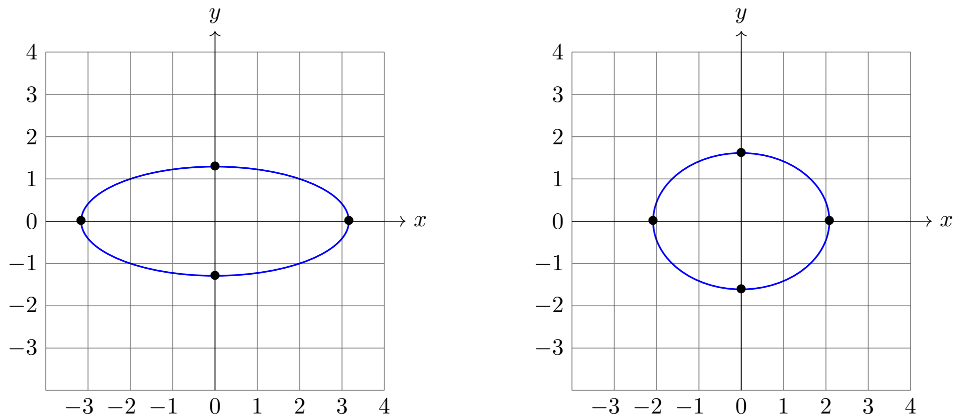

FIGURE 6. The ellipses $x^2 + 6y^2 = 10$ (left) and $3x^2 + 5y^2 = 13$ (right).

(b) The pictures for (i) and (ii) are respectively on the left and right in Figure 6, where the intercepts lie between the consecutive integers are as determined. Both ellipses are longer along the x-axis direction.

The pictures for (iii) and (iv) are respectively on the left and right in Figure 7, where the intercepts lie between the consecutive integers are as determined. Both ellipses are longer along the y-axis direction.

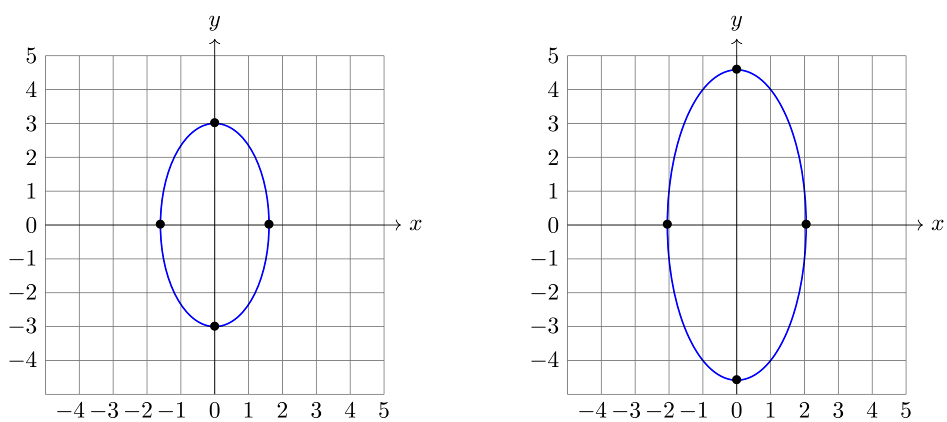

FIGURE 7. The ellipses $7x^2 + 2y^2 = 18$ (left) and $5x^2 + y^2 = 21$ (right).

(c) The pattern is that in all cases, whichever of x^2 or y^2 has the *smaller* coefficient corresponds to the axis line along which the ellipse is *longer*. $\Diamond$

8. (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,

$$(Df)(\mathbf{c}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{mn} \end{bmatrix} = A. \qquad \Diamond$$

9. (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). ◊

10. (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:

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

11. (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. $\diamondsuit$

12. (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 8.

