2.5.8 False: while the image of g is obviously in the kernel of f (or the composition would not be 0), the image if g is not necessarily all of the kernel of f. (For example, if both f and g are 0 functions, with $m \neq 0$ then $\operatorname{img} g = 0 \neq \ker\, f = \mathbb{R}^m$.)

2.5.9 a. The matrix of a linear transformation has as its columns the images of the standard basis vectors, in this case identified to the polynomials $p_1(x) = 1,\ p_2(x) = x,\ p_3(x) = x^2$. Since

$$T(p_1)(x) = 0, \quad T(p_2)(x) = x, \quad T(p_3)(x) = 4x^2,$$

the matrix of T is $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$.

b. The matrix $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ row reduces to $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Thus the image has dimension 2 (the number of pivotal columns) and has a basis made up of the polynomials $ax + 4bx^2$, the linear combinations of the second and third columns of the matrix of T. The kernel has dimension 1 (the number of nonpivotal columns), and consists precisely of the constant polynomials.

2.5.10 Recall from equation 2.5.15 that the linear transformation $T_k : P_k \to \mathbb{R}^{k+1}$ given by

$$T_k(p) = \begin{bmatrix} p(0) \\ \vdots \\ p(k) \end{bmatrix}$$

is invertible, i.e., there exists $T_k^{-1} : \mathbb{R}^{k+1} \to P_k$ such that

$$T_k^{-1} \begin{bmatrix} p(0) \\ \vdots \\ p(k) \end{bmatrix} = p.$$

The linear transformation $\mathbb{R}^{k+1} \to \mathbb{R}$ given by

$$\vec{\mathbf{a}} \mapsto \int_0^n \left(T_k^{-1}(\vec{\mathbf{a}})\right)(t)dt$$

has a matrix, which is a line matrix $[c_0, \dots, c_k]$. The assertion is exactly that

$$[c_0, \dots, c_k] \begin{bmatrix} p(0) \\ \vdots \\ p(k) \end{bmatrix} = \int_0^n p(t)dt.$$

Remark. The entries $c_0, \dots, c_k$ depend on the interval over which one is integrating; we are integrating over the interval from 0 to n, and there are different c_i for each n.

Actually computing the numbers c_i by which each "sampled value" $p(i)$ of the polynomial must be weighted can be quite involved if done by hand. In the case where $k = 2$ we can use MATLAB or the equivalent to compute $T_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -3/2 & 2 & -1/2 \\ 1/2 & -1 & 1/2 \end{bmatrix}$, but computing it from the definition (the ith column of T_2^{-1} is $T_2^{-1}(\vec{\mathbf{e}}_i)$) is already fairly involved; for example, $T_2^{-1}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = a_0 + a_1 x + a_2 x^2$ tells us that $a_0 = 1, a_0 + a_1 + a_2 = 0$, and $a_0 + 2a_1 + 4a_2 = 0$, i.e., $a_0 = 1, a_1 = -3/2, a_2 = 1/2$. Thus to compute c_0 we would compute

$$c_0 = \int_0^n \left(1 - \frac{3}{2}x + \frac{1}{2}x^2\right) dx.$$

Another approach would be to use the Lagrange interpolation formula.

2.5.11 The sketch is shown at left.

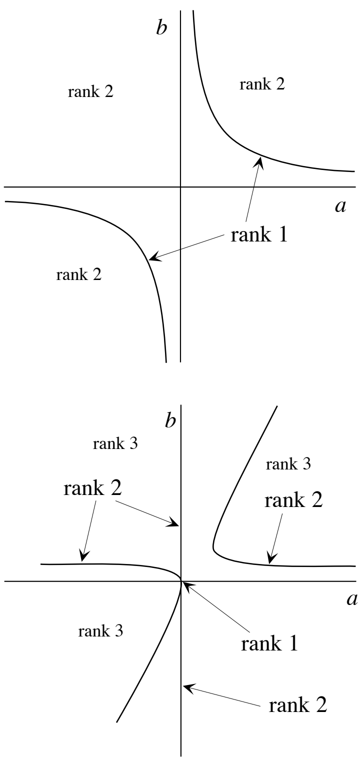

FIGURE FOR SOLUTION 2.5.11

TOP: On the curves, the kernel of A has dimension 1 and its image has dimension 1. Elsewhere, the rank (dimension of the image) is 2, so by the dimension formula the kernel has dimension 0. The rank is never 0 or 3.

