Student Solution Manual

to accompany the 5th edition of

Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach

John Hamal Hubbard

Barbara Burke Hubbard

Cornell University
Université Aix-Marseille

Matrix Editions Ithaca, NY 14850 MatrixEditions.com

214 University Ave. Ithaca, NY 14850
www.MatrixEditions.com

Printed in the United States of America

10 9 8 7 6 5 4 3 2

ISBN: 978-0-9715766-9-8

Contents

Note to Students

This manual gives solutions to all odd-numbered exercises in the fifth edition of *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach*.

Please think about the exercises before looking at the solutions! You will get more out of them that way. Some exercises are straightforward; their solutions are correspondingly straightforward. Others are "entertaining".

You may be tempted go straight to the solution manual to look for problems similar to those assigned as your homework, without reading the text. This is a mistake. You may learn some of the simpler material that way, but if you don't read the text, soon you will find you cannot understand the solutions. The textbook contains a lot of challenging material; you really have to come to terms with the underlying ideas and definitions.

Errata for this manual and for the textbook will be posted at

www.matrixeditions.com/errata.html

If you think you have found errors either in the text or in this manual, please email us at

hubbard@matrixeditions.com

or

jhh8@cornell.edu

so that we can make the errata lists as complete as possible.

You should not attempt to use this manual with a previous edition of the textbook, as some sections of the textbook have been renumbered, new sections have been added, some exercises have been moved from one section to another, some earlier exercises have been discarded, and some new ones have been added.

In this manual, we do not number every equation. When we do, we use (1), (2), . . . , starting again with (1) in a new solution.

John H. Hubbard and Barbara Burke Hubbard

Solutions for Chapter 0

0.2.1 a. The negation of the statement is: There exists a prime number such that if you divide it by 4 you have a remainder of 1, but which is not the sum of two squares.

b. The negation is: There exist $x \in \mathbb{R}$ and $\epsilon > 0$ such that for all $\delta > 0$ there exists $y \in \mathbb{R}$ with $|y - x| < \delta$ and $|y^2 - x^2| \geq \epsilon$.

c. The negation is: There exists $\epsilon > 0$ such that for all $\delta > 0$, there exist $x, y \in \mathbb{R}$ with $|x - y| < \delta$ and $|y^2 - x^2| \geq \epsilon$.

Solution 0.2.1: The negation in part a is false; the original statement is true: $5 = 4+1$, $13 = 9+4$, $17 = 16+1$, $29 = 25+4, \dots$, $97 = 81+16, \dots$.

If you divide a whole number (prime or not) by 4 and get a remainder of 3, then it is never the sum of two squares: 3 is not, 7 is not, 11 is not, etc. You may well be able to prove this; it isn't all that hard. But the original statement about primes is pretty tricky.

The negation in part b is false also, and the original statement is true: it is the definition of the mapping $x \mapsto x^2$ being continuous. But in part c, the negation is true, and the original statement is false. Indeed, if you take $\epsilon = 1$ and any $\delta > 0$ and set $x = \frac{1}{\delta}$, $y = \frac{1}{\delta} + \frac{\delta}{2}$, then

$$\begin{aligned} |y^2 - x^2| &= |y - x||y + x| \\ &= \frac{\delta}{2}\left(\frac{2}{\delta} + \frac{\delta}{2}\right) > 1. \end{aligned}$$

The original statement says that the function $x \mapsto x^2$ is *uniformly continuous*, and it isn't.

0.3.1 a. $(A * B) * (A * B)$ b. $(A * A) * (B * B)$ c. $A * A$

0.4.1 a. No: Many people have more than one aunt, and some have none.

b. No, $\frac{1}{0}$ is not defined.

c. No.

0.4.3 Here are some examples:

a. "DNA of" from people (excluding clones and identical twins) to DNA patterns.

b. $f(x) = x$.

0.4.5 The following are well defined: $g \circ f : A \to C$, $h \circ k : C \to C$, $k \circ g : B \to A$, $k \circ h : A \to A$, and $f \circ k : C \to B$. The others are not, unless some of the sets are subsets of the others. For example, $f \circ g$ is not because the codomain of g is C, which is not the domain of f, unless $C \subset A$.

0.4.7 a. $f\big(g(h(3))\big) = f\big(g(-1)\big) = f(-3) = 8$

b. $f\big(g(h(1))\big) = f\big(g(-2)\big) = f(-5) = 25$

0.4.9 If the argument of the square root is nonnegative, the square root can be evaluated, so the open first and the third quadrants are in the natural domain. The x-axis is not (since $y = 0$ there), but the y-axis with the origin removed is in the natural domain, since $x/y = 0$ there.

0.4.11 The function is defined for $\{\, x \in \mathbb{R} \mid -1 < x < 0, \text{ or } 0 < x \,\}$. It is also defined for the negative odd integers.

0.5.1 Without loss of generality we may assume that the polynomial is of odd degree d and that the coefficient of x^d is 1. Write the polynomial

$$x^d + a_{d-1}x^{d-1} + \cdots + a_0.$$

Let $A = |a_0| + \cdots + |a_{d-1}| + 1$. Then

$$\left| a_{d-1}A^{d-1} + a_{d-2}A^{d-2} + \cdots + a_0 \right| \leq (A-1)A^{d-1},$$

and similarly,

$$\left|a_{d-1}(-A)^{d-1} + a_{d-2}(-A)^{d-2} + \cdots + a_0\right| \le (A-1)A^{d-1}.$$

Therefore,

$$p(A) \ge A^d - (A-1)A^{d-1} > 0 \quad \text{and} \quad p(-A) \le (-A)^d + (A-1)A^{d-1} < 0.$$

By the intermediate value theorem, there must exist c with $|c| < A$ such that $p(c) = 0$.

0.5.3 Suppose $f : [a,b] \to [a,b]$ is continuous. Then the function $g(x) = x - f(x)$ satisfies the hypotheses of the intermediate value theorem (Theorem 0.5.9): $g(a) \le 0$ and $g(b) \ge 0$. So there must exist $x \in [a,b]$ such that $g(x) = 0$, i.e., $f(x) = x$.

0.5.5 *First solution*

Let $\sum_{k=1}^{\infty} b_k$ be a rearrangement of $\sum_{k=1}^{\infty} a_k$, which means that there exists a bijective map $\alpha : \mathbb{N} \to \mathbb{N}$ such that $b_{\alpha(k)} = a_k$.

We give two solutions to Exercise 0.5.5.

Since the series $\sum_{k=1}^{\infty} a_k$ is absolutely convergent, we know it is convergent. Let $A = \sum_{k=1}^{\infty} a_k$. We must show that

$$(\forall \epsilon > 0)(\exists M)\left(m \ge M \implies \left|\sum_{k=1}^{m} b_k - A\right| < \epsilon\right).$$

First choose N such that $\sum_{N+1}^{\infty} |a_k| < \epsilon/2$; in that case,

$$(n > N) \implies \left|\sum_{k=1}^{n} a_k - A\right| \le \sum_{k=n}^{\infty} |a_k| \le \sum_{k=N+1}^{\infty} |a_k| < \epsilon/2.$$

Now set

$$M = \max\{\alpha(1), \dots, \alpha(N)\},$$

i.e., M is so large that all of $a_1, \dots, a_N$ appear among $b_1, \dots, b_M$. Then

$$m > M \implies \left|\sum_{k=1}^{m} b_k - A\right| \le \left|\sum_{k=1}^{N} a_k - A\right| + \sum_{k=N+1}^{\infty} |a_k| < \epsilon.$$

Second solution

The series $\sum_{k=1}^{\infty} a_k + |a_k|$ is a convergent series of positive numbers, so its sum is the sup of all finite sums of terms.

If $\sum_{k=1}^{\infty} b_k$ is a rearrangement of the same series, then $\sum_{k=1}^{\infty} b_k + |b_k|$ is also a rearrangement of $\sum_{k=1}^{\infty} a_k + |a_k|$.

But the finite sums of terms of $\sum_{k=1}^{\infty} a_k + |a_k|$ and $\sum_{k=1}^{\infty} b_k + |b_k|$ are exactly the same set of numbers, so these two series converge to the same limit. The same argument says that

$$\sum_{k=1}^{\infty} |a_k| = \sum_{k=1}^{\infty} |b_k|.$$

Thus

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty}(b_k + |b_k|) - \sum_{k=1}^{\infty}|b_k| = \sum_{k=1}^{\infty}(a_k + |a_k|) - \sum_{k=1}^{\infty}|a_k| = \sum_{k=1}^{\infty} a_k.$$

The first equality is justified because both series

$$\sum_{k=1}^{\infty}(b_k + |b_k|) \quad \text{and} \quad \sum_{k=1}^{\infty}|b_k|$$

converge, so their difference does also.

0.6.1 a. Begin by listing the integers between -1 and 1, then list the numbers between -2 and 2 that can be written with denominators ≤ 2 and which haven't already been listed, then list the rationals between -3 and 3 that can be written with denominators ≤ 3 and haven't already been listed, etc. This will eventually list all rationals. Here is the beginning of the list:

$$-1, 0, 1, -2, -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}, 2, -3, -\frac{8}{3}, -\frac{5}{2}, -\frac{7}{3}, -\frac{5}{3}, -\frac{4}{3}, -\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \dots$$

Solution 0.6.1, part a: There are infinitely many ways of solving this problem, and many seem just as natural as the one we propose.

b. Just as before, list first the finite decimals in $[-1, 1]$ with at most one digit after the decimal point (there are 21 of them), then the ones in $[-2, 2]$ with at most two digits after the decimal, and which haven't been listed earlier (there are 380 of these), etc.

0.6.3 It is easy to write a bijective map $(-1, 1) \to \mathbb{R}$. For instance,

$$f(x) = \frac{x}{1-x^2} \quad \text{or} \quad g(x) = \tan\frac{\pi x}{2}.$$

Solution 0.6.3: The same argument holds if one uses the function g, with derivative

$$g'(x) = \frac{\pi}{2(1+\cos^2\frac{\pi x}{2})}.$$

The derivative of the first mapping is

$$f'(x) = \frac{(1-x^2) + 2x^2}{(1-x^2)^2} = \frac{1+x^2}{(1-x^2)^2} > 0$$

so the mapping is monotone increasing on $(-1, 1)$, hence injective. Since it is continuous on $(-1, 1)$ and

$$\lim_{x \searrow -1} f(x) = -\infty \quad \text{and} \quad \lim_{x \nearrow +1} f(x) = +\infty,$$

it is surjective by the intermediate value theorem.

0.6.5 a. To the right, if $a \in A$ is an element of a chain, then it can be continued:

$$a, f(a), g\big(f(a)\big), f\big(g(f(a))\big), \dots,$$

and this is clearly the only possible continuation.

Similarly, if b is an element of a chain, then the chain can be continued

$$b, g(b), f\big(g(b)\big), g\big(f(g(b))\big), \dots,$$

and again this is the only possible continuation.

Let us set $A_1 = g(B)$ and $B_1 = f(A)$. Let $f_1 : A_1 \to B$ be the unique map such that $f_1(g(b)) = b$ for all $b \in B$, and let $g_1 : B_1 \to A$ be the unique map such that $g_1(f(a)) = a$ for all $a \in A$. To the left, if $a \in A_1$ is an element of a chain, we can extend to $f_1(a), a$. Then if $f(a_1) \in B_1$, we can extend one further, to $g_1(f_1(a))$, and if $g_1(f_1(a)) \in A_1$, we can extend further to

$$f_1\big(g_1(f_1(a))\big), \quad g_1\big(f_1(a)\big), \quad f_1(a), \quad a.$$

Either this will continue forever or at some point we will run into an element of A that is not in A_1 or into an element of B that is not in B_1; the chain necessarily ends there.

b. As we saw, any element of A or B is part of a unique infinite chain to the right and of a unique finite or infinite chain to the left.

Part c:

1. can be uniquely continued forever
2. can be continued to an element of A that is not in the image of g
3. can be continued to an element of B that is not in the image of f

c. Since every element of A and of B is an element of a unique chain, and since this chain satisfies either (1), (2), or (3) and these are exclusive, the mapping h is well defined.

If $h(a_1) = h(a_2)$ and a_1 belongs to a chain of type 1 or 2, then so does $f(a_1)$, hence so does $f(a_2)$, hence so does a_2, and then $h(a_1) = h(a_2)$ implies $a_1 = a_2$, since f is injective.

Now suppose that a_1 belongs to a chain of type 3. Then a_1 is not the first element of the list, so $a_1 \in A_1$, and $h(a_1) = f_1(a_1)$ is well defined. The element $h(a_2)$ is a part of the same chain, hence also of type 3, and $h(a_2) = f_1(a_2)$. But then $a_1 = g(f_1(a_1)) = g(f_1(a_2)) = a_2$. So h is injective. Now any element $b \in B$ belongs to a maximal chain. If this chain is of type 1 or 2, then b is not the first element of the chain, so $b \in B_1$ and $h(g_1(b)) = f(g_1(b)) = b$, so b is in the image. If b is in a chain of type 3, then $b = f_1(g(b)) = h(g(b))$, so again b is in the image of h, proving that h is surjective.

d. This is the only entertaining part of the problem. In this case, there is only one chain of type 1, the chain of all 0's. There are only two chains of type 2, which are

$$+1 \xrightarrow{f} \frac{1}{2} \xrightarrow{g} \frac{1}{2} \xrightarrow{f} \frac{1}{4} \xrightarrow{g} \frac{1}{4} \xrightarrow{f} \frac{1}{8} \dots$$

$$-1 \xrightarrow{f} -\frac{1}{2} \xrightarrow{g} -\frac{1}{2} \xrightarrow{f} -\frac{1}{4} \xrightarrow{g} -\frac{1}{4} \xrightarrow{f} -\frac{1}{8} \dots$$

All other chains are of type 3, and end to the left with a number in $(1/2, 1)$ or in $(-1, -1/2)$, which are the points in $B - f(A)$. Such a sequence might be

$$\frac{3}{4} \xrightarrow{g} \frac{3}{4} \xrightarrow{f} \frac{3}{8} \xrightarrow{g} \frac{3}{8} \xrightarrow{f} \frac{3}{16} \dots$$

The existence of this function h is *Bernstein's theorem*.

Following our definition of h, we see that

$$h(x) = \begin{cases} 0 & \text{if } x = 0 \\ x/2 & \text{if } x = \pm 1/2^k \text{ for some } k \geq 0 \\ x & \text{if } x \neq \pm 1/2^k \text{ for some } k \geq 0. \end{cases}$$

Solution 0.6.7: Two sets A and B have the same cardinality if there exists an invertible mapping $A \to B$. Here, finding such an invertible map would be difficult, so instead we use Bernstein's theorem, which says that if there is an injective map $A \to B$ and an injective map $B \to A$, then there is a invertible map $A \to B$.

0.6.7 a. There is an obvious injective map $[0, 1) \to [0, 1) \times [0, 1)$, given simply by $g : x \mapsto (x, 0)$.