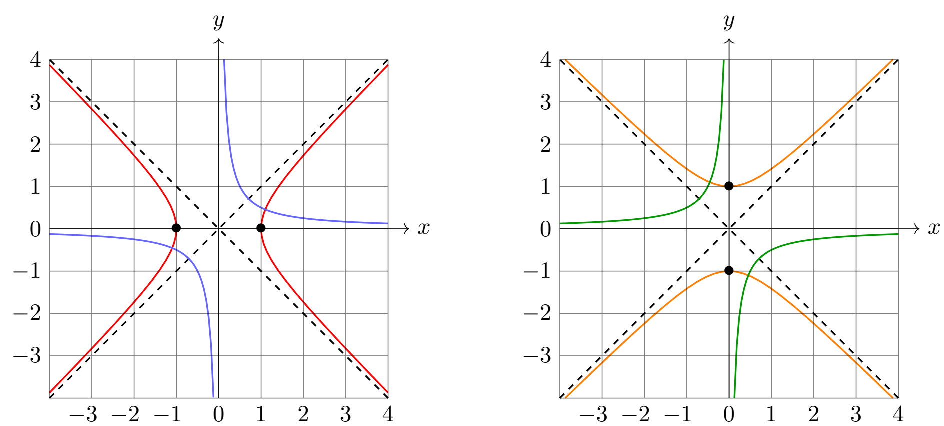

FIGURE 8. The curves H_+ (red) and H_- (orange), with dotted asymptotes $y = \pm x$ and black dots where each crosses the coordinate axes. The graphs of $1/(2x)$ (blue) and $-1/(2x)$ (green) with coordinate axes as their asymptotes are overlaid on the curves H_+ and H_- respectively, for comparison purposes.

(c) The reasoning as in Exercise 6 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 9.

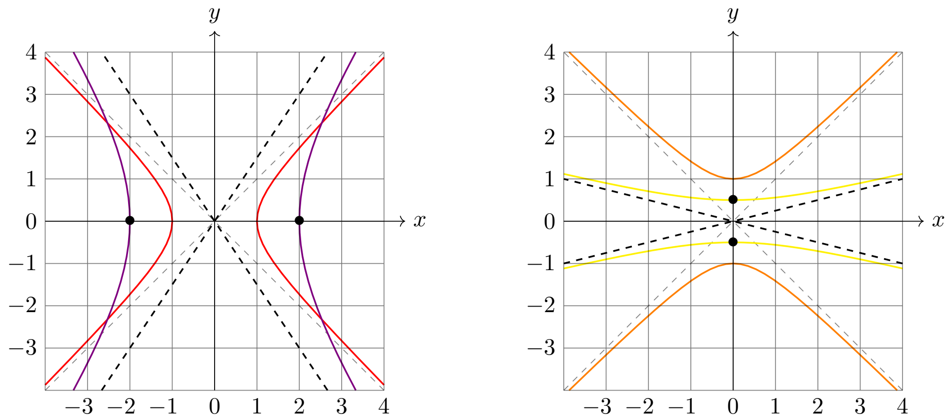

FIGURE 9. The curves $x^2/4 - y^2/9 = 1$ (purple) and $x^2/4 - 4y^2 = -1$ (yellow) and their respective dotted asymptotes $y = \pm(3/2)x$ and $y = \pm x/4$, and black dots where each graph crosses the coordinate axes. The original curves H_+ (red) and H_- (orange) and their asymptotes (in a lighter tone)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

$\Diamond$

13. 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:

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

$\Diamond$

14. (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. ♢

15. The key thing for us to note is that A is diagonal, so multiplying by it on the left or right rescales the rows or columns of the other matrix, respectively. Thus, $AB = \begin{bmatrix} 2b_1 & 2b_2 & 0 \\ 2b_4 & 2b_5 & 0 \\ 0 & 0 & 6b_9 \end{bmatrix} = BA$. Also, the top right entry of AB' is $2 \cdot 1 = 2$, whereas the top right entry of $B'A$ is $6 \cdot 1 = 6$. Therefore, $AB' \neq B'A$. ♢