BOTTOM: On the b-axis and on the hyperbola, the image of B has dimension 2, i.e., its kernel has dimension 1. At the origin the rank is 1 and the dimension of the kernel is 2. Elsewhere, the kernel has dimension 0 and the rank is 3.

a. If $ab \neq 2$, then $\dim(\ker(A)) = 0$, so in that case the image has dimension 2. If $ab = 2$, the image and the kernel have dimension 1.

b. This is more complicated. By row operations, we can bring the matrix B to

$$\begin{bmatrix} 1 & 2 & a \\ 0 & b & ab - a \\ 0 & 2a - b & a \end{bmatrix}.$$

We now separate the case $b \neq 0$ and $b = 0$.

• If $b \neq 0$, then we can do further row operations to bring the matrix to the form

$\begin{bmatrix} 1 & 0 & a - 2\frac{ab-a}{b} \\ 0 & 1 & \frac{ab-a}{b} \\ 0 & 0 & -a - (b - 2a)\frac{ab-a}{b} \end{bmatrix}$. The entry in the 3rd row, 3rd column is $-\frac{a}{b}(b^2 - 2ab + 2a)$.

So if $b \neq 0$, and the point $\begin{pmatrix} a \\ b \end{pmatrix}$ is neither on the line $a = 0$ nor on the hyperbola of equation $b^2 - 2ab + 2a = 0$, the matrix has rank 3, whereas if $b \neq 0$ and the point $\begin{pmatrix} a \\ b \end{pmatrix}$ is on one of these curves, the matrix has rank 2.

• If $b = 0$, the matrix is $\begin{bmatrix} 1 & 2 & a \\ 0 & -2a & -a \\ 0 & 0 & a \end{bmatrix}$, which evidently has rank 3 unless $a = 0$, in which case it has rank 1.

2.5.12 a. If we put the right side on a common denominator, we find

$$x + x^2 = A(x^2 + 5x + 6) + B(x^2 + 4x + 3) + C(x^2 + 3x + 2)$$

which leads to the system of linear equations

$$\begin{aligned} A + B + C &= 1 \\ 5A + 4B + 3C &= 1 \\ 6A + 3B + 2C &= 0. \end{aligned}$$

One way to solve this system of equations is to see that the matrix of coefficients

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 5 & 4 & 3 \\ 6 & 3 & 2 \end{bmatrix}$$

is invertible, with inverse

$$M^{-1} = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ -4 & 2 & -1 \\ 9/2 & -3/2 & 1/2 \end{bmatrix},$$

and that the solution is

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ -4 & 2 & -1 \\ 9/2 & -3/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = [\,0 \quad -2 \quad 3\,].$$

If you now look back at the problem, you will observe that $x^2+x = x(x+1)$, and that the $x+1$'s cancel. That explains why $A = 0$.

b. This time, if you put the right side on a common denominator, you find

$$\begin{aligned} &(A+C)x^4 + (-3A + B - 2C + D)x^3 + (3A - 3B + C - 2D + F)x^2 \\ &\qquad + (-A + 3B + D - 2F)x + (-B + F) = x + x^3, \end{aligned}$$

which leads to the system of equations

$$\begin{aligned} A + C &= 0 \\ -3A + B - 2C + D &= 1 \\ 3A - 3B + C - 2D + F &= 0 \\ -A + 3B + D - 2F &= 1 \\ -B + F &= 0. \end{aligned}$$

This time let us solve the system in the obvious way, by row reduction. The matrix

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ -3 & 1 & 2 & 1 & 0 & 1 \\ 3 & -3 & 1 & 2 & 1 & 0 \\ -1 & 3 & 0 & 1 & 2 & 1 \\ 0 & -1 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{row reduces to} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1/8 \\ 0 & 1 & 0 & 0 & 0 & 1/8 \\ 0 & 0 & 1 & 0 & 0 & 1/8 \\ 0 & 0 & 0 & 1 & 0 & 1/4 \\ 0 & 0 & 0 & 0 & 1 & 1/8 \end{bmatrix}.$$

In particular, the matrix of coefficients is invertible, since it row reduces to the identity. This gives the answer:

$$\frac{x + x^3}{(x+1)^2(x-1)^3} = -\frac{1}{8}\frac{x-1}{(x+1)^2} + \frac{1}{8}\frac{x^2+2x+1}{(x-1)^3}.$$

2.5.13 a. As in the example following proposition 2.5.14, we need to put the right side on a common denominator and consider the resulting system of linear equations. Row reduction then tells us for what values of a the system has no solutions. So:

$$\begin{aligned}\frac{x-1}{(x+1)(x^2+ax+5)} &= \frac{A_0}{x+1} + \frac{B_1x+B_0}{x^2+ax+5} \\ &= \frac{A_0x^2+aA_0x+5A_0+B_1x^2+B_1x+B_0x+B_0}{(x+1)(x^2+ax+5)}.\end{aligned}$$