We need to construct an injective map in the opposite direction; that is a lot harder. Take a point in $(x, y) \in [0, 1) \times [0, 1)$, and write both coordinates as decimals:

$$x = .a_1 a_2 a_3 \dots \quad \text{and} \quad y = .b_1 b_2 b_3 \dots;$$

if either number can be written in two ways, one ending in 0's and the other in 9's, use the one ending in 0's.

Remember that
injective = one to one
surjective = onto
bijective = invertible.

Now consider the number $f(x,y) = .a_1b_1a_2b_2a_3b_3\ldots$. The mapping f is injective: using the even and the odd digits of $f(x,y)$ allows you to reconstruct x and y. The only problem you might have is if $f(x,y)$ could be written in two different ways, but as constructed $f(x,y)$ will never end in all 9's, so this doesn't happen.

Bernstein's theorem is the object of Exercise 0.6.5.

Note that this mapping is not surjective; for instance, $.191919\ldots$ is not in the image. But Bernstein's theorem guarantees that since there are injective maps both ways, there is a bijection between $[0,1)$ and $[0,1)\times[0,1)$.

b. The proof of part a works also to construct a bijective mapping $(0,1)\to(0,1)\times(0,1)$. But it is easy to construct a bijective mapping $(0,1)\to\mathbb{R}$, for instance $x\mapsto\cot(\pi x)$. If we compose these mappings, we find bijective maps

$$\mathbb{R}\to(0,1)\to(0,1)\times(0,1)\to\mathbb{R}\times\mathbb{R}.$$

c. We can use part b repeatedly to construct bijective maps

Part c: The map

$$(f,\mathrm{id}):\mathbb{R}\times\mathbb{R}\to\mathbb{R}$$

maps the first $\mathbb{R}$ to $\mathbb{R}\times\mathbb{R}$ and the second $\mathbb{R}$ to itself.

$$\begin{aligned}\mathbb{R}\xrightarrow{f}\mathbb{R}\times\mathbb{R}\xrightarrow{(f,\mathrm{id})}(\mathbb{R}\times\mathbb{R})\times\mathbb{R}=\mathbb{R}\times\mathbb{R}^2\xrightarrow{(f,\mathrm{id})}(\mathbb{R}\times\mathbb{R})\times\mathbb{R}^2\\ =\mathbb{R}\times\mathbb{R}^3\xrightarrow{(f,\mathrm{id})}(\mathbb{R}\times\mathbb{R})\times\mathbb{R}^3=\mathbb{R}\times\mathbb{R}^4\cdots.\end{aligned}$$

Continuing this way, it is easy to get a bijective map $\mathbb{R}\to\mathbb{R}^n$.

0.6.9 It is not possible. For suppose it were, and consider as in equation 0.6.3 the decimal made up of the entries on the diagonal. If this number is rational, then the digits are eventually periodic, and if you do anything systematic to these digits, such as changing all digits that are not 7 to 7's, and changing 7's to 5's, the sequence of digits obtained will still be eventually periodic, so it will represent a rational number. This rational number must appear someplace in the sequence, but it doesn't, since it has a different kth digit than the kth number for every k.

The only weakness in this argument is that it might be a number that can be written in two different ways, but it isn't, since it has only 5's and 7's as digits.

Solution 0.7.3, part b: If you had trouble with this, note that it follows from Proposition 0.7.5 (geometrical representation of multiplication of complex numbers) that

$$\left|\frac{1}{z}\right|=\frac{1}{|z|},$$

since $zz^{-1}=1$, which has length 1. The modulus (absolute value) of $3+4i$ is $\sqrt{25}=5$, so the modulus of $(3+4i)^{-1}$ is $1/5$. It follows from the second statement in Proposition 0.7.5 that the polar angle of z^{-1} is minus the polar angle of z, since the two angles sum to 0, which is the polar angle of the product $zz^{-1}=1$. The polar angle of $3+4i$ is $\arccos(3/5)$.

0.7.1 "Modulus of z," "absolute value of z," and $|z|$ are synonyms. "Real part of z" is the same as $\operatorname{Re} z=a$. "Imaginary part of z" is the same as $\operatorname{Im} z=b$. The "complex conjugate of z" is the same as $\bar{z}$.

0.7.3 a. The absolute value of $2+4i$ is $|2+4i|=2\sqrt{5}$. The argument (polar angle) of $2+4i$ is $\arccos 1/\sqrt{5}$, which you could also write as $\arctan 2$.

b. The absolute value of $(3+4i)^{-1}$ is $1/5$. The argument (polar angle) is $-\arccos(3/5)$.

c. The absolute value of $(1+i)^5$ is $4\sqrt{2}$. The argument is $5\pi/4$. (The complex number $1+i$ has absolute value $\sqrt{2}$ and polar angle $\pi/4$. De Moivre's formula says how to compute these for $(1+i)^5$.)

d. The absolute value of $1+4i$ is $\sqrt{17}$; the argument is $\arccos 1/\sqrt{17}$.

0.7.5 Parts 1–4 are immediate. For part 5, we find

$$\begin{aligned}(z_1z_2)z_3 &= \big((x_1x_2-y_1y_2)+i(y_1x_2+x_1y_2)\big)(x_3+iy_3)\\ &= (x_1x_2x_3-y_1y_2x_3-y_1x_2y_3-x_1y_2y_3)\\ &\qquad + i(x_1x_2y_3-y_1y_2y_3+y_1x_2x_3+x_1y_2x_3),\end{aligned}$$

which is equal to

$$\begin{aligned}z_1(z_2z_3) &= (x_1+iy_1)((x_2x_3-y_2y_3)+i(y_2x_3+x_2y_3))\\ &= (x_1x_2x_3-x_1y_2y_3-y_1y_2x_3-y_1x_2y_3)\\ &\qquad + i(y_1x_2x_3-y_1y_2y_3+x_1y_2x_3+x_1x_2y_3).\end{aligned}$$

Parts 6 and 7 are immediate. For part 8, multiply out:

$$\begin{aligned}(a+ib)\left(\frac{a}{a^2+b^2}-i\frac{b}{a^2+b^2}\right) &= \frac{a^2}{a^2+b^2}+\frac{b^2}{a^2+b^2}+i\left(\frac{ab}{a^2+b^2}-\frac{ab}{a^2+b^2}\right)\\ &= 1+i0=1.\end{aligned}$$

Part 9 is also a matter of multiplying out:

$$\begin{aligned}z_1(z_2+z_3) &= (x_1+iy_1)\big((x_2+iy_2)+(x_3+iy_3)\big)\\ &= (x_1+iy_1)((x_2+x_3)+i(y_2+y_3))\\ &= x_1(x_2+x_3)-y_1(y_2+y_3)+i\big(y_1(x_2+x_3)+x_1(y_2+y_3)\big)\\ &= x_1x_2-y_1y_2+i(y_1x_2+x_1y_2)+x_1x_3-y_1y_3+i(y_1x_3+x_1y_3)\\ &= z_1z_2+z_1z_3.\end{aligned}$$

0.7.7 a. The equation $|z-u|+|z-v|=c$ represents an ellipse with foci at u and v, at least if $c>|u-v|$. If $c=|u-v|$ it is the degenerate ellipse consisting of the segment $[u,v]$, and if $c<|u-v|$ it is empty, by the triangle inequality, which asserts that if there is a z satisfying the equality, then

$$c<|u-v|\le|u-z|+|z-v|=c.$$

Solution 0.7.7: Remember that the set of points such that the sum of their distances to two points is constant is an ellipse, with foci at those points.

b. Set $z=x+iy$; the inequality $|z|<1-\operatorname{Re} z$ becomes

$$\sqrt{x^2+y^2}<1-x,$$

corresponding to a region bounded by the curve of equation

$$\sqrt{x^2+y^2}=1-x.$$

If we square this equation, we will get the curve of equation

$$x^2+y^2=1-2x+x^2, \quad \text{i.e.,} \quad x=\frac{1}{2}(1-y^2),$$

which is a parabola lying on its side. The original inequality corresponds to the inside of the parabola.

We should worry that squaring the equation might have introduced parasitic points, where $-\sqrt{x^2+y^2}<1-x$. This is not the case, since $1-x$ is positive throughout the region.

0.7.9 a. The quadratic formula gives $x=\dfrac{-i\pm\sqrt{-1-8}}{2}$, so the solutions are $x=i$ and $x=-2i$.

b. In this case, the quadratic formula gives

$$x^2 = \frac{-1 \pm \sqrt{1-8}}{2} = \frac{-1 \pm i\sqrt{7}}{2}.$$

Each of these numbers has two square roots, which we still need to find.

One way, probably the best, is to use the polar form; this gives

$$x^2 = r(\cos\theta \pm i \sin\theta),$$

where

$$r = \frac{\sqrt{1+7}}{2} = \sqrt{2}, \quad \theta = \pm \arccos -\frac{1}{2\sqrt{2}} \approx 1.2094\ldots \text{ radians.}$$

Thus the four roots are

$$\pm \sqrt[4]{2}(\cos\theta/2 + i \sin\theta/2) \quad \text{and} \quad \pm \sqrt[4]{2}(\cos\theta/2 - i\sin\theta/2).$$

c. Multiplying the first equation by $(1+i)$ and the second by i gives

$$\begin{aligned} i(1+i)x - (2+i)(1+i)y &= 3(1+i) \\ i(1+i)x - \qquad\qquad\quad y &= 4i, \end{aligned}$$

which gives

$$-(2+i)(1+i)y + y = 3 - i, \quad \text{i.e.,} \quad y = i + \frac{1}{3}.$$

Substituting this value for y then gives $x = \frac{7}{3} - \frac{8}{3}i$.

0.7.11 a. These are the vertical line $x = 1$ and the circle centered at the origin of radius 3.

b. Use $Z = X + iY$ as the variable in the codomain. Then

$$(1+iy)^2 = 1 - y^2 + 2iy = X + iY$$

gives $1 - X = y^2 = Y^2/4$. Thus the image of the line is the curve of equation $X = 1 - (Y^2/4)$, which is a parabola with horizontal axis.

The image of the circle is another circle, centered at the origin, of radius 9, i.e., the curve of equation $X^2 + Y^2 = 81$.

c. This time use $Z = X + iY$ as the variable in the domain. Then the inverse image of the line $= \operatorname{Re} z = 1$ is the curve of equation

$$\operatorname{Re}(X+iY)^2 = X^2 - Y^2 = 1,$$

which is a hyperbola. The inverse image of the curve of equation $|z| = 3$ is the curve of equation $|Z^2| = |Z|^2 = 3$, i.e., $|Z| = \sqrt{3}$, the circle of radius $\sqrt{3}$ centered at the origin.

0.7.13 a. The cube roots of 1 are

$$1, \quad \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + \frac{i\sqrt{3}}{2}, \quad \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - \frac{i\sqrt{3}}{2}.$$

b. The fourth roots of 1 are 1, i, -1, $-i$.

c. The sixth roots of 1 are

$$1, \quad -1, \quad \frac{1}{2}+i\frac{\sqrt{3}}{2}, \quad \frac{1}{2}-i\frac{\sqrt{3}}{2}, \quad -\frac{1}{2}+i\frac{\sqrt{3}}{2}, \quad -\frac{1}{2}-i\frac{\sqrt{3}}{2}$$

0.7.15 a. The fifth roots of 1 are

$$\cos 2\pi k/5 + i \sin 2\pi k/5, \quad \text{for } k = 0, 1, 2, 3, 4.$$

The point of the question is to find these numbers in some more manageable form. One possible approach is to set $\theta = 2\pi/5$, and to observe that $\cos 4\theta = \cos\theta$. If you set $x = \cos\theta$, this leads to the equation

$$2(2x^2-1)^2 - 1 = x \quad \text{i.e.,} \quad 8x^4 - 8x^2 - x + 1 = 0.$$

This still isn't too manageable, until you start asking what other angles satisfy $\cos 4\theta = \cos\theta$. Of course $\theta = 0$ does, meaning that $x = 1$ is one root of our equation. But $\theta = 2\pi/3$ does also, meaning that $-1/2$ is also a root. Thus we can divide:

$$\frac{8x^4 - 8x^2 - x + 1}{(x-1)(2x+1)} = 4x^2 + 2x - 1,$$

and $\cos 2\pi/5$ is the positive root of that quadratic equation, i.e.,

$$\cos\frac{2\pi}{5} = \frac{\sqrt{5}-1}{4}, \text{ which gives } \sin\frac{2\pi}{5} = \frac{\sqrt{10+2\sqrt{5}}}{4}.$$

The fifth roots of 1 are now

$$1, \quad \frac{\sqrt{5}-1}{4} \pm i\frac{\sqrt{10+2\sqrt{5}}}{4}, \quad -\frac{\sqrt{5}+1}{4} \pm i\frac{\sqrt{10-2\sqrt{5}}}{4}.$$

b. It is straightforward to draw a line segment of length $(\sqrt{5}-1)/4$: construct a rectangle with sides 1 and 2, so the diagonal has length $\sqrt{5}$. Then subtract 1 and divide twice by 2, as shown in the figure in the margin.

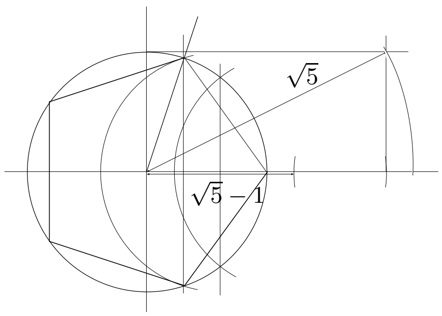


FIGURE FOR SOLUTION 0.7.15, part b

Solutions for Chapter 1

1.1.1 a. $\begin{bmatrix}1\\3\end{bmatrix}+\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}3\\4\end{bmatrix}$ b. $2\begin{bmatrix}2\\4\end{bmatrix}=\begin{bmatrix}4\\8\end{bmatrix}$

c. $\begin{bmatrix}1\\3\end{bmatrix}-\begin{bmatrix}2\\1\end{bmatrix}=\begin{bmatrix}1-2\\3-1\end{bmatrix}=\begin{bmatrix}-1\\2\end{bmatrix}$ d. $\begin{bmatrix}3\\2\end{bmatrix}+\vec{\mathbf{e}}_1=\begin{bmatrix}3\\2\end{bmatrix}+\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}4\\2\end{bmatrix}$

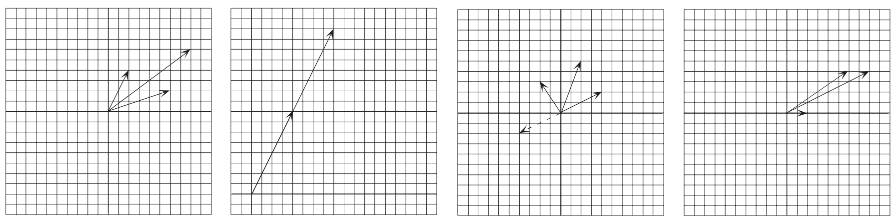
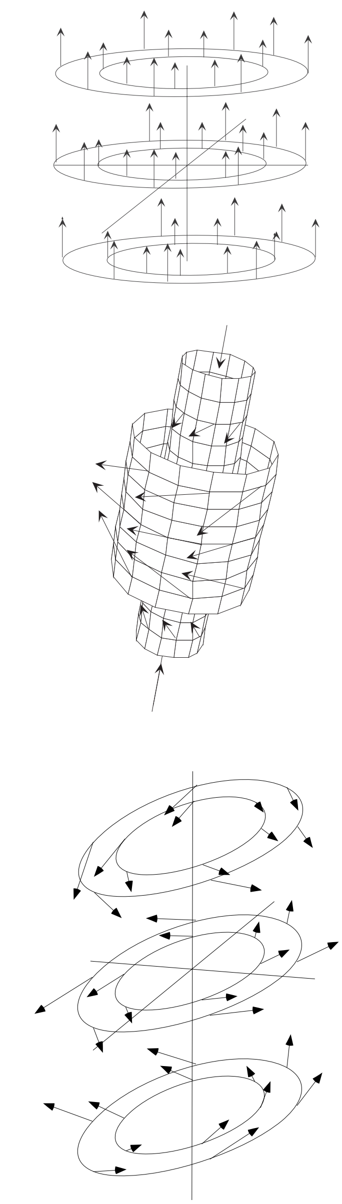


Figure for Solution 1.1.1. From left: (a), (b), (c), and (d).