This gives

$$x-1 = A_0x^2+aA_0x+5A_0+B_1x^2+B_1x+B_0x+B_0,$$

i.e.,

$$\begin{aligned}5A_0+B_0 &= -1 \\ aA_0+B_1+B_0 &= 1\ , \\ A_0+B_1 &= 0.\end{aligned} \quad \text{which we can write as} \quad \begin{bmatrix}5 & 0 & 1 & -1 \\ a & 1 & 1 & 1 \\ 1 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}A_0 \\ B_1 \\ B_0\end{bmatrix}.$$

Row reduction gives $\begin{bmatrix}1 & 0 & 0 & \frac{2}{a-6} \\ 0 & 1 & 0 & \frac{-2}{a-6} \\ 0 & 0 & 1 & \frac{-4-a}{a-6}\end{bmatrix}$, so the fraction in question cannot be written as a partial fraction when $a=6$.

b. This does not contradict proposition 2.5.14 because that proposition requires that p be factored as

$$p\,(x) = (x-a_1)^{n_1}\cdots(x-a_k)^{n_k}.$$

with the a_i distinct. If you substitute 6 for a in x^2+ax+5 you get $x^2+6x+5=(x+1)(x+5)$, so factoring p/q as

$$\frac{p(x)}{q(x)} = \frac{x-1}{(x+1)(x^2+ax+5)} = \frac{A_0}{x+1} + \frac{B_1x+B_0}{x^2+ax+5} = \frac{A_0}{x+1} + \frac{B_1x+B_0}{(x+1)(x+5)}$$

does not meet that requirement; both terms contain $(x+1)$ in the denominator.

Note that you could avoid this by using a different factorization:

$$\frac{p(x)}{q(x)} = \frac{x-1}{(x+1)(x^2+6x+5)} = \frac{x-1}{(x+1)^2(x+5)} = \frac{A_1x+A_0}{(x+1)^2} + \frac{B_0}{x+5}$$

2.5.14 a. We have

$$g\circ f(x) = x+(A+\alpha)x^2+(2A\alpha+B+\beta)x^3+O(x^4),$$

where $O(x^4)$ represents terms of degree 4 or higher, which we will ignore. If $A+\alpha=0=2A\alpha+B+\beta$, then $g(f(x))-x$ will have x^4 (or a higher power of x) as its lowest order term. These two equations are simple to solve for α and β: $\alpha=-A$, and $\beta=2A^2-B$. Thus the g with the specified properties is $g(x)=x-Ax^2+(2A^2-B)x^3$.

b. Consider the composition

$$g \circ f(x) = (x + a_2x^2 + \ldots + a + kx^k) + b_2(x + a_2x^2 + \ldots + a + kx^k)^2 + \ldots + b_k(x + a_2x^2 + \ldots + a + kx^k)^k$$
$$= x + c_2x^2 + \ldots + c_kx^k + \ldots,$$

and notice that the coefficients $c_2, \ldots, c_k$ depend in a complicated way on the a's, but are of degree 1 as functions of the b's, say

$$c_j = c_{j,1}(\mathbf{a}) + c_{j,2}(\mathbf{a})b_2 + \ldots c_{j,k}(\mathbf{a})b_k.$$

The $c_{i,j}, 2 \le i, j \le k$ form a square matrix C, and if it is invertible, then it will be possible to choose the b's so that $c_j = 0, j = 2, \ldots, k$, which is the point of the exercise. It is enough to prove that its kernel is 0.

Suppose $C\mathbf{b} = 0$, and that j is the smallest index such that $b_j \neq 0$. Then the term

$$b_j(x + a_2x^2 + \cdots + a_kx^k)^j$$

will contribute a term b_jx^j, which cannot be canceled by any other term, since all others are of degree $> j$. This contradicts the statement $C\mathbf{b} = 0$, so the kernel of C is 0.

2.5.15 Note first the following results: if $T_1, T_2 : \mathbb{R}^n \to \mathbb{R}^n$ are linear transformations, then

1. the image of T_1 contains the image of $T_1 \circ T_2$, and
2. the kernel of $T_1 \circ T_2$ contains the kernel of T_2.

The first is true because for any $\vec{\mathbf{v}} \in \operatorname{img} T_1 \circ T_2$, there exists a vector $\vec{\mathbf{w}}$ such that $(T_1 \circ T_2)\vec{\mathbf{w}} = \vec{\mathbf{v}}$ (by the definition of image). Since $T_1(T_2(\vec{\mathbf{w}})) = \vec{\mathbf{v}}$, the vector $\vec{\mathbf{v}}$ is also in the image of T_1.

Recall that the *nullity* of a linear transformation is the dimension of its kernel.

The second is true because for any $\vec{\mathbf{v}} \in \ker T_2$, we have $T_2(\vec{\mathbf{v}}) = \vec{\mathbf{0}}$. Since $T_1(\vec{\mathbf{0}}) = \vec{\mathbf{0}}$, we see that

$$(T_1 \circ T_2)(\vec{\mathbf{v}}) = T_1\big(T_2(\vec{\mathbf{v}})\big) = T_1(\vec{\mathbf{0}}) = \vec{\mathbf{0}},$$

so $\vec{\mathbf{v}}$ is also in the kernel of $T_1 \circ T_2$.

By proposition 1.2.15, since the matrices A^{-1} and AB are invertible, so is the product

$$B = (A^{-1}A)B = A^{-1}(AB),$$

and $B^{-1} = (AB)^{-1}A$. Note that this uses associativity of matrix mutiplication (corollary 1.3.12)

If AB is invertible, then the image of A contains the image of AB by statement 1. So A has rank n, hence nullity 0 by the dimension formula, so A is invertible. Since $B = A^{-1}(AB)$, we have $B^{-1} = (AB)^{-1}A$.

For B, one could argue that $\ker B \subset \ker AB = \{\mathbf{0}\}$, so B has nullity 0, and thus rank n, so B is invertible.

2.5.16

***2.5.17** a. Since $p(0) = a + 0b + 0^2c = 1$, we have $a = 1$. Since $p(1) = a + b + c = 4$, we have $b + c = 3$. Since $p(3) = a + 3b + 9c = -2$, we have $3b + 9c = -3$. It follows that $c = -2$ and $b = 5$.

b. Let $M_{\mathbf{x}}$ be the linear transformation from the space of P_n of polynomials of degree at most n to $\mathbb{R}^{n+1}$ given by

$$p \mapsto \begin{bmatrix} p(x_0) \\ \vdots \\ p(x_n) \end{bmatrix}, \quad \text{where } \mathbf{x} = \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix}.$$

Assume that the polynomial q is in the kernel of $M_{\mathbf{x}}$. Then q vanishes at $n+1$ distinct points, so that either q is the zero polynomial or it has degree at least $n+1$. Since q cannot have degree greater than n, it must be the zero polynomial. So ker $(M_{\mathbf{x}}) = \{0\}$, hence $M_{\mathbf{x}}$ is injective, so by corollary 2.5.10 it is surjective. It follows that a solution of $M_{\mathbf{x}}(p) = \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix}$ exists and is unique.

Solution 2.5.17, part b : The polynomials themselves are almost certainly nonlinear, but $M_{\mathbf{x}}$ is linear since it handles polynomials that have already been evaluated at the given points.

c. Take the linear transformation $M'_{\mathbf{x}}$ from P_k to R^{2n+2} defined by

$$p \mapsto \begin{bmatrix} p(x_0) \\ \vdots \\ p(x_n) \\ p'(x_0) \\ \vdots \\ p'(x_n) \end{bmatrix}. \text{ If ker } (M'_{\mathbf{x}}) = 0, \text{ then a solution of } M'_{\mathbf{x}}(p) = \begin{bmatrix} a_0 \\ \vdots \\ a_n \\ b_0 \\ \vdots \\ b_n \end{bmatrix}$$

exists and so a value for k is $2n+2$ (as shown above). In fact this is the lowest value for k that always has a solution.

2.5.18

2.5.19 a. $$H_2 = \begin{bmatrix} 1 & 1 \\ 1/2 & 1/4 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 1 & 1 \\ 1/2 & 1/4 & 1/8 \\ 1/3 & 1/9 & 1/27 \end{bmatrix}.$$

b. $$H_n = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1/2 & 1/4 & 1/8 & \dots & 1/2^n \\ 1/3 & 1/9 & 1/27 & \dots & 1/3^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/n^2 & 1/n^3 & \dots & 1/n^n \end{bmatrix}.$$

c. If H_n is not invertible, then there exist numbers $a_1, \dots, a_n$ not all zero such that

$$f_{\vec{\mathbf{a}}}(1) = \cdots = f_{\vec{\mathbf{a}}}(n) = 0.$$

But we can write

$$f_{\vec{\mathbf{a}}}(x) = \frac{a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n}{x^n},$$

and the only way this function can vanish at the integers $1, \dots, n$ is if the numerator vanishes at all these points. But it is a polynomial of degree $n-1$, and cannot vanish at n different points without vanishing identically.