Solution 1.1.7. Top: The vector field in part a points up. Middle: The vector field in part b is rotation in the (x, y)-plane. Bottom: The vector field in part c spirals out in the (x, y)-plane.

1.1.3 a. $\vec{\mathbf{v}} \in \mathbb{R}^3$ b. $L \subset \mathbb{R}^2$ c. $C \subset \mathbb{R}^3$ d. $\mathbf{x} \in \mathbb{C}^2$
e. $B_0 \subset B_1 \subset B_2, \dots$.

1.1.5 $\begin{bmatrix}1\\1\\\vdots\\1\\1\end{bmatrix}=\sum_{i=1}^n \vec{\mathbf{e}}_i$ $\begin{bmatrix}1\\2\\\vdots\\n-1\\n\end{bmatrix}=\sum_{i=1}^n i\vec{\mathbf{e}}_i$ $\begin{bmatrix}0\\0\\3\\4\\\vdots\\n-1\\n\end{bmatrix}=\sum_{i=3}^n i\vec{\mathbf{e}}_i$

1.1.7 As shown in the figure in the margin, the vector field in part a points straight up everywhere. Its length depends only on how far you are from the z-axis, and it gets longer and longer the further you get from the z-axis; it vanishes on the z-axis. The vector field in part b is simply rotation in the (x, y)-plane, like (f) in Exercise 1.1.6. But the z-component is down when $z > 0$ and up when $z < 0$. The vector field in part c spirals out in the (x, y)-plane, like (h) in Exercise 1.1.6. Again, the z-component is down when $z > 0$ and up when $z < 0$.

1.1.9 The number 0 is in the set, since $\mathrm{Re}\,(w0) = 0$. If a and b are in the set, then $a+b$ is also in the set, since $\mathrm{Re}\,(w(a+b)) = \mathrm{Re}\,(wa)+\mathrm{Re}\,(wb) = 0$. If a is in the set and c is a real number, then ca is in the set, since then $\mathrm{Re}\,(wca) = c\mathrm{Re}\,(wa) = 0$. So the set is a subspace of $\mathbb{C}$. The subspace is a line in $\mathbb{C}$ with a polar angle θ such that $\theta+\varphi = k\pi/2$, where φ is the polar angle of w and k is an odd integer.

1.2.1 a. i. 2×3 ii. 2×2 iii. 3×2 iv. 3×4 v. 3×3

b. The matrices i and v can be multiplied on the right by the matrices iii, iv, v; the matrices ii and iii on the right by the matrices i and ii.

1.2.3 a. $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ b. $[6 \quad 16 \quad 2]$

1.2.5 a. True: $(AB)^\top = B^\top A^\top = B^\top A$

b. True: $(A^\top B)^\top = B^\top (A^\top)^\top = B^\top A = B^\top A^\top$

c. False: $(A^\top B)^\top = B^\top (A^\top)^\top = B^\top A \neq BA$

d. False: $(AB)^\top = B^\top A^\top \neq A^\top B^\top$

1.2.7 The matrices a and d have no transposes here. The matrices b and f are transposes of each other. The matrices c and e are transposes of each other.

1.2.9 a. $A^\top = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $B^\top = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$

b. $(AB)^\top = B^\top A^\top = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$

c. $(AB)^\top = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}^\top = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$

d. The matrix multiplication $A^\top B^\top$ is impossible.

1.2.11 The expressions b, c, d, f, g, and i make no sense.

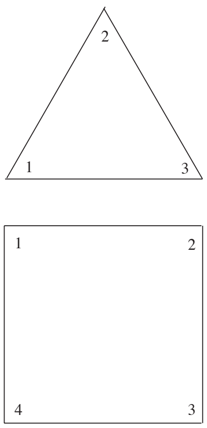
1.2.13 The trivial case is when $a = b = c = d = 0$; then obviously $ad - bc = 0$ and the matrix is not invertible. Let us suppose $d \neq 0$. (If we suppose that any other entry is nonzero, the proof would work the same way.) If $ad = bc$, then the first row is a multiple of the second: we can write $a = \frac{b}{d}c$ and $b = \frac{b}{d}d$, so the matrix is $A = \begin{bmatrix} \frac{b}{d}c & \frac{b}{d}d \\ c & d \end{bmatrix}$.

To show that A is not invertible, we need to show that there is no matrix $B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$ such that $AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. But if the upper left corner of AB is 1, then we have $\frac{b}{d}(a'c + c'd) = 1$, so the lower left corner, which is $a'c + c'd$, cannot be 0.

1.2.15
$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & a+x & az+b+y \\ 0 & 1 & c+z \\ 0 & 0 & 1 \end{bmatrix}$$

So $x = -a$, $z = -c$, and $y = ac - b$.

1.2.17 With the labeling shown in the margin, the adjacency matrices are



Labeling for Solution 1.2.17

a. $$A_T = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad A_S = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

b. $$A_T^2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad A_T^3 = \begin{bmatrix} 2 & 3 & 3 \\ 3 & 2 & 3 \\ 3 & 3 & 2 \end{bmatrix} \quad A_T^4 = \begin{bmatrix} 6 & 5 & 5 \\ 5 & 6 & 5 \\ 5 & 5 & 6 \end{bmatrix}$$

$$A_T^5 = \begin{bmatrix} 10 & 11 & 11 \\ 11 & 10 & 11 \\ 11 & 11 & 10 \end{bmatrix}$$

$$A_S^2 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix} \quad A_S^3 = \begin{bmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{bmatrix} \quad A_S^4 = \begin{bmatrix} 8 & 0 & 8 & 0 \\ 0 & 8 & 0 & 8 \\ 8 & 0 & 8 & 0 \\ 0 & 8 & 0 & 8 \end{bmatrix}$$

$$A_S^5 = \begin{bmatrix} 0 & 16 & 0 & 16 \\ 16 & 0 & 16 & 0 \\ 0 & 16 & 0 & 16 \\ 16 & 0 & 16 & 0 \end{bmatrix}$$

The diagonal entries of A^n are the number of walks we can take of length n that take us back to our starting point.

c. In a triangle, by symmetry there are only two different numbers: the number a_n of walks of length n from a vertex to itself, and the number b_n of walks of length n from a vertex to a different vertex. The recurrence relation relating these is

$$a_{n+1} = 2b_n \quad \text{and} \quad b_{n+1} = a_n + b_n.$$

These reflect that to walk from a vertex V_1 to itself in time $n+1$, at time n we must be at either V_2 or V_3, but to walk from a vertex V_1 to a different vertex V_2 in time $n+1$, at time n we must be either at V_1 or at V_3. If $|a_n - b_n| = 1$, then $a_{n+1} - b_{n+1} = |2b_2 - (a_n + b_n)| = |b_n - a_n| = 1$.

d. Color two opposite vertices of the square black and the other two white. Every move takes you from a vertex to a vertex of the opposite color. Thus if you start at time 0 on black, you will be on black at all even times, and on white at all odd times, and there will be no walks of odd length from a vertex to itself.

e. Suppose such a coloring in black and white exists; then every walk goes from black to white to black to white . . . , in particular the (B,B) and the (W,W) entries of A^n are 0 for all odd n, and the (B,W) and (W,B) entries are 0 for all even n. Moreover, since the graph is connected, for any pair of vertices there is a walk of some length m joining them, and then the corresponding entry is nonzero for m, $m+2$, $m+4, \dots$ since you can go from the point of departure to the point of arrival in time m, and then bounce back and forth between this vertex and one of its neighbors.

Conversely, suppose the entries of A^n are zero or nonzero as described, and look at the top line of A^n, where n is chosen sufficiently large so that any entry that is ever nonzero is nonzero for A^{n-1} or A^n. The entries correspond to pairs of vertices (V_1, V_i); color in white the vertices V_i for which the $(1,i)$ entry of A^n is zero, and in black those for which the $(1,i)$ entry of A^{n+1} is zero. By hypothesis, we have colored all the vertices. It remains to show that adjacent vertices have different colors. Take a path of length m from V_1 to V_i. If V_j is adjacent to V_i, then there certainly exists a path of length $m+1$ from V_1 to V_j, namely the previous path, extended by one to go from V_i to V_j. Thus V_i and V_j have opposite colors.

Solution 1.2.19, part b: There are three ways to go from V_i to an adjacent vertex V_j and back again in three steps: go, return, stay; go, stay, return; and stay, go, return. Each vertex V_i has three adjacent vertices that can play the role of V_j. That brings us to nine. The couch potato itinerary "stay, stay, stay" brings us to 10.

1.2.19

$$\text{a.}\quad B = \begin{bmatrix} 1&1&0&1&0&1&0&0\\ 1&1&1&0&0&0&1&0\\ 0&1&1&1&0&0&0&1\\ 1&0&1&1&1&0&0&0\\ 0&0&0&1&1&1&0&1\\ 1&0&0&0&1&1&1&0\\ 0&1&0&0&0&1&1&1\\ 0&0&1&0&1&0&1&1 \end{bmatrix}$$

$$\text{b.}\quad B^2 = \begin{bmatrix} 4&2&2&2&2&2&2&0\\ 2&4&2&2&0&2&2&2\\ 2&2&4&2&2&0&2&2\\ 2&2&2&4&2&2&0&2\\ 2&0&2&2&4&2&2&2\\ 2&2&0&2&2&4&2&2\\ 2&2&2&0&2&2&4&2\\ 0&2&2&2&2&2&2&4 \end{bmatrix} \qquad B^3 = \begin{bmatrix} 10&10&6&10&6&10&6&6\\ 10&10&10&6&6&6&10&6\\ 6&10&10&10&6&6&6&10\\ 10&6&10&10&10&6&6&6\\ 6&6&6&10&10&10&6&10\\ 10&6&6&6&10&10&10&6\\ 6&10&6&6&6&10&10&10\\ 6&6&10&6&10&6&10&10 \end{bmatrix}$$

The diagonal entries of B^3 correspond to the fact that there are exactly 10 loops of length 3 going from any given vertex back to itself.

Solution 1.2.21, part b: The first column of an adjacency matrix corresponds to vertex 1, the second to vertex 2, and so on, and the same for the rows. If the matrix is upper triangle, for example, $\begin{bmatrix} 1&1&1&0\\ 0&0&1&0\\ 0&0&1&1\\ 0&0&0&1 \end{bmatrix}$, you can go from vertex 1 to vertex 2, but not from 2 to 1; from 2 to 3, but not from 3 to 2, and so on: once you have gone from a lower-numbered vertex to a higher-numbered vertex, there is no going back.

1.2.21 a. The proof is the same as with unoriented walks (Proposition 1.2.23): first we state that if B_n is the $n \times n$ matrix whose i,jth entry is the number of walks of length n from V_i to V_j, then $B_1 = A^1 = A$ for the same reasons as in the proof of Proposition 1.2.23. Here again if we assume the proposition true for n, we have:

$$(B_{n+1})_{i,j} = \sum_{k=1}^{n} (B_n)_{i,k}(B_1)_{k,j} = \sum_{k=1}^{n} (A^n)_{i,k} A_{k,j} = (A^{n+1})_{i,j}.$$

So $A^n = B_n$ for all n.

b. If the adjacency matrix is upper triangular, then you can only go from a lower number vertex to a higher number vertex; if it is lower triangular, you can only go from a higher number vertex to a lower number vertex. If it is diagonal, you can never go from any vertex to any other.

1.2.23 a. Let $A = \begin{bmatrix} a & 1 & 0 \\ b & 0 & 1 \end{bmatrix}$ and let $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $AB = I$.

b. Whatever matrix one multiplies B by on the right, the top left corner of the resultant matrix will always be 0 when we need it to be 1. So the matrix B has no right inverse.

c. With A and B as in part a, write

$$I^\top = I = AB = (AB)^\top = B^\top A^\top.$$

So $A^\top$ is a right inverse for $B^\top$, so $B^\top$ has infinitely many right inverses.

Solution 1.3.1: The first question on one exam at Cornell University was "What is a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$?" Several students gave answers like "A function taking some vector $\vec{\mathbf{v}} \in \mathbb{R}^n$ and producing some vector $\vec{\mathbf{w}} \in \mathbb{R}^m$."

This is wrong! The student who gave this answer may have been thinking of matrix multiplication. But a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ need not be given by matrix multiplication: consider the mapping $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x^2 \\ y^2 \end{bmatrix}$. In any case, defining a linear transformation by saying that it is given by matrix multiplication begs the question, because it does not explain why matrix multiplication is linear.

The correct answer was given by another student in the class:

"A linear transformation

$$T : \mathbb{R}^n \to \mathbb{R}^m$$

is a mapping $\mathbb{R}^n \to \mathbb{R}^m$ such that for all $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathbb{R}^n$,

$$T(\vec{\mathbf{a}} + \vec{\mathbf{b}}) = T(\vec{\mathbf{a}}) + T(\vec{\mathbf{b}}),$$

and for all $\vec{\mathbf{a}} \in \mathbb{R}^n$ and all scalar r,

$$T(r\vec{\mathbf{a}}) = rT(\vec{\mathbf{a}})."$$

Linearity and the approximation of nonlinear mappings by linear mappings are key motifs of this book. You must know the definition, which gives you a foolproof way to check whether or not a given mapping is linear.

1.3.1 a. Every linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^2$ is given by a 2×4 matrix. For example, $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 7 \end{bmatrix}$ is a linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^2$.

b. Any row matrix 3 wide will do, for example, $[1, -1, 2]$; such a matrix takes a vector in $\mathbb{R}^3$ and gives a number.