2.5.20 a. The matrices are

$$H_2 = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

b. More generally, the matrix is

$$H_n = \begin{bmatrix} 1 & 1/2 & 1/3 & \dots & 1/n \\ 1/2 & 1/3 & 1/4 & \dots & 1/(n+1) \\ 1/3 & 1/4 & 1/5 & \dots & 1/(n+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/n & 1/(n+1) & 1/(n+2) & \dots & 1/(2n-1) \end{bmatrix}.$$

c. The question is whether H_n is onto, which will happen if and only if it is one to one, i.e., if and only if its kernel is $\{0\}$. Thus the question is whether

$$f_{\vec{\mathbf{a}}}(1) = f_{\vec{\mathbf{a}}}(2) = \cdots = f_{\vec{\mathbf{a}}}(n) = 0$$

implies that $a_1 = a_2 = \cdots = a_n = 0$. This is indeed true: by putting all the terms of $f_{\vec{\mathbf{a}}}$ on a common denominator, we can write

$$f_{\vec{\mathbf{a}}}(x) = \frac{p_{\vec{\mathbf{a}}}(x)}{x(x-1)\cdots(x+n-1)}$$

where $p_{\vec{\mathbf{a}}}$ is a polynomial of degree at most $n-1$; requiring it to vanish at the n points $1, 2, \dots, n$ is saying that it is the zero polynomial, or equivalently, that $f_{\vec{\mathbf{a}}}$ is the zero function. But if any a_i is nonzero, then

$$\lim_{x \to -i+1} f_{\vec{\mathbf{a}}}(x) = \infty,$$

so $f_{\vec{\mathbf{a}}}$ cannot be the zero function.

2.5.21 a. If $P_{[\mathbf{v}]}$ is one to one, then $P_{[\mathbf{v}]}$ has kernel $\{\mathbf{0}\}$. It then follows that $\sum(a_i\vec{\mathbf{v}}_i) = \mathbf{0}$ has as its only solution $a_i = 0, \forall i$, so $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ are linearly independent.

Conversely, if the vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ are linearly independent, then the equation $\sum(a_i\vec{\mathbf{v}}_i) = \mathbf{0}$ has as its only solution $a_i = 0, \forall i$. This means that $P_{[\mathbf{v}]}$ has kernel $\{\mathbf{0}\}$ and so is one to one.

b. If $P_{[\mathbf{v}]}$ is onto, then $\forall \vec{\mathbf{w}} \in \mathbb{R}^m$, $\exists \vec{\mathbf{a}} \in \mathbb{R}^n$ such that

$$P_{[\mathbf{v}]}(\vec{\mathbf{a}}) = \sum(a_i\vec{\mathbf{v}}_i) = \vec{\mathbf{w}},$$

so the vectors $\vec{\mathbf{v}}_1, ...\vec{\mathbf{v}}_n$ span $\mathbb{R}^m$.

Conversely, if $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ span $\mathbb{R}^m$ then

$$\forall \vec{\mathbf{w}} \in \mathbb{R}^m, \exists a_1, ...a_n \quad \text{such that} \sum(a_i\vec{\mathbf{v}}_i) = \vec{\mathbf{w}},$$

so $\forall \vec{\mathbf{w}} \in \mathbb{R}^m$, $\exists \vec{\mathbf{a}} \in \mathbb{R}^n$ such that $P_{[\mathbf{v}]}(\vec{\mathbf{a}}) = \vec{\mathbf{w}}$. Therefore, $P_{[\mathbf{v}]}$ is onto.

c. The vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ form a basis of $\mathbb{R}^m$ if and only if they are linearly independent and they span $\mathbb{R}^m$. By parts a and b, this is equivalent to $P_{[\mathbf{v}]}$ being one to one (part a) and onto (part b).

***2.5.22** a. First, we will show that if there exists $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $T_1 = S \circ T_2$ then ker $T_2 \subset$ ker T_1:

Indeed, if $T_2(\vec{\mathbf{v}}) = 0$, then $(S \circ T_2)(\vec{\mathbf{v}}) = (T_1)(\vec{\mathbf{v}}) = 0$.

Now we will show that if ker $T_2 \subset$ ker T_1 then there exists $S : \mathbb{R}^n \to \mathbb{R}^n$ such that $T_1 = S \circ T_2$:

For any $\mathbf{v} \in \operatorname{img} T_2$, choose $\tilde{\mathbf{v}} \in \mathbb{R}^n$ such that $T_2(\tilde{\mathbf{v}}) = \mathbf{v}$, and set $S(\mathbf{v}) = T_1(\tilde{\mathbf{v}})$. We need to show that this does not depend on the choice of $\tilde{\mathbf{v}}$. If $\tilde{\mathbf{v}}_1$ also satisfies $T_2(\tilde{\mathbf{v}}_1) = \mathbf{v}$, then $\tilde{\mathbf{v}} - \tilde{\mathbf{v}}_1 \in$ ker $T_2 \subset$ ker T_1, so $T_1(\tilde{\mathbf{v}}) = T_1(\tilde{\mathbf{v}}_1)$, showing that S is well defined on $\operatorname{img} T_2$. Now extend it to $\mathbb{R}^n$ in any way, for instance by choosing a basis for $\operatorname{img} T_2$, extending it to a basis of $\mathbb{R}^n$, and setting it equal to 0 on all the new basis vectors.