1.3.3 A. $\mathbb{R}^4 \to \mathbb{R}^3$ B. $\mathbb{R}^2 \to \mathbb{R}^5$ C. $\mathbb{R}^4 \to \mathbb{R}^2$ D. $\mathbb{R}^4 \to \mathbb{R}$.

1.3.5 Multiply the original matrix by the vector $\vec{\mathbf{v}} = \begin{bmatrix} .25 \\ .025 \\ .025 \\ .025 \\ .025 \\ .025 \\ .025 \\ .025 \\ .025 \\ .025 \\ .025 \\ .5 \end{bmatrix}$, putting $\vec{\mathbf{v}}$ on the right.

1.3.7 It is enough to know what T gives when evaluated on the three standard basis vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. The matrix of T is $\begin{bmatrix} 3 & -1 & 0 \\ 1 & 1 & 2 \\ 2 & 3 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

1.3.9 No, T is not linear. If it were linear, the matrix would be $[T] = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$, but $[T] \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ 5 \end{bmatrix}$, which contradicts the definition of the transformation.

1.3.11 The rotation matrix is $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. This transformation takes $\vec{\mathbf{e}}_1$ to $\begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix}$, which is thus the first column of the matrix, by Theorem 1.3.4; it takes $\vec{\mathbf{e}}_2$ to $\begin{bmatrix} \cos(90^\circ-\theta)=\sin\theta \\ \sin(90^\circ-\theta)=\cos\theta \end{bmatrix}$, which is the second column.

One could also write this mapping as the rotation matrix of Example 1.3.9, applied to $-\theta$:

$$\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}.$$

1.3.13 The expressions a, e, f, and j are not well-defined compositions. For the others:

b. $C\circ B:\mathbb{R}^m\to\mathbb{R}^n$ (domain $\mathbb{R}^m$, codomain $\mathbb{R}^n$) c. $A\circ C:\mathbb{R}^k\to\mathbb{R}^m$

d. $B\circ A\circ C:\mathbb{R}^k\to\mathbb{R}^k$ g. $B\circ A:\mathbb{R}^n\to\mathbb{R}^k$

h. $A\circ C\circ B:\mathbb{R}^m\to\mathbb{R}^m$ i. $C\circ B\circ A:\mathbb{R}^n\to\mathbb{R}^n$

1.3.15 We need to show that $A(\vec{\mathbf{v}}+\vec{\mathbf{w}})=A\vec{\mathbf{v}}+A\vec{\mathbf{w}}$ and $A(c\vec{\mathbf{v}})=cA\vec{\mathbf{v}}$. By Definition 1.2.4,

Solution 1.3.15: With the dot product, introduced in the next section, the solution is simpler: Every entry in the vector $A\vec{\mathbf{v}}$ is the dot product of $\vec{\mathbf{v}}$ with one of the rows of A. The dot product is linear with respect to both of its arguments, so the mapping $A\vec{\mathbf{v}}$ is also linear with respect to $\vec{\mathbf{v}}$.

$$(A\vec{\mathbf{v}})_i=\sum_{k=1}^n a_{i,k}v_k,\quad (A\vec{\mathbf{w}})_i=\sum_{k=1}^n a_{i,k}w_k,\ \text{and}$$

$$\begin{aligned}\big(A(\vec{\mathbf{v}}+\vec{\mathbf{w}})\big)_i&=\sum_{k=1}^n a_{i,k}(v+w)_k=\sum_{k=1}^n a_{i,k}(v_k+w_k)\\&=\sum_{k=1}^n a_{i,k}v_k+\sum_{k=1}^n a_{i,k}w_k=(A\vec{\mathbf{v}})_i+(A\vec{\mathbf{w}})_i.\end{aligned}$$

Similarly, $\big(A(c\vec{\mathbf{v}})\big)_i=\sum_{k=1}^n a_{i,k}(cv)_k=\sum_{k=1}^n a_{i,k}cv_k=c\sum_{k=1}^n a_{i,k}v_k=c(A\vec{\mathbf{v}})_i.$

1.3.17

$$\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}^2=\begin{bmatrix} \cos(2\theta)^2+\sin(2\theta)^2 & \cos(2\theta)\sin(2\theta)-\sin(2\theta)\cos(2\theta) \\ \sin(2\theta)\cos(2\theta)-\cos(2\theta)\sin(2\theta) & \sin(2\theta)^2+\cos(2\theta)^2 \end{bmatrix}$$

$$\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}^2=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=I$$

1.3.19 By commutativity of matrix addition, $\frac{AB+BA}{2}=\frac{BA+AB}{2}$, so the Jordan product is commutative. By non-commutativity of matrix multiplication:

$$\begin{aligned}\frac{\frac{AB+BA}{2}C+C\frac{AB+BA}{2}}{2}&=\frac{ABC+BAC+CAB+CBA}{4}\\&\neq\frac{ABC+ACB+BCA+CBA}{4}=\frac{A\frac{BC+CB}{2}+\frac{BC+CB}{2}A}{2}\end{aligned}$$

so the Jordan product is not associative.

1.4.1 a. Numbers: $\vec{\mathbf{v}} \cdot \vec{\mathbf{w}}$, $|\vec{\mathbf{v}}|$, $|A|$, and $\det A$. (If A consists of a single row, then $A\vec{\mathbf{v}}$ is also a number.)

Vectors: $\vec{\mathbf{v}} \times \vec{\mathbf{w}}$ and $A\vec{\mathbf{v}}$ (unless A consists of a single row).

So far we have defined only determinants of 2×2 and 3×3 matrices; in Section 4.8 we will define the determinant in general.

b. In the expression $\vec{\mathbf{v}} \times \vec{\mathbf{w}}$, the vectors must each have three entries. In the expression $\det A$, the matrix A must be square.

1.4.3 To normalize a vector, divide it by its length. This gives:

a. $\frac{1}{\sqrt{17}}\begin{bmatrix}0\\1\\4\end{bmatrix}$ b. $\frac{1}{\sqrt{58}}\begin{bmatrix}-3\\7\end{bmatrix}$ c. $\frac{1}{\sqrt{31}}\begin{bmatrix}\sqrt{2}\\-2\\-5\end{bmatrix}$

1.4.5 a. $\cos(\theta) = \frac{1}{1\times\sqrt{3}}\left(\begin{bmatrix}1\\0\\0\end{bmatrix} \cdot \begin{bmatrix}1\\1\\1\end{bmatrix}\right) = \frac{1}{\sqrt{3}}$, so

$$\theta = \arccos\left(\frac{1}{\sqrt{3}}\right) \approx .95532.$$

b. $\cos(\theta) = 0$, so $\theta = \pi/2$.

1.4.7 a. $\det = 1$; the inverse is $\begin{bmatrix}0 & 1\\-1 & 2\end{bmatrix}$

b. $\det = 0$; no inverse

c. $\det = ad$; if $a, d \neq 0$, the inverse is $\frac{1}{ad}\begin{bmatrix}d & -b\\0 & a\end{bmatrix}$.

d. $\det = 0$; no inverse

1.4.9 a. $\begin{bmatrix}-6yz\\3xz\\5xy\end{bmatrix}$ b. $\begin{bmatrix}6\\7\\-4\end{bmatrix}$ c. $\begin{bmatrix}-2\\-22\\3\end{bmatrix}$

1.4.11 a. True by Theorem 1.4.5, because $\vec{\mathbf{w}} = -2\vec{\mathbf{v}}$.

Solution 1.4.11, part c: Our answer depended on the vectors chosen. In general,

$$\det[\vec{\mathbf{a}}, \vec{\mathbf{b}}, \vec{\mathbf{c}}] = -\det[\vec{\mathbf{a}}, \vec{\mathbf{c}}, \vec{\mathbf{b}}];$$

the result here is true only because both sides are 0, since $\vec{\mathbf{v}}$, $\vec{\mathbf{w}}$ are linearly dependent.

b. False; $\vec{\mathbf{u}} \cdot (\vec{\mathbf{v}} \times \vec{\mathbf{w}})$ is a number; $|\vec{\mathbf{u}}|(\vec{\mathbf{v}} \times \vec{\mathbf{w}})$ is a number times a vector, i.e., a vector.

c. True: since $\vec{\mathbf{w}} = -2\vec{\mathbf{v}}$, we have $\det[\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}}] = 0$ and $\det[\vec{\mathbf{u}}, \vec{\mathbf{w}}, \vec{\mathbf{v}}] = 0$.

d. False, since $\vec{\mathbf{u}}$ is not necessarily (in fact almost surely not) a multiple of $\vec{\mathbf{w}}$; the correct statement is $|\vec{\mathbf{u}} \cdot \vec{\mathbf{w}}| \leq |\vec{\mathbf{u}}||\vec{\mathbf{w}}|$.

e. True. f. True.

1.4.13 $\begin{bmatrix}xa\\xb\\xc\end{bmatrix} \times \begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}xbc - xbc\\-(xac - xac)\\xab - xab\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$.

Solutions 1.4.13 and 1.4.15: Of course, the vectors must be in $\mathbb{R}^3$ for the cross product to be defined.

1.4.15 $\begin{bmatrix}a\\b\\c\end{bmatrix} \times \begin{bmatrix}d\\e\\f\end{bmatrix} = \begin{bmatrix}bf - ce\\cd - af\\ae - bd\end{bmatrix} = -\begin{bmatrix}d\\e\\f\end{bmatrix} \times \begin{bmatrix}a\\b\\c\end{bmatrix} = -\begin{bmatrix}ec - bf\\af - cd\\bd - ae\end{bmatrix}$

1.4.17 a. It is the line of equation $\begin{bmatrix} x \\ y \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 2x - y = 0$.

b. It is the line of equation $\begin{bmatrix} x-2 \\ y-3 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -4 \end{bmatrix} = 2x - 4 - 4y + 12 = 0$, which you can rewrite as $2x - 4y + 8 = 0$.

1.4.19 a. The length of $\vec{\mathbf{v}}_n$ is $|\vec{\mathbf{v}}_n| = \sqrt{1 + \cdots + 1} = \sqrt{n}$.

b. The angle is $\arccos \dfrac{1}{\sqrt{n}}$, which tends to $\pi/2$ as $n \to \infty$.

Solution 1.4.19, part b: We find it surprising that the diagonal vector $\vec{\mathbf{v}}_n$ is almost orthogonal to all the standard basis vectors when n is large.

1.4.21 a. $|A| = \sqrt{1+1+4} = \sqrt{6}$; $\quad |B| = \sqrt{5}$; $\quad |\vec{\mathbf{c}}| = \sqrt{10}$

b. $|AB| = \left| \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix} \right| = \sqrt{12} \ \le \ \sqrt{30} = |A||B|$

$|A\vec{\mathbf{c}}| = \sqrt{50} \ \le \ |A||\vec{\mathbf{c}}| = \sqrt{60}$; $\quad |B\vec{\mathbf{c}}| = \sqrt{13} \ \le \ |B||\vec{\mathbf{c}}| = \sqrt{50}$

1.4.23 a. The length is $|\vec{\mathbf{w}}_n| = \sqrt{1 + 4 + \cdots + n^2} = \sqrt{\dfrac{n^3}{3} + \dfrac{n^2}{2} + \dfrac{n}{6}}$.

b. The angle $\alpha_{n,k}$ is $\arccos \dfrac{k}{\sqrt{\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}}}$.

c. In all three cases, the limit is $\pi/2$. Clearly $\lim_{n\to\infty} \alpha_{n,k} = \pi/2$, since the cosine tends to 0.

The limit of $\alpha_{n,n}$ is $\lim_{n\to\infty} \arccos \frac{n}{\sqrt{\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}}} = \arccos 0 = \pi/2$.

The limit of $\alpha_{n,[n/2]}$ is also $\pi/2$, since it is the arccos of

$$\frac{[n/2]}{\sqrt{\frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}}}, \quad \text{which tends to 0 as } n \to \infty.$$

Solution 1.4.23, part a: This uses the fact that

$$1^2 + 2^2 + \cdots + n^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}.$$

1.4.25

$$\begin{aligned}\left(\sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}\right)^2 - \left(x_1 y_1 + x_2 y_2\right)^2 &= (x_2 y_1)^2 + (x_1 y_2)^2 - 2x_1 x_2 y_1 y_2 \\ &= (x_1 y_2 - x_2 y_1)^2 \ge 0,\end{aligned}$$

so $(x_1 y_1 + x_2 y_2)^2 \le \left(\sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}\right)^2$, so

$$|x_1 y_1 + x_2 y_2| \le \sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}.$$

1.4.27 a. To show that the transformation is linear, we need to show that

$$T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}}) + T_{\vec{\mathbf{a}}}(\vec{\mathbf{w}}) \quad \text{and} \quad \alpha T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}}) = T_{\vec{\mathbf{a}}}(\alpha\vec{\mathbf{v}}).$$

For the first,

$$T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = \vec{\mathbf{v}} + \vec{\mathbf{w}} - 2\big(\vec{\mathbf{a}} \cdot (\vec{\mathbf{v}} + \vec{\mathbf{w}})\big)\vec{\mathbf{a}} = \vec{\mathbf{v}} + \vec{\mathbf{w}} - 2(\vec{\mathbf{a}} \cdot \vec{\mathbf{v}} + \vec{\mathbf{a}} \cdot \vec{\mathbf{w}})\vec{\mathbf{a}} = T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}}) + T_{\vec{\mathbf{a}}}(\vec{\mathbf{w}}).$$

For the second,

$$\alpha T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}}) = \alpha\vec{\mathbf{v}} - 2\alpha(\vec{\mathbf{a}} \cdot \vec{\mathbf{v}})\vec{\mathbf{a}} = \alpha\vec{\mathbf{v}} - 2(\vec{\mathbf{a}} \cdot \alpha\vec{\mathbf{v}})\vec{\mathbf{a}} = T_{\vec{\mathbf{a}}}(\alpha\vec{\mathbf{v}}).$$

b. We have $T_{\vec{\mathbf{a}}}(\vec{\mathbf{a}}) = -\vec{\mathbf{a}}$, since $\vec{\mathbf{a}} \cdot \vec{\mathbf{a}} = a^2 + b^2 + c^2 = 1$:

$$T_{\vec{\mathbf{a}}}(\vec{\mathbf{a}}) = \vec{\mathbf{a}} - 2(\vec{\mathbf{a}} \cdot \vec{\mathbf{a}})\vec{\mathbf{a}} = \vec{\mathbf{a}} - 2\vec{\mathbf{a}} = -\vec{\mathbf{a}}.$$

If $\vec{\mathbf{v}}$ is orthogonal to $\vec{\mathbf{a}}$, then $T_{\vec{\mathbf{a}}}(\vec{\mathbf{v}}) = \vec{\mathbf{v}}$, since in that case $\vec{\mathbf{a}} \cdot \vec{\mathbf{v}} = 0$. Thus $T_{\vec{\mathbf{a}}}$ is reflection in the plane that goes through the origin and is perpendicular to $\vec{\mathbf{a}}$.