b. If $\exists S : \mathbb{R}^n \to \mathbb{R}^n$ such that $T_1 = T_2 \circ S$ then $\operatorname{img} T_1 \subset \operatorname{img} T_2$:

For each $\vec{\mathbf{w}} \in \operatorname{img} T_1$ there is a vector $\vec{\mathbf{v}}$ such that $T_1\vec{\mathbf{v}} = \vec{\mathbf{w}}$ (by definition of image). If $T_1 = T_2 \circ S$, $T_2(S(\vec{\mathbf{v}})) = \vec{\mathbf{w}}$, so $\vec{\mathbf{w}} \in \operatorname{img} T_2$.

Conversely, we need to show that if $\operatorname{img} T_1 \subset \operatorname{img} T_2$ then $\exists S : \mathbb{R}^n \to \mathbb{R}^n$ such that $T_1 = T_2 \circ S$.

Choose, for each i, a vector $\vec{\mathbf{v}}_i$ such that

$$T_2\vec{\mathbf{v}}_i = T_1(\vec{\mathbf{e}}_i).$$

This is possible, since $\operatorname{img} T_2 \supset \operatorname{img} T_1$.

Set $S = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n]$. Then $T_1 = T_2 \circ S$, since

$$(T_2 \circ S)(\vec{\mathbf{e}}_i) = T_2(\vec{\mathbf{v}}_i) = T_1(\vec{\mathbf{e}}_i).$$

2.6.1 a. It corresponds to the basis $\underline{\mathbf{v}}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\underline{\mathbf{v}}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\underline{\mathbf{v}}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\underline{\mathbf{v}}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. We have $\begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} = 2\underline{\mathbf{v}}_1 + \underline{\mathbf{v}}_2 + 5\underline{\mathbf{v}}_3 + 4\underline{\mathbf{v}}_4$.

b. It corresponds to the basis $\underline{\mathbf{v}}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\underline{\mathbf{v}}_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\underline{\mathbf{v}}_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\underline{\mathbf{v}}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. We have $\begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} = 2\underline{\mathbf{v}}_1 + 5\underline{\mathbf{v}}_2 + \underline{\mathbf{v}}_3 + 4\underline{\mathbf{v}}_4$.

2.6.2 There is almost nothing to this: everything is true about functions $f \in \mathcal{C}(0,1)$ because it is true of $f(x)$ for each $x \in (0,1)$. Remember that $0 \in \mathcal{C}(0,1)$ stands for the zero *function*; to distinguish it from the number

0 we will denote it by $\tilde{0}$.

$$\begin{aligned}
\big((f+g)+h\big)(x) &= (f+g)(x)+h(x) = \big(f(x)+g(x)\big)+h(x)\\
&= f(x)+(g(x)+h(x)) = f(x)+(g+h)(x) = \big(f+(g+h)\big)(x)\\
(f+g)(x) &= f(x)+g(x) = g(x)+f(x) = (g+f)(x)\\
(f+\tilde{0})(x) &= f(x)+0 = f(x)\\
(f+(-f))(x) &= f(x)-f(x) = 0 = \tilde{0}(x)
\end{aligned}$$

$$\begin{aligned}
\big((ab)f\big)(x) &= (ab)f(x) = a(b(f(x))) = (a(b(f)))(x)\\
\big(a(f+g)\big)(x) &= (af+ag))(x) = af(x)+ag(x) = ((af)+(ag))(x)\\
\big((a+b)f\big)(x) &= (a+b)(f(x)) = af(x)+bf(x) = (af+bg)(x)\\
(1f)(x) &= 1(f(x)) = f(x).
\end{aligned}$$

2.6.3

$$\Phi_{\{\underline{\mathbf{v}}\}}\left(\begin{bmatrix} a\\ b\\ c\\ d\end{bmatrix}\right) = a\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + b\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} + c\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} + d\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}a+b & c-d\\ c+d & a-b\end{bmatrix}$$

2.6.4 Define the linear transformation $T : V \times W \to \mathbb{R}^n$ by $T(\vec{\mathbf{v}}, \vec{\mathbf{w}}) \mapsto \vec{\mathbf{v}} - \vec{\mathbf{w}}$. The kernel of T is $V \cap W$. So by the dimension formula,

$$\dim\ker T + \dim\operatorname{img} T = \dim(V \times W) = \dim V + \dim W.$$

Since the image of T is a subspace of $\mathbb{R}^n$ and thus has dimension at most n,

$$\dim\ker T = \dim V + \dim W - \dim\operatorname{img} T \geq \dim V + \dim W - n.$$

2.6.5 a. The ith column of $[R_A]$ is $[R_A]\vec{\mathbf{e}}_i$:

$$[R_A]\vec{\mathbf{e}}_1 = \begin{bmatrix} a\\ b\\ 0\\ 0\end{bmatrix}, \quad \text{which corresponds to } \begin{bmatrix}a & b\\ 0 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

$$[R_A]\vec{\mathbf{e}}_2 = \begin{bmatrix} c\\ d\\ 0\\ 0\end{bmatrix}, \quad \text{which corresponds to } \begin{bmatrix}c & d\\ 0 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

$$[R_A]\vec{\mathbf{e}}_3 = \begin{bmatrix} 0\\ 0\\ a\\ b\end{bmatrix}, \quad \text{which corresponds to } \begin{bmatrix}0 & 0\\ a & b\end{bmatrix} = \begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

$$[R_A]\vec{\mathbf{e}}_4 = \begin{bmatrix} 0 \\ 0 \\ c \\ d \end{bmatrix}, \quad \text{which corresponds to } \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Similarly, the first column of $[L_A]$ corresponds to $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$; the second column to $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, the third to $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, and the fourth to $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

b. From part a. we have

$$|R_A| = |L_A| = \sqrt{2a^2 + 2b^2 + 2c^2 + 2d^2} = \sqrt{2}|A|.$$

2.6.6 a. The matrix for the transformation $L_A : B \to AB$ that multiplies a 3×3 matrix on the left by $\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix}$ is

$$L_A = \begin{bmatrix} a_1 & 0 & 0 & a_2 & 0 & 0 & a_3 & 0 & 0 \\ 0 & a_1 & 0 & 0 & a_2 & 0 & 0 & a_3 & 0 \\ 0 & 0 & a_1 & 0 & 0 & a_2 & 0 & 0 & a_3 \\ a_4 & 0 & 0 & a_5 & 0 & 0 & a_6 & 0 & 0 \\ 0 & a_4 & 0 & 0 & a_5 & 0 & 0 & a_6 & 0 \\ 0 & 0 & a_4 & 0 & 0 & a_5 & 0 & 0 & a_6 \\ a_7 & 0 & 0 & a_8 & 0 & 0 & a_9 & 0 & 0 \\ 0 & a_7 & 0 & 0 & a_8 & 0 & 0 & a_9 & 0 \\ 0 & 0 & a_7 & 0 & 0 & a_8 & 0 & 0 & a_9 \end{bmatrix}$$

The matrix for the transformation $R_A : B \mapsto AB$ is

$$\begin{bmatrix} a_1 & a_4 & a_7 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_2 & a_5 & a_8 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_3 & a_6 & a_9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_1 & a_4 & a_7 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_2 & a_5 & a_8 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 & a_6 & a_9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_1 & a_4 & a_7 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_2 & a_5 & a_8 \\ 0 & 0 & 0 & 0 & 0 & 0 & a_3 & a_6 & a_9 \end{bmatrix}.$$

b. The matrix for the transformation L_A that multiples an $n \times n$ matrix on the left is an $n^2 \times n^2$ matrix constructed as follows. The main diagonal consists of the diagonal entries of A, each appearing n times: first $a_{1,1}$, then $a_{2,2}$, etc. On either side of the main diagonal are $n-1$ smaller diagonals, whose entries are all 0. The next diagonal below the main diagonal contains the entries on the diagonal of A that is next to, and below, the main diagonal. Each entry appears n times. The next diagonal above the main diagonal contains the entries on the diagonal of A next to and above the main diagonal. Then we again have $n-1$ diagonals whose entries are all 0.

We continue until every entry of A has appeared n times in a row, always on a diagonal.

The matrix R_A is easier to describe. On the diagonal put n copies of $A^\top$, positioned so that the diagonal entries of each $A^\top$ are on the main diagonal of R_A. All other entries are 0.

2.6.7 a. This is not a subspace, since 0 is not in it.

b. This is a subspace: If f, g satisfy the differential equation, then so does $af + bg$:

$$(af + bg)(x) = af(x) + bg(x) = axf'(x) + bxg'(x) = x(af + bg)'(x).$$

c. This is not a vector space: the function $f(x) = x^2/4$ is in it, but $x^2 = 4(x^2/4)$ is not, so it isn't closed under multiplication by scalars.