The transformation $T_{\vec{\mathbf{a}}}$ is the 3-dimensional version of the transformation shown in Figure 1.3.4.

c. The matrix of $T_{\vec{\mathbf{a}}}$ is

$$M = [T_{\vec{\mathbf{a}}}(\vec{\mathbf{e}}_1), T_{\vec{\mathbf{a}}}(\vec{\mathbf{e}}_2), T_{\vec{\mathbf{a}}}(\vec{\mathbf{e}}_3)] = \begin{bmatrix} 1-2a^2 & -2ab & -2ac \\ -2ab & 1-2b^2 & -2bc \\ -2ac & -2bc & 1-2c^2 \end{bmatrix}$$

Squaring the matrix gives the 3×3 identity matrix: if you reflect a vector, then reflect it again, you are back to where you started.

Solution 1.5.1: Note that part c is the same as part a.

1.5.1 a. The set $\{\, x \in \mathbb{R} \mid 0 < x \leq 1 \,\}$ is neither open nor closed: the point 1 is in the set, but $1 + \epsilon$ is not for every ϵ, showing it isn't open, and 0 is not but $0 + \epsilon$ is for every $\epsilon > 0$, showing that the complement is also not open, so the set is not closed.

b. open c. neither d. closed e. closed f. neither

g. Both. That the empty set ϕ is closed is obvious. Showing that it is open is an "eleven-legged alligator" argument (see Section 0.2). Is it true that for every point of the empty set, there exists $\epsilon > 0$ such that … ? Yes, because there are no points in the empty set.

1.5.3 a. Suppose $A_i, i \in I$ is some collection (probably infinite) of open sets. If $\mathbf{x} \in \bigcup_{i \in I} A_i$, then $\mathbf{x} \in A_j$ for some j, and since A_j is open, there exists $\epsilon > 0$ such that $B_\epsilon(\mathbf{x}) \subset A_j$. But then $B_\epsilon(\mathbf{x}) \subset \bigcup_{i \in I} A_i$.

Solution 1.5.3, part b: There is a smallest ϵ_i, because there are finitely many of them, and it is positive. If there were infinitely many, then there would be a greatest lower bound, but it could be 0.

Part c: In fact, every closed set is a countable intersection of open sets.

b. If $A_1, \dots, A_j$ are open and $\mathbf{x} \in \cap_{i=1}^k A_i$, then there exist $\epsilon_1, \dots, \epsilon_k > 0$ such that $B_{\epsilon_i}(\mathbf{x}) \subset A_i$, for $i = 1, \dots, k$. Set ϵ to be the smallest of $\epsilon_1, \dots, \epsilon_k$. Then $B_\epsilon(\mathbf{x}) \subset B_{\epsilon_i}(\mathbf{x}) \subset A_i$.

c. The infinite intersection of open sets $(-1/n, 1/n)$, for $n = 1, 2, \dots$, is not open; as $n \to \infty$, $-1/n \to 0$ and $1/n \to 0$; the set $\{0\}$ is not open.

1.5.5 a. This set is open. Indeed, if you choose $\begin{pmatrix} x \\ y \end{pmatrix}$ in your set, then $1 < \sqrt{x^2 + y^2} < \sqrt{2}$. Set

$$r = \min\left\{ \sqrt{x^2+y^2} - 1, \sqrt{2} - \sqrt{x^2+y^2} \right\} > 0.$$

Then the ball of radius r around $\begin{pmatrix} x \\ y \end{pmatrix}$ is contained in the set, since if $\begin{pmatrix} u \\ v \end{pmatrix}$ is in that ball, then, by the triangle inequality,

Equation 1 uses the familiar form of the triangle inequality: if $\mathbf{a} = \mathbf{b} + \mathbf{c}$, then

$$|\mathbf{a}| \leq |\mathbf{b}| + |\mathbf{c}|.$$

Equation 2 uses the variant

$$|\mathbf{a}| \geq \Big| |\mathbf{b}| - |\mathbf{c}| \Big|.$$

$$\left| \begin{bmatrix} u \\ v \end{bmatrix} \right| \leq \left| \begin{bmatrix} u-x \\ v-y \end{bmatrix} \right| + \left| \begin{bmatrix} x \\ y \end{bmatrix} \right| < r + \left| \begin{bmatrix} x \\ y \end{bmatrix} \right| \leq \sqrt{2} \tag{1}$$

$$\left| \begin{bmatrix} u \\ v \end{bmatrix} \right| \geq \left| \begin{bmatrix} x \\ y \end{bmatrix} \right| - \left| \begin{bmatrix} u-x \\ v-y \end{bmatrix} \right| > \left| \begin{bmatrix} x \\ y \end{bmatrix} \right| - r \geq 1. \tag{2}$$

b. The locus $xy \neq 0$ is also open. It is the complement of the two axes, so that if $\begin{pmatrix} x \\ y \end{pmatrix}$ is in the set, then $r = \min\{|x|, |y|\} > 0$, and the ball B of radius r around $\begin{pmatrix} x \\ y \end{pmatrix}$ is contained in the set. Indeed, if $\begin{pmatrix} u \\ v \end{pmatrix}$ is B, then $|u| = |x + u - x| > |x| - |u - x| > |x| - r \geq 0$, so u is not 0, and neither is v, by the same argument.

c. This time our set is the x-axis, and it is closed. We will use the criterion that a set is closed if the limit of a convergent sequence of elements of the set is in the set (Proposition 1.5.17). If $n \mapsto \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ is a sequence in the set, and converges to $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, then all $y_n = 0$, so $y_0 = \lim_{n\to\infty} y_n = 0$, and the limit is also in the set.

d. The rational numbers are neither open nor closed. Any rational number x is the limit of the numbers $x + \sqrt{2}/n$, which are all irrational, so the irrationals aren't closed. Any irrational number is the limit of the finite decimals used to write it, which are all rational, so the rationals aren't closed either.

Many students have found Exercise 1.5.5 difficult, even though they also thought it was obvious, but didn't know how to say it. If this applies to you, you should check carefully where we used the triangle inequality, and how. Almost everything concerning inequalities requires the triangle inequality.

1.5.7 a. The natural domain is $\mathbb{R}^2$ minus the union of the two axes; it is open.

b. The natural domain is that part of $\mathbb{R}^2$ where $x^2 > y$ (i.e., the area "outside" the parabola of equation $y = x^2$). It is open since its "fence" x^2 belongs to its neighbor.

c. The natural domain of $\ln \ln x$ is $\{x | x > 1\}$, since we must have $\ln x > 0$. This domain is open.

d. The natural domain of arcsin is $[-1, 1]$. Thus the natural domain of $\arcsin \frac{3}{x^2+y^2}$ is $\mathbb{R}^2$ minus the open disc $x^2 + y^2 < 3$. Since this domain is the complement of an open disc it is closed (and not open, since it isn't $\mathbb{R}^2$ or the empty set).

e. The natural domain is all of $\mathbb{R}^2$, which is open.

f. The natural domain is $\mathbb{R}^3$ minus the union of the three coordinate planes of equation $x = 0$, $y = 0$, $z = 0$; it is open.

1.5.9 For any $n > 0$ we have $\left|\sum_{i=1}^{n} \mathbf{x}_i\right| \leq \sum_{i=1}^{n} |\mathbf{x}_i|$ by the triangle inequality (Theorem 1.4.9). Because $\sum_{i=1}^{\infty} \mathbf{x}_i$ converges, $\sum_{i=1}^{n} \mathbf{x}_i$ converges as $n \to \infty$. So :

$$\left|\sum_{i=1}^{\infty} \mathbf{x}_i\right| \leq \sum_{i=1}^{\infty} |\mathbf{x}_i|.$$

1.5.11 a. First let us see that

$$\Big((\forall \epsilon > 0)(\exists N)(n > N) \implies |\mathbf{a}_n - \mathbf{a}| < \varphi(\epsilon)\Big) \implies \big(\mathbf{a}_n \text{ converges to } \mathbf{a}\big).$$

Choose $\eta > 0$. Since $\lim_{t\to 0}\varphi(t) = 0$, there exists $\delta > 0$ such that when $0 < t \le \delta$ we have $\varphi(t) < \eta$. Our hypothesis guarantees that there exists N such that when $n > N$, then $|\mathbf{a}_n - \mathbf{a}| \le \varphi(\delta) = \eta$.

Now for the converse:

$$\big(\mathbf{a}_n \text{ converges to } \mathbf{a}\big) \implies \Big((\forall \epsilon > 0)(\exists N)(n > N) \implies |\mathbf{a}_n - \mathbf{a}| < \varphi(\epsilon)\Big).$$

For any $\epsilon > 0$, we also have $\varphi(\epsilon) > 0$, so there exists N such that

$$n > N \implies |\mathbf{a}_n - \mathbf{a}| < \varphi(\epsilon).$$

b. The analogous statement for limits of functions is:

Let $\varphi : [0,\infty) \to [0,\infty)$ be a function such that $\lim_{t\to 0}\varphi(t) = 0$. Let $U \subset \mathbb{R}^n$, $f : U \to \mathbb{R}^m$, and $\mathbf{x}_0 \in \overline{U}$. Then $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{a}$ if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that when $\mathbf{x} \in U$ and $|\mathbf{x} - \mathbf{x}_0| < \delta$, we have $|f(\mathbf{x}) - \mathbf{a}| < \varphi(\epsilon)$.

1.5.13 Choose a point $\mathbf{a} \in \mathbb{R}^n - C$. Suppose that the ball $B_{1/n}(\mathbf{a})$ of radius $1/n$ around $\mathbf{a}$ satisfies $B_{1/n}(\mathbf{a}) \cap C \neq \phi$ for every n. Choose $\mathbf{a}_n$ in $B_{1/n}(\mathbf{a}) \cap C$. Then the sequence $n \mapsto \mathbf{a}_n$ converges to $\mathbf{a}$, since for any $\epsilon > 0$ we can find N such that for $n > N$ we have $1/n < \epsilon$, so for $n > N$ we have $|\mathbf{a} - \mathbf{a}_n| < 1/n < \epsilon$. Then our hypothesis implies that $\mathbf{a} \in C$, a contradiction. Thus there exists N such that $B_{1/N}(\mathbf{a}) \cap C = \phi$. This shows that $\mathbb{R}^n - C$ is open, so C is closed.

Solution 1.5.15: Here, x plays the role of the alligators in Section 0.2, and $x \ge 0$ satisfying $|-2-x| < \delta$ plays the role of eleven-legged alligators; the conclusion $|\sqrt{x} - 5| < \epsilon$ (i.e., that 5 is the limit) is the conclusion "are orange with blue spots" and the conclusion $|\sqrt{x} - 3| < \epsilon$ (3 is the limit) is the conclusion "are black with white stripes".

1.5.15 Both statements are true. To show that the first is true, we say: for every $\epsilon > 0$, there exists $\delta > 0$ such that for all x satisfying $x \ge 0$ and $|-2-x| < \delta$, then $|\sqrt{x} - 5| < \epsilon$. For any $\epsilon > 0$, choose $\delta = 1$. Then there is no $x \ge 0$ satisfying $|-2-x| < \delta$. So for those nonexistent x satisfying $|-2-x| < 1$, it is true that $|\sqrt{x} - 5| < \epsilon$. By the same argument the second statement is true.

1.5.17 Set $\mathbf{a} = \lim_{m\to\infty}\mathbf{a}_m$, $\mathbf{b} = \lim_{m\to\infty}\mathbf{b}_m$ and $c = \lim_{m\to\infty} c_m$.

1. Choose $\epsilon > 0$ and find M_1 and M_2 such that if $m \ge M_1$ then we have $|\mathbf{a}_m - \mathbf{a}| \le \epsilon/2$, and if $m \ge M_2$ then $|\mathbf{b}_m - \mathbf{b}| \le \epsilon/2$. Set

$$M = \max(M_1, M_2).$$

If $m > M$, we have

$$|\mathbf{a}_m + \mathbf{b}_m - \mathbf{a} - \mathbf{b}| \le |\mathbf{a}_m - \mathbf{a}| + |\mathbf{b}_m - \mathbf{b}| \le \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

So the sequence $m \mapsto (\mathbf{a}_m + \mathbf{b}_m)$ converges to $\mathbf{a} + \mathbf{b}$.

2. Choose $\epsilon > 0$. Find M_1 such that if

$$m \ge M_1, \quad \text{then} \quad |\mathbf{a}_m - \mathbf{a}| \le \frac{1}{2}\inf\left(\frac{\epsilon}{|c|}, \epsilon\right).$$

The inf is there to guard against the possibility that $|c| = 0$. In particular, if $m \ge M_1$, then $|\mathbf{a}_m| \le |\mathbf{a}| + \epsilon$. Next find M_2 such that if

$$m \ge M_2, \quad \text{then} \quad |c_m - c| \le \frac{\epsilon}{2(|\mathbf{a}| + \epsilon)}.$$

If $M = \max(M_1, M_2)$ and $m \geq M$, then

$$\begin{aligned} |c_m \mathbf{a}_m - c\mathbf{a}| &= |c(\mathbf{a}_m - \mathbf{a}) + (c_m - c)\mathbf{a}_m| \\ &\leq |c(\mathbf{a}_m - \mathbf{a})| + |(c_m - c)\mathbf{a}_m| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon, \end{aligned}$$

so the sequence $m \mapsto c_m \mathbf{a}_m$ converges and the limit is $c\mathbf{a}$.

3. We can either repeat the argument above, or use parts 1 and 2 as follows:

$$\begin{aligned} \lim_{m\to\infty} \mathbf{a}_m \cdot \mathbf{b}_m &= \lim_{m\to\infty} \sum_{i=1}^{n} a_{m,i} b_{m,i} = \sum_{i=1}^{n} \lim_{m\to\infty} (a_{m,i} b_{m,i}) \\ &= \sum_{i=1}^{n} \left(\lim_{m\to\infty} a_{m,i} \right) \left(\lim_{m\to\infty} b_{m,i} \right) = \sum_{i=1}^{n} a_i b_i = \mathbf{a} \cdot \mathbf{b}. \end{aligned}$$

4. Find C such that $|\mathbf{a}_m| \leq C$ for all m; saying that $\mathbf{a}_m$ is bounded means exactly that such a C exists. Choose $\epsilon > 0$, and find M such that when $m > M$, then $|c_m| < \epsilon/C$ (this is possible since the c_m converge to 0). Then when $m > M$ we have

$$|c_m \mathbf{a}_m| = |c_m||\mathbf{a}_m| \leq \frac{\epsilon}{C} C = \epsilon.$$

1.5.19 a. Suppose $I - A$ is invertible, and write

$$I - A + C = (I - A) + C(I - A)^{-1}(I - A) = \big(I + C(I - A)^{-1}\big)(I - A),$$

so

$$\begin{aligned} (I - A + C)^{-1} &= (I - A)^{-1} \Big(I + C(I - A)^{-1} \Big)^{-1} \\ &= (I - A)^{-1} \Bigg(\underbrace{I - (C(I - A)^{-1}) + (C(I - A)^{-1})^2 - (C(I - A)^{-1})^3 + \cdots}_{\text{geometric series if } |C(I-A)^{-1}|<1} \Bigg) \end{aligned}$$

so long as the series is convergent. By Proposition 1.5.38, this occurs if

$$|C(I - A)^{-1}| < 1, \quad \text{in particular if} \quad |C| < \frac{1}{|(I - A)^{-1}|}.$$

Thus every point of U is the center of a ball contained in U.