2.6.8 a. Immediate from $(f + g)' = f' + g'$.

b. We must compute the polynomials $T(1) = 2,\ T(x) = x, T(x^2) = 2x^2 + 2 - 2x^2 + 2x^2 = 2 + 2x^2$. Now the coefficients of these polynomials are the desired matrix.

c. If we compute, we find $T(x^n) = (n^2 - 2n + 2)x^n + n(n-1)x^{n-2}$, which leads to

$$T = \begin{bmatrix} 2 & 0 & 2 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 6 & 0 & 0 & \dots \\ 0 & 0 & 3 & 0 & 12 & 0 & \dots \\ 0 & 0 & 0 & 5 & 0 & 20 & \dots \\ 0 & 0 & 0 & 0 & 10 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 17 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}.$$

2.6.9 a. Take any basis $\vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_n$ of V, and discard from the ordered set of vectors

$$\vec{\mathbf{v}}_1, \dots \vec{\mathbf{v}}_k, \vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_n$$

any vectors $\vec{\mathbf{w}}_i$ that are linear combinations of earlier vectors. At all stages, the set of vectors obtained will span V, since they do when you start and discarding a vector that is a linear combination of others doesn't change the span. When you are through, the vectors obtained will be linearly independent, so they satisfy condition 3 of definition 2.4.12.

b. The approach is identical: eliminate from $\vec{\mathbf{v}}_1, \dots \vec{\mathbf{v}}_k$ any vectors that depend linearly on earlier vectors; this never changes the span, and you end up with linearly independent vectors that span V.

2.6.10 Clearly,

$$A[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] = [A\mathbf{x}_1, A\mathbf{x}_2, \dots, A\mathbf{x}_n].$$

This can be rewritten

$$L_A \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{x}_1 \\ A\mathbf{x}_2 \\ \vdots \\ A\mathbf{x}_n \end{bmatrix},$$

where L_A is a linear transformation $L_A : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$. In this representation, it is clear that

$$L_A = \begin{bmatrix} A & 0 & \dots & 0 \\ 0 & A & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A \end{bmatrix}$$

so $|L_A|^2 = n|A|^2$.

The result is the same for R_A, but this time you have to take the entries of the matrix X by rows, i.e., write

$$\begin{bmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \\ \vdots \\ \mathbf{x}_n^\top \end{bmatrix} A = \begin{bmatrix} \mathbf{x}_1^\top A \\ \mathbf{x}_2^\top A \\ \vdots \\ \mathbf{x}_n^\top A \end{bmatrix} = \begin{bmatrix} (A^\top \mathbf{x}_1)^\top \\ (A^\top \mathbf{x}_2)^\top \\ \vdots \\ (A^\top \mathbf{x}_n)^\top \end{bmatrix}.$$

Thus in this basis, the matrix of R_A is

$$\begin{bmatrix} A^\top & 0 & \dots & 0 \\ 0 & A^\top & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A^\top \end{bmatrix},$$

and again $|R_A|^2 = n|A^\top|^2 = n|A|^2$.

2.6.11 We have

$$AB = \begin{bmatrix} 1+ab & a \\ b & 1 \end{bmatrix}, \quad BA = \begin{bmatrix} 1 & a \\ b & 1+ab \end{bmatrix}.$$

Thus we are asking about the rank of the matrix

$$\begin{bmatrix} 1 & 1 & 1+ab & 1 \\ a & 0 & a & a \\ 0 & b & b & b \\ 1 & 1 & 1 & 1+ab \end{bmatrix}.$$

We need to row reduce this matrix, but before starting let us see what happens if $a = 0$, or $b = 0$, or both. If $a = 0$, the matrix is $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & b & b & b \\ 1 & 1 & 1 & 1 \end{bmatrix}$, which evidently has rank 2 if $b \neq 0$, and rank 1 if $b = 0$. Similarly, if $b = 0$ and $a \neq 0$, the matrix has rank 2. Now let us suppose that $ab \neq 0$. Then row reduction gives

$$\begin{bmatrix} 1 & 1 & 1+ab & 1 \\ a & 0 & a & a \\ 0 & b & b & b \\ 1 & 1 & 1 & 1+ab \end{bmatrix} \to \begin{bmatrix} 1 & 1 & 1+ab & 1 \\ 0 & -a & -a^2b & 0 \\ 0 & b & b & b \\ 0 & 0 & -ab & ab \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & ab & 0 \\ 0 & a & 1 & 1 \\ 0 & 0 & a-a^2b & a \\ 0 & 0 & 1 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & ab \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & a(2-ab) \end{bmatrix}.$$