For the second part of the question, the matrices

$$C_n = \begin{bmatrix} 1 - 1/n & 0 \\ 0 & 1 - 1/n \end{bmatrix}, \quad n = 1, 2, \ldots$$

converge to I, and C_n is in U, since $I - C_n = \begin{bmatrix} 1/n & 0 \\ 0 & 1/n \end{bmatrix}$ is invertible.

b. Simply factor: $(A + I)(A - I) = A^2 + A - A - I = A^2 - I$, so

$$(A^2 - I)(A - I)^{-1} = (A + I)(A - I)(A - I)^{-1} = A + I,$$

which converges to $2I$ as $A \to I$.

c. Showing that V is open is very much like showing that U is open (part a). Suppose $B-A$ is invertible, and write

$$B - A + C = (I + C(B-A)^{-1})(B-A), \quad \text{so}$$

$$\begin{aligned}(B-A+C)^{-1} &= (B-A)^{-1}\big(I + C(B-A)^{-1}\big)^{-1}\\ &= (B-A)^{-1}\Big(I - (C(B-A)^{-1}) + (C(B-A)^{-1})^2 - (C(B-A)^{-1})^3 + \cdots\Big),\end{aligned}$$

so long as the series is convergent. This will happen if

$$|C(B-A)^{-1}| < 1, \text{ in particular, if } |C| < \frac{1}{|(B-A)^{-1}|}.$$

Thus every point of V is the center of a ball contained in V. Again, the matrices

$$\begin{bmatrix} 1+1/n & 0 \\ 0 & -1+1/n \end{bmatrix}, \quad n = 1, 2, \dots$$

do the trick.

d. This time the limit does not exist. Note that you cannot factor $A^2 - B^2 = (A+B)(A-B)$ if A and B do not commute.

Part d: You may wonder how we came by the matrices A_n; we observed that

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

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix},$$

so these matrices do not commute.

First set

$$A_n = \begin{bmatrix} 1/n+1 & 1/n \\ 0 & -1+1/n \end{bmatrix}.$$

Then

$$A_n^2 - B^2 = \begin{bmatrix} 2/n + 1/n^2 & 2/n^2 \\ 0 & -2/n+1/n^2 \end{bmatrix} \text{ and } (A-B)^{-1} = \begin{bmatrix} n & -n \\ 0 & n \end{bmatrix}.$$

Thus we find

$$(A_n^2 - B^2)(A_n - B)^{-1} = \begin{bmatrix} 2+1/n & -2+1/n \\ 0 & -2+1/n \end{bmatrix} \to \begin{bmatrix} 2 & -2 \\ 0 & -2 \end{bmatrix}$$

as $n \to \infty$.

Do the same computation with $A'_n = \begin{bmatrix} 1/n+1 & 0 \\ 0 & -1+1/n \end{bmatrix}$. This time we find

$$({A'}_n^2 - B^2)({A'}_n - B)^{-1} = \begin{bmatrix} 2+1/n & 0 \\ 0 & -2+1/n \end{bmatrix} \to \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} = 2B$$

as $n \to \infty$.

Since both sequences $n \mapsto A_n$ and $n \mapsto A'_n$ converge to B, this shows that there is no limit.

1.5.21 a. This is a quotient of continuous functions where the denominator does not vanish at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, so it is continuous at the origin.

b. If we approach the origin along the x-axis, $f = 1$, and if we approach the origin along the y-axis, $f = |y|^{\frac{2}{3}}$ goes to 0, so f is not continuous at the origin. There is no way of choosing a value of $f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ that will make f continuous at the origin.

Part c of Solution 1.5.21 uses the following statement from one-variable calculus :

$$\lim_{u\to 0} u \ln |u| = 0.$$

This can be proved by applying l'Hôpital's rule to $\dfrac{\ln |u|}{1/u}$.

c. When $0 < x^2 + 2y^2 < 1$,

$$0 > (x^2 + y^2) \ln(x^2 + 2y^2) \geq (x^2 + y^2) \ln(x^2 + y^2),$$

which tends to 0 using the equation in the margin. So if we choose $f\begin{pmatrix}0\\0\end{pmatrix} = 0$, then f is continuous.

d. The function is not continuous near the origin. Since $\ln 0$ is undefined, the diagonal $x + y = 0$ is not part of the function's domain of definition. However, the function is defined at points arbitrarily close to that line, e.g., the point $\begin{pmatrix}x\\-x+e^{-1/x^3}\end{pmatrix}$. At this point we have

$$\left(x^2 + \left(-x + e^{-1/x^3}\right)^2\right) \ln\left|x - x + e^{-1/x^3}\right| \geq x^2 \left|\frac{1}{x^3}\right| = \frac{1}{|x|},$$

which tends to infinity as x tends to 0. But if we approach the origin along the x-axis (for instance), the function is $x^2 \ln |x|$, which tends to 0 as x tends to 0.

1.5.23 a. To say that $\lim_{B\to A}(A-B)^{-1}(A^2-B^2)$ exists means that there is a matrix C such that for all $\epsilon > 0$, there exists $\delta > 0$, such that when $|B - A| < \delta$ and $B - A$ is invertible, then

$$\left|(B - A)^{-1}(B^2 - A^2) - C\right| < \epsilon.$$

b. We will show that the limit exists, and is $\begin{bmatrix}2&0\\0&2\end{bmatrix} = 2I$. Write $B = I + H$, with H invertible, and choose $\epsilon > 0$. We need to show that there exists $\delta > 0$ such that if $|H| < \delta$, then

$$\left|(I + H - I)^{-1}\left((I + H)^2 - I^2\right) - 2I\right| < \epsilon. \tag{1}$$

Indeed,

$$\begin{aligned}\left|(I + H - I)^{-1}\left((I + H)^2 - I^2\right) - 2I\right| &= \left|H^{-1}\left(I^2 + IH + HI + H^2 - I^2\right) - 2I\right| \\ &= \left|H^{-1}\left(2H + H^2\right) - 2I\right| = |H|.\end{aligned}$$

So if you set $\delta = \epsilon$, and $|H| \leq \delta$, then equation 1 is satisfied.

c. We will show that the limit does not exist. In this case, we find

$$\begin{aligned}(A + H - A)^{-1}(A + H)^2 - A^2 &= H^{-1}(I^2 + AH + HA + H^2 - I^2) \\ &= H^{-1}(AH + HA + H^2) = A + H^{-1}AH + H^2.\end{aligned}$$

If the limit exists, it must be $2A$: choose $H = \epsilon I$ so that $H^{-1} = \epsilon^{-1}I$; then

$$A + H^{-1}AH + H^2 = 2A + \epsilon I$$

is close to $2A$.

But if you choose $H = \epsilon \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, you will find that

$$H^{-1}AH = \begin{bmatrix} 1/\epsilon & 0 \\ 0 & -1/\epsilon \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \epsilon & 0 \\ 0 & \epsilon \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = -A.$$

So with this H we have

$$A + H^{-1}AH + H^2 = A - A + \epsilon H,$$

which is close to the zero matrix.

1.6.1 Let B be a set contained in a ball of radius R centered at a point $\mathbf{a}$. Then it is also contained in a ball of radius $R + |\mathbf{a}|$ centered at the origin; thus it is bounded.

1.6.3 The polynomial $p(z) = 1 + x^2y^2$ has no roots because 1 plus something positive cannot be 0. This does not contradict the fundamental theorem of algebra because although p is a polynomial in the real variables x and y, it is not a polynomial in the complex variable z: it is a polynomial in z and $\bar{z}$. It is possible to write $p(z) = 1 + x^2y^2$ in terms of z and $\bar{z}$. You can use

$$x = \frac{z + \overline{z}}{2} \quad \text{and} \quad y = \frac{z - \overline{z}}{2i},$$

and find

$$p(z) = 1 + \frac{z^4 - 2|z|^4 + \overline{z}^4}{-16} \tag{1}$$

but you simply cannot get rid of the $\overline{z}$.

1.6.5 a. Suppose $|z| > 3$. Then (by the triangle inequality)

Solution 1.6.5: How did we land on the number 3? We started the computation, until we got to the expression $|z|^2 - 7$, which we needed to be positive. The number 3 works, and 2 does not; 2.7 works too.

$$\begin{aligned} |z|^6 - |q(z)| &\geq |z|^6 - (4|z|^4 + |z| + 2) \geq |z|^6 - (4|z|^4 + |z|^4 + 2|z|^4) \\ &= |z|^4(|z|^2 - 7) \geq (9 - 7) \cdot 3^4 = 162. \end{aligned}$$

b. Since $p(0) = 2$, but when $|z| > 3$ we have $|p(z)| \geq |z|^6 - |q(z)| \geq 162$, the minimum value of $|p|$ on the disc of radius $R_1 = 3$ around the origin must be the absolute minimum value of $|p|$. Notice there must be a point at which this minimum value is achieved, since $|p|$ is a continuous function on the closed and bounded set $|z| \leq 3$ of $\mathbb{C}$.

1.6.7 Consider the function $g(x) = f(x) - mx$. This is a continuous function on the closed and bounded set $[a, b]$, so it has a minimum at some point $c \in [a, b]$. Let us see that $c \neq a$ and $c \neq b$. Since $g'(a) = f'(a) - m < 0$, we have

$$\lim_{h \to 0} \frac{g(a + h) - g(a)}{h} < 0.$$

Thus for every $\epsilon > 0$, there exists $\delta > 0$ such that $0 < |h| < \delta$ implies

$$\left| \frac{g(a+h) - g(a)}{h} - g'(a) \right| < \epsilon.$$

Choose $\epsilon = |g'(a)|/2$, and find a corresponding $\delta > 0$. Then for h satisfying $0 < h < \delta$ the inequality

$$\left| \frac{g(a+h) - g(a)}{h} - g'(a) \right| < \frac{|g'(a)|}{2}$$

implies that

$$\frac{g(a+h) - g(a)}{h} < \frac{g'(a)}{2} < 0$$

and since $h > 0$ we have $g(a+h) < g(a)$, so a is not the minimum of g.

Similarly, b is not the minimum:

$$\lim_{h \to 0} \frac{g(b+h) - g(b)}{h} = g'(b) - m > 0.$$

Express this again in terms of ϵ's and δ's, choose $\epsilon = g'(b)/2$, and choose h satisfying $-\delta < h < 0$. As above, we have

$$\frac{g(b+h) - g(b)}{h} > \frac{g'(b)}{2} > 0,$$

and since $h < 0$, this implies $g(b+h) < g(b)$.

So $c \in (a, b)$, and in particular c is a minimum on (a, b), so $g'(c) = f'(c) - m = 0$ by Proposition 1.6.12.

Solution 1.6.7: Although our function g is differentiable on a neighborhood of a and b, we cannot apply Proposition 1.6.12 if the minimum occurs at one of those points, since c would not be a maximum on a neighborhood of the point.

1.6.9 "Between 0 and a" means that if you plot a as a point in $\mathbb{R}^2$ in the usual way (real part of a on the x-axis, imaginary part on the y-axis), then $a + bu^j$ lies on the line segment connecting the origin and the point a; see the figure in the margin. For this to happen, bu^j must be a negative real multiple of a, and we must have $|bu^j| < |a|$. Write

$$\begin{aligned} a &= r_1(\cos\omega_1 + i\sin\omega_1) \\ b &= r_2(\cos\omega_2 + i\sin\omega_2) \\ u &= p(\cos\theta + i\sin\theta). \end{aligned}$$

Then

$$a + bu^j = r_1(\cos\omega_1 + i\sin\omega_1) + r_2p^j\big(\cos(\omega_2 + j\theta) + i\sin(\omega_2 + j\theta)\big).$$

Then bu^j will point in the opposite direction from a if

$$\omega_2 + j\theta = \omega_1 + \pi + 2k\pi \quad \text{for some } k, \text{ i.e., } \theta = \frac{1}{j}(\omega_1 - \omega_2 + \pi + 2k\pi),$$

and we find j distinct such angles by taking $k = 0, 1, \dots, j-1$.

The condition $|bu^j| < |a|$ becomes $r_2p^j < r_1$, so we can take

$$0 < p < (r_1/r_2)^{1/j} \stackrel{\text{def}}{=} p_0.$$

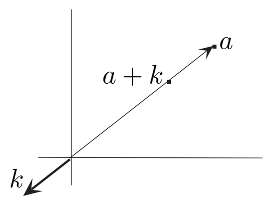


FIGURE FOR SOLUTION 1.6.9

A first error to avoid is writing "$a + bu^j$ is between 0 and a" as

"$0 < a + bu^j < a$."

Remember that a, b, and u are complex numbers so that writing that sort of inequality doesn't make sense. If we set $k = bu^j$ to simplify notation, then $a+k$ is between 0 and a if $a-(a+k) = k$ is on the same line as a and points in the opposite direction, with $|k| < |a|$.

The proof given essentially reproves Proposition 0.7.7. If you want to use that proposition instead, you could say:

If $a + bu^j$ is between 0 and a, then there exists ρ with $0 < \rho < 1$ such that

$$a + bu^j = \rho a, \text{ i.e., } u^j = \frac{(\rho - 1)a}{b}.$$

This equation has j solutions by Proposition 0.7.7, and

$$|u| = (1-\rho)|a/b| < |a/b|,$$

so we can take $p_0 = |a/b|^{1/j}$.

1.6.11 Set $p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$ with k odd. Choose

$$C = \sup\{1, |a_{k-1}|, \dots, |a_0|\}$$

and set $A = kC + 1$. Then if $x \le -A$ we have

$$\begin{aligned} p(x) &= x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \\ &\le (-A)^k + CA^{k-1} + \cdots + C \le -A^k + kCA^{k-1} \\ &= A^{k-1}(kC - A) = -A^{k-1} \le 0. \end{aligned}$$

Similarly, if $x \ge A$ we have

$$\begin{aligned} p(x) &= x^k + a_{k-1}x^{k-1} + \cdots + a_1 x + a_0 \\ &\ge (A)^k - CA^{k-1} - \cdots - C \ge A^k - kCA^{k-1} \\ &= A^{k-1}(A - kC) = A^{k-1} \ge 0. \end{aligned}$$

Since $p : [-A, A] \to \mathbb{R}$ is a continuous function (Corollary 1.5.31) and we have $p(-A) \le 0$ and $p(A) \ge 0$, then by the intermediate value theorem there exists $x_0 \in [-A, A]$ such that $p(x_0) = 0$.

1.7.1 a. $f(a) = 0$, $f'(a) = \cos(a) = 1$, so the tangent is $g(x) = x$.

b. $f(a) = \frac{1}{2}$, $f'(a) = -\sin(a) = -\frac{\sqrt{3}}{2}$, so the tangent is

$$g(x) = -\frac{\sqrt{3}}{2}(x - \frac{\pi}{3}) + \frac{1}{2}.$$

c. $f(a) = 1$, $f'(a) = -\sin(a) = 0$, so the tangent is $g(x) = 1$.

d. $f(a) = 2$, $f'(a) = -\frac{1}{a^2} = -4$, so the tangent is

$$g(x) = -4(x - 1/2) + 2 = -4x + 4.$$

1.7.3 a. $f'(x) = \Big(3\sin^2(x^2 + \cos x)\Big)\Big(\cos(x^2 + \cos x)\Big)\Big(2x - \sin x\Big)$

b. $f'(x) = \Big(2\cos((x + \sin x)^2)\Big)\Big(-\sin((x + \sin x)^2)\Big)\Big(2(x + \sin x)\Big)\Big(1 + \cos x\Big)$

c. $f'(x) = \Big((\cos x)^5 + \sin x\Big)\Big(4(\cos x)^3\Big)(-\sin(x)) = (\cos x)^5 - 4(\sin x)^2(\cos x)^3$

d. $f'(x) = 3(x + \sin^4 x)^2(1 + 4\sin^3 x \cos x)$

e. $f'(x) = \dfrac{\sin^3 x(\cos x^2 * 2x)}{2 + \sin(x)} + \dfrac{\sin x^2(3\sin^2 x \cos x)}{2 + \sin(x)} - \dfrac{(\sin x^2 \sin^3 x)(\cos x)}{\big(2 + \sin(x)\big)^2}$

f. $f'(x) = \cos\left(\dfrac{x^3}{\sin x^2}\right)\left(\dfrac{3x^2}{\sin x^2} - \dfrac{(x^3)(\cos x^2 * 2x)}{(\sin x^2)^2}\right)$

1.7.5 a. Compute the partial derivatives:

$$D_1 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x}{\sqrt{x^2 + y}} \quad \text{and} \quad D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2\sqrt{x^2 + y}}.$$

This gives

$$D_1 f\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{2}{\sqrt{2^2 + 1}} = \frac{2}{\sqrt{5}} \quad \text{and} \quad D_2 f\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{2\sqrt{2^2 + 1}} = \frac{1}{2\sqrt{5}}.$$

At the point $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$, we have $x^2 + y < 0$, so the function is not defined there, and neither are the partial derivatives.

b. Similarly, $D_1 f \begin{pmatrix} x \\ y \end{pmatrix} = 2xy$ and $D_2 f \begin{pmatrix} x \\ y \end{pmatrix} = x^2 + 4y^3$. This gives

$$D_1 f \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 4 \quad \text{and} \quad D_2 f \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 4 + 4 = 8;$$

$$D_1 f \begin{pmatrix} 1 \\ -2 \end{pmatrix} = -4 \quad \text{and} \quad D_2 f \begin{pmatrix} 1 \\ -2 \end{pmatrix} = 1 + 4 \cdot (-8) = -31.$$

c. Compute

$$\begin{aligned} D_1 f \begin{pmatrix} x \\ y \end{pmatrix} &= -y \sin xy \\ D_2 f \begin{pmatrix} x \\ y \end{pmatrix} &= -x \sin xy + \cos y - y \sin y. \end{aligned}$$

This gives

$$\begin{aligned} D_1 f \begin{pmatrix} 2 \\ 1 \end{pmatrix} &= -\sin 2 \quad \text{and} \quad D_2 f \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -2 \sin 2 + \cos 1 - \sin 1 \\ D_1 f \begin{pmatrix} 1 \\ -2 \end{pmatrix} &= -2 \sin 2 \quad \text{and} \quad D_2 f \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \sin 2 + \cos 2 - 2 \sin 2 = \cos 2 - \sin 2 \end{aligned}$$

d. Since

$$D_1 f \begin{pmatrix} x \\ y \end{pmatrix} = \frac{xy^2 + 2y^4}{2(x + y^2)^{3/2}} \quad \text{and} \quad D_2 f \begin{pmatrix} x \\ y \end{pmatrix} = \frac{2x^2 y + xy^3}{(x + y^2)^{3/2}},$$

we have

$$\begin{aligned} D_1 f \begin{pmatrix} 2 \\ 1 \end{pmatrix} &= \frac{4}{2\sqrt{27}} \quad \text{and} \quad D_2 f \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{10}{\sqrt{27}}; \\ D_1 f \begin{pmatrix} 1 \\ -2 \end{pmatrix} &= \frac{36}{10\sqrt{5}} \quad \text{and} \quad D_2 f \begin{pmatrix} 1 \\ -2 \end{pmatrix} = -\frac{12}{5\sqrt{5}}. \end{aligned}$$

1.7.7 Just pile up the partial derivative vectors side by side:

$$\text{a.} \quad \left[\mathbf{D}\vec{\mathbf{f}}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{bmatrix} -\sin x & 0 \\ 2xy & x^2 + 2y \\ 2x \cos(x^2 - y) & -\cos(x^2 - y) \end{bmatrix}$$

$$\text{b.} \quad \left[\mathbf{D}\vec{\mathbf{f}}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ y & x \\ 2y \sin xy \ \cos xy & 2x \sin xy \ \cos xy \end{bmatrix}.$$

1.7.9 a. The derivative is an $m \times n$ matrix

b. a 1×3 matrix (line matrix)

c. a 4×1 matrix (vector 4 high)

1.7.11 a. $[y \cos(xy),\ x \cos(xy)]$ b. $[2xe^{x^2+y^3},\ 3y^2 e^{x^2+y^3}]$

c. $\begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix}$ d. $\begin{bmatrix} \cos\theta & -r \sin\theta \\ \sin\theta & r \cos\theta \end{bmatrix}$

1.7.13 For the first part, $|x|$ and mx are continuous functions, hence so is $f(0+h)-f(0)-mh = |h|-mh$.

For the second, we have

$$\begin{aligned} \frac{|h|-mh}{h} &= \frac{-h-mh}{h} = -1-m \quad \text{when } h<0 \\ \frac{|h|-mh}{h} &= \frac{h-mh}{h} \quad = \quad 1-m \quad \text{when } h>0. \end{aligned}$$

The difference between these values is always 2, and cannot be made small by taking h small.

Solution 1.7.15, part a: Writing the numerator as a length is optional, but the length in the denominator is not optional: you cannot divide by a matrix.

Since H is an $n\times m$ matrix, the $[0]$ in $\lim_{H\to[0]}$ is the $n\times m$ matrix with all entries 0.

1.7.15 a. There exists a linear transformation $[\mathbf{D}F(A)]$ such that

$$\lim_{H\to[0]} \frac{\big|F(A+H)-F(A)-[\mathbf{D}F(A)]H\big|}{|H|} = 0.$$

b. The derivative is $[\mathbf{D}F(A)]H = AH^\top + HA^\top$. We found this by looking for linear terms in H of the difference

$$\begin{aligned} F(A+H)-F(A) &= (A+H)(A+H)^\top - AA^\top \\ &= (A+H)(A^\top + H^\top) - AA^\top \\ &= AH^\top + HA^\top + HH^\top; \end{aligned}$$

see Remark 1.7.6. The linear terms $AH^\top + HA^\top$ are the derivative. Indeed,

$$\lim_{H\to[0]} \frac{|(A+H)(A+H)^\top - AA^\top - AH^\top - HA^\top|}{|H|}$$
$$= \lim_{H\to[0]} \frac{|HH^\top|}{|H|} \le \lim_{H\to[0]} \frac{|H|\,|H^\top|}{|H|} = \lim_{H\to[0]} |H| = 0.$$

1.7.17 The derivative of the squaring function is given by

$$[\mathbf{D}S(A)]H = AH + HA;$$

substituting $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $H = \begin{bmatrix} 0 & 0 \\ \epsilon & 0 \end{bmatrix}$ gives

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ \epsilon & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \epsilon & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \epsilon & 0 \\ \epsilon & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \epsilon & \epsilon \end{bmatrix} = \begin{bmatrix} \epsilon & 0 \\ 2\epsilon & \epsilon \end{bmatrix}.$$

Computing $(A+H)^2 - A^2$ gives the same result;

$$(A+H)^2 - A^2 = \begin{bmatrix} 1+\epsilon & 2 \\ 2\epsilon & 1+\epsilon \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \epsilon & 0 \\ 2\epsilon & \epsilon \end{bmatrix}.$$

Solution 1.7.17: This may look like a miracle: the expressions should not be equal, they should differ by terms in ϵ^2. The reason why they are exactly equal here is that

$$\begin{bmatrix} 0 & 0 \\ \epsilon & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

1.7.19 Since $\lim_{\vec{\mathbf{h}}\to\mathbf{0}} \frac{|\vec{\mathbf{h}}|\vec{\mathbf{h}}}{|\vec{\mathbf{h}}|} = [0]$, the derivative exists at the origin and is the 0 linear transformation, represented by the $n\times n$ matrix with all entries 0.

1.7.21 We will work directly from the definition of the derivative:

$$\begin{aligned}\det(I+H)-\det(I)-&(h_{1,1}+h_{2,2})\\&=(1+h_{1,1})(1+h_{2,2})-h_{1,2}h_{2,1}-1-(h_{1,1}+h_{2,2})\\&=h_{1,1}h_{2,2}-h_{1,2}h_{2,1}.\end{aligned}$$

Each $h_{i,j}$ satisfies $|h_{i,j}|\le|H|$, so we have

$$\frac{|\det(I+H)-\det(I)-(h_{1,1}+h_{2,2})|}{|H|}=\frac{|h_{1,1}h_{2,2}-h_{1,2}h_{2,1}|}{|H|}\le\frac{2|H|^2}{|H|}=2|H|.$$

Thus

$$\lim_{H\to 0}\frac{|\det(I+H)-\det(I)-(h_{1,1}+h_{2,2})|}{|H|}\le\lim_{H\to 0}2|H|=0.$$

1.8.1 Three make sense:

c. $\mathbf{g}\circ\mathbf{f}:\mathbb{R}^2\to\mathbb{R}^2$; the derivative is a 2×2 matrix
d. $\mathbf{f}\circ\mathbf{g}:\mathbb{R}^3\to\mathbb{R}^3$; the derivative is a 3×3 matrix
e. $f\circ\mathbf{f}:\mathbb{R}^2\to\mathbb{R}$; the derivative is a 1×2 matrix

1.8.3 Yes: We have a composition of sine, the exponential function, and the function $\begin{pmatrix}x\\y\end{pmatrix}\mapsto xy$, all of which are differentiable everywhere.

1.8.5 One must also show that fg is differentiable, working from the definition of the derivative.

1.8.7 Since $[\mathbf{D}f(x)]=[\,x_2\quad x_1+x_3\quad x_2+x_4\quad\cdots\quad x_{n-2}+x_n\quad x_{n-1}\,]$, we have
$[\mathbf{D}\big(f(\gamma(t))\big)]=[\,t^2\quad t+t^3\quad t^2+t^4\quad\cdots\quad t^{n-2}+t^n\quad t^{n-1}\,]$. In addition, $[\mathbf{D}\gamma(t)]=\begin{bmatrix}1\\2t\\\vdots\\nt^{n-1}\end{bmatrix}$. So the derivative of the function $t\to f(\gamma(t))$ is

$$\underbrace{[\mathbf{D}(f\circ\gamma)(t)]}_{\text{deriv. of comp. at }t}=\underbrace{[\mathbf{D}f(\gamma(t))]}_{\substack{\text{deriv. of }f\\\text{at }\gamma(t)}}\underbrace{[\mathbf{D}\gamma(t)]}_{\substack{\text{deriv. of }\gamma\\\text{at }t}}=t^2+\left(\sum_{i=2}^{n-1}it^{i-1}(t^{i-1}+t^{i+1})\right)+nt^{2(n-1)}$$

Solution 1.8.9: It is easiest to solve this problem without using the chain rule in several variables.

1.8.9 Clearly

$$D_1f\begin{pmatrix}x\\y\end{pmatrix}=2xy\varphi'(x^2-y^2);\qquad D_2f\begin{pmatrix}x\\y\end{pmatrix}=-2y^2\varphi'(x^2-y^2)+\varphi(x^2-y^2).$$

Thus

$$\begin{aligned}\frac{1}{x}D_1f\begin{pmatrix}x\\y\end{pmatrix}+\frac{1}{y}D_2f\begin{pmatrix}x\\y\end{pmatrix}&=2y\varphi'(x^2-y^2)-2y\varphi'(x^2-y^2)+\frac{1}{y}\varphi(x^2-y^2)\\&=\frac{1}{y^2}f\begin{pmatrix}x\\y\end{pmatrix}.\end{aligned}$$

To use the chain rule, write $f = k \circ \mathbf{h} \circ \mathbf{g}$, where

$$\mathbf{g}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ y \end{pmatrix}, \qquad \mathbf{h}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \varphi(u) \\ v \end{pmatrix}, \qquad k\begin{pmatrix} s \\ t \end{pmatrix} = st.$$

This leads to

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \end{pmatrix}\right] = [t, s]\begin{bmatrix} \varphi'(u) & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2x & -2y \\ 0 & 1 \end{bmatrix} = [2xt\varphi'(u),\ -2yt\varphi'(u) + s].$$

Insert the values of the variables; you find

$$D_1 f\begin{pmatrix} x \\ y \end{pmatrix} = 2xy\,\varphi'(x^2 - y^2) \text{ and } D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = -2y^2\varphi'(x^2 - y^2) + \varphi(x^2 - y^2).$$

Now continue as above.

1.8.11 Using the chain rule in one variable,

$$D_1 f = \varphi'\left(\frac{x+y}{x-y}\right)\left(\frac{1(x-y) - 1(x+y)}{(x-y)^2}\right) = \varphi'\left(\frac{x+y}{x-y}\right)\left(\frac{-2y}{(x-y)^2}\right)$$

and

$$D_2 f = \varphi'\left(\frac{x+y}{x-y}\right)\left(\frac{1(x-y) - (-1)(x+y)}{(x-y)^2}\right) = \varphi'\left(\frac{x+y}{x-y}\right)\left(\frac{2x}{(x-y)^2}\right)$$

so

$$xD_1 f + yD_2 f = \varphi'\left(\frac{x+y}{x-y}\right)\left(\frac{-2xy}{(x-y)^2}\right) + \varphi'\left(\frac{x+y}{x-y}\right)\left(\frac{2yx}{(x-y)^2}\right) = 0$$

Solution 1.8.13: Recall (Proposition 1.7.18) that if $T(A) = A^{-1}$ then T is a differentiable map on the space of invertible matrices, and

$$[\mathbf{D}T(A)]H = -A^{-1}HA^{-1}.$$

Here we replace A by $A^2 + A$ and H by $AH + HA + H$.

1.8.13 Call $S(A) = A^2 + A$ and $T(A) = A^{-1}$. We have $F = T \circ S$, so

$$\begin{aligned} [\mathbf{D}F(A)]H &= [\mathbf{D}T(A^2 + A)][\mathbf{D}S(A)]H \\ &= [\mathbf{D}T(A^2 + A)](AH + HA + H) \\ &= -(A^2 + A)^{-1}(AH + HA + H)(A^2 + A)^{-1}. \end{aligned}$$

It isn't really possible to simplify this much.

Solution 1.9.1: Simplify notation by setting $A = \sqrt{x^2 + y^2}$. Then, for instance, $D_1 f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is

$$\begin{aligned} &\frac{2x^5 + 4x^3y^2 - 2xy^4}{(x^2 + y^2)^2} \\ &\le \frac{2A^5 + 4A^5 + 2A^5}{A^4} \\ &\le 8A = 8\sqrt{x^2 + y^2}. \end{aligned}$$

We can't write $-2xy^4$ as $-2A^5$ and compute $2A^5 - 2A^5 = 0$ because x may be negative.

1.9.1 Except at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, the partial derivatives of f are given by

$$D_1 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{2x^5 + 4x^3y^2 - 2xy^4}{(x^2 + y^2)^2} \quad \text{and} \quad D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{4x^2y^3 - 2x^4y + 2y^5}{(x^2 + y^2)^2}.$$

At the origin, they are both given by

$$D_1 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lim_{h \to 0}\frac{1}{h}\left(\frac{h^4 + 0^4}{h^2 + 0^2}\right) = 0; \quad D_2 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lim_{h \to 0}\frac{1}{h}\left(\frac{0^4 + h^4}{0^2 + h^2}\right) = 0.$$

Thus there are partial derivatives everywhere, and we need to check that they are continuous. The only problem is at the origin. One easy way to show this is to remember that

$$|x| \le \sqrt{x^2 + y^2} \quad \text{and} \quad |y| \le \sqrt{x^2 + y^2}.$$

Then both partial derivatives satisfy

$$|D_i f\begin{pmatrix} x \\ y \end{pmatrix}| \le 8\frac{(x^2 + y^2)^{5/2}}{(x^2 + y^2)^2} = 8\sqrt{x^2 + y^2}.$$

Thus the limit of both partials at the origin is 0, so the partials are continuous and f is differentiable everywhere.

1.9.3 a. It means that there is a line matrix $[a\ b]$ such that

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}\left(\frac{\sin(h_1^2h_2^2)}{h_1^2+h_2^2} - ah_1 - bh_2\right) = 0.$$

Solution 1.9.3: The line matrix $[a\ b]$ is the derivative; it is the linear transformation L of Proposition and Definition 1.7.9.

b. Since f vanishes identically on both axes, both partials exist, and are 0 at the origin. In fact, D_1f vanishes on the x-axis and D_2f vanishes on the y-axis.

c. Since $D_1f(0) = 0$ and $D_2f(0) = 0$, we see that f is differentiable at the origin if and only if

$$\lim_{|\vec{\mathbf{h}}|\to 0} \frac{\sin(h_1^2h_2^2)}{(h_1^2+h_2^2)(h_1^2+h_2^2)^{1/2}} = 0,$$

and this is indeed true, since

$$\left|\sin(h_1^2h_2^2)\right| \le h_1^2h_2^2 \le \frac{1}{4}\left(h_1^2+h_2^2\right)^2 \qquad (1)$$

and

$$\lim_{|\vec{\mathbf{h}}|\to 0} \frac{1}{4}\frac{(h_1^2+h_2^2)^2}{(h_1^2+h_2^2)^{3/2}} = \frac{1}{4}\lim_{|\mathbf{h}|\to 0}(h_1^2+h_2^2)^{1/2} = 0.$$

Inequality 1, first inequality: For any x, we have $|\sin x| \le |x|$. The second inequality follows from

$$0 \le (x-y)^2 = x^2 - 2xy + y^2;$$

for any $x, y \in \mathbb{R}$ we have

$$|xy| \le \frac{1}{2}(x^2+y^2).$$

Solutions for review exercises, chapter 1

1.1 a. Not a subspace: $\vec{\mathbf{0}}$ is not on the line.

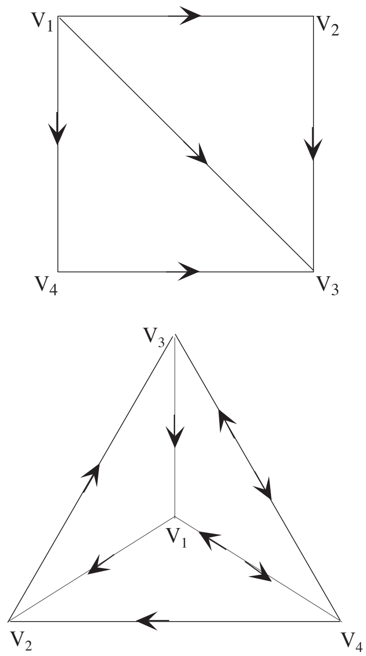
b. Not a subspace: $\vec{\mathbf{0}}$ is not on the line.

c. A subspace.

1.3 If A and B are upper triangular matrices, then if $i > j$ we know that $a_{i,j} = b_{i,j} = 0$. Using the definition of matrix multiplication:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j},$$

we see that if $i > j$, then in the summation either $a_{i,k} = 0$ or $b_{k,j} = 0$, so $c_{i,j} = \sum_{k=1}^{n} 0 = 0$. If for all $i > j$ we have $c_{i,j} = 0$, then C is upper triangular, so if A and B are upper triangular, AB is also upper triangular.



Labeling of vertices for Solution 1.5, parts b and c.

1.5 a. Labeling the vertices in the direction of the arrows: $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

b. With the labels shown in the figure: $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

c. Again, with the labeling as shown in the figure: $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

1.7 Except when $x = 0$, $y \neq 0$, there is no problem with continuity: the formula describes a product of continuous functions, divided by a function that doesn't vanish.

To understand the behavior near $x = 0$, $y \neq 0$, we can write

$$\frac{|y| e^{-|y|/x^2}}{x^2} = \frac{|y|}{x^2 \left(1 + \frac{|y|}{x^2} + \frac{1}{2} \left(\frac{|y|}{x^2} \right)^2 + \cdots \right)} < \frac{|y|}{y^2/2x^2}$$

which clearly tends to 0 when $x \to 0$ with $y \neq 0$. We used the fact that the series defining the exponential is a series of positive numbers, hence bounded below by any one of its terms.

Now for the directional derivatives at the origin. On the axes, f is identically 0, so the directional derivatives exist and are 0. On any other line we can set $y = mx$; at fixed m the function becomes

$$f \begin{pmatrix} x \\ mx \end{pmatrix} = \frac{|mx| e^{-|m/x|}}{x^2},$$

and we need to show that

$$\lim_{x\to 0}\frac{1}{x} f\begin{pmatrix} x \\ mx \end{pmatrix} = 0.$$

We find

$$\left|\frac{1}{x}\frac{|mx|e^{-|m/x|}}{x^2}\right| = \frac{|m|}{|x|^2\left(1+\left|\frac{m}{x}\right|+\frac{1}{2}\left|\frac{m}{x}\right|^2+\frac{1}{6}\left|\frac{m}{x}\right|^3+\cdots\right)} \le \frac{6|x|}{|m|^2},$$

which certainly tends to 0 with $|x|$.

1.9 a. Yes there is; its matrix is $[T] = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & -1 \\ 0 & 2 & 1 & -1 \end{bmatrix}$. We computed this matrix by determining what combinations of the four input vectors give the four standard basis vectors in $\mathbb{R}^4$. For example, if we call the input vectors $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4$, then $\vec{\mathbf{e}}_4 = \vec{\mathbf{v}}_3 - \vec{\mathbf{v}}_2$, so the fourth column of the matrix is $T\vec{\mathbf{e}}_4 = T\vec{\mathbf{v}}_3 - T\vec{\mathbf{v}}_2 = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$. Similarly, $\vec{\mathbf{e}}_1 = \vec{\mathbf{v}}_4 - \vec{\mathbf{v}}_3 + \vec{\mathbf{v}}_2$, so

Solution 1.9: This uses Theorem 1.3.4; the ith column of $[T]$ is $T\vec{\mathbf{e}}_i$.

$$T(\vec{\mathbf{e}}_1) = T\vec{\mathbf{v}}_4 - T\vec{\mathbf{v}}_3 + T\vec{\mathbf{v}}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

It is easy to confirm that this matrix does indeed satisfy the four equations of the exercise.

b. No, it is not linear: we have $[T]\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$, not $\begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}$. Another way to see this is to say that if S were linear, then by part a we would have

$$S\vec{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad S\vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \quad S\vec{\mathbf{e}}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad S\vec{\mathbf{e}}_4 = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix},$$

which by linearity should give $S\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = S(\vec{\mathbf{e}}_1 + \vec{\mathbf{e}}_2 + \vec{\mathbf{e}}_3 + \vec{\mathbf{e}}_4) = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

1.11 a. The matrices of S and T are

$$[S] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad [T] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

b. The matrices of the compositions are given by matrix multiplication:

$$[S\circ T] = [S][T] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad [T\circ S] = [T][S] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

c. The matrices $[S \circ T]$ and $[T \circ S]$ are inverses of each other: you can either compute it out, or note that since S and T are reflections, we have $S \circ S = T \circ T = I$, so

$$T \circ (S \circ S) \circ T = T \circ T = I \quad \text{and} \quad S \circ (T \circ T) \circ S = S \circ S = I.$$

d. They are the rotations by $2\pi/3$ and $-2\pi/3$ around the line $x = y = z$, counterclockwise if you look from a point of this line with positive coordinates towards the origin.

1.13 Below we denote by $|\overrightarrow{\text{side}}|$ the length of the side.

a. Because the side and the diagonal define a right triangle,

$$\begin{aligned}
\text{angle between } \overrightarrow{\text{side}} \text{ and } \overrightarrow{\text{diagonal}} &= \arccos\left(\frac{|\overrightarrow{\text{side}}|}{|\overrightarrow{\text{diagonal}}|}\right)\\
\theta_x &= \arccos\left(\frac{a}{\sqrt{a^2+b^2+c^2}}\right)\\
\theta_y &= \arccos\left(\frac{b}{\sqrt{a^2+b^2+c^2}}\right)\\
\theta_z &= \arccos(\frac{c}{\sqrt{a^2+b^2+c^2}})
\end{aligned}$$

b. Volume (parallelepiped) = abc = area(base) $\times$ height, but height = length(diagonal) $\times$ sin(angle of diagonal with face), so

$$\begin{aligned}
\text{angle } &= \arcsin\left(\frac{abc}{\text{area(base)}\sqrt{a^2+b^2+c^2}}\right)\\
\theta_{x-y} &= \arcsin\left(\frac{abc}{ab\sqrt{a^2+b^2+c^2}}\right) = \arcsin\left(\frac{c}{\sqrt{a^2+b^2+c^2}}\right)\\
\theta_{x-z} &= \arcsin\left(\frac{abc}{ac\sqrt{a^2+b^2+c^2}}\right) = \arcsin\left(\frac{b}{\sqrt{a^2+b^2+c^2}}\right)\\
\theta_{y-z} &= \arcsin\left(\frac{abc}{bc\sqrt{a^2+b^2+c^2}}\right) = \arcsin\left(\frac{a}{\sqrt{a^2+b^2+c^2}}\right)
\end{aligned}$$

1.15 a. The normalized vectors are:

$$\text{i. } \frac{1}{\sqrt{14}}\begin{bmatrix}2\\1\\3\end{bmatrix}, \qquad \text{ii. } \frac{1}{\sqrt{13}}\begin{bmatrix}-2\\3\end{bmatrix}, \qquad \text{iii. } \frac{1}{\sqrt{7}}\begin{bmatrix}\sqrt{3}\\0\\2\end{bmatrix}.$$

b. The angle θ satisfies $\cos\theta = \dfrac{2\sqrt{3}+6}{7\sqrt{2}}$, i.e., $\theta = \arccos \dfrac{2\sqrt{3}+6}{7\sqrt{2}}$.

Solution 1.17: There are many ways to write sequences. A sequence $j \mapsto \mathbf{x}_j$ can also be written as $\mathbf{x}_1, \mathbf{x}_2, \dots$ or as $(\mathbf{x}_j)_{j\in\mathbb{N}}$ or as $(\mathbf{x}_j)$ or even as $\mathbf{x}_j$, although this last could also refer to some collection of elements of the sequence. We prefer the notation $j \mapsto \mathbf{x}_j$.

1.17 a. Let $C_i,\ i \in I$ be some collection of closed subsets of $\mathbb{R}^n$. If the intersection is empty, it is automatically closed. So suppose the intersection is nonempty; we will use Proposition 1.5.17 to show that their intersection is closed. Indeed, let $j \mapsto \mathbf{x}_j$ be a convergent sequence in $\cap_{i\in I} C_i$, converging in $\mathbb{R}^n$ to some $\mathbf{a}$. Then every element of the sequence $j \mapsto \mathbf{x}_j$ belongs to each C_i and since the C_i are closed, we have $\mathbf{a} \in C_i$ for each $i \in I$. Therefore $\mathbf{a} \in \cap_{i \in I} C_i$.

Part b: We write "there exists C_k such that for infinitely many j's, we have $\mathbf{x}_j \in C_k$" instead of "infinitely many of the entries of the sequence must be elements of a single C_k" because we can't guarantee that the $\mathbf{x}_j$ are distinct; for instance, the sequence could eventually become constant.

b. Again, we will use Proposition 1.5.17. Let $C_1, \dots, C_m$ be a finite collection of closed subsets of $\mathbb{R}^m$. Suppose that $j \mapsto \mathbf{x}_j$ is a convergent sequence in the union $\cup_{i=1}^m C_i$, converging in $\mathbb{R}^n$ to some $\mathbf{a}$. Then there exists C_k such that for infinitely many j's, we have $x_j \in C_k$; these form a subsequence, which still converges to $\mathbf{a}$ by Proposition 1.5.19. Hence $\mathbf{a}$ is an element of C_k, hence also an element of $\cup_{i=1}^m C_i$. It follows from Proposition 1.5.17 that the union is closed.

c. The union of the closed sets $[0, (n-1)/n]$, $n = 2, 3, 4, \dots$ is the nonclosed set $[0, 1)$.

1.19 a. The derivative of the function zy^2 is $[0\ 2yz\ y^2]$; at $\mathbf{p} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ it is $[0\ 2\ 1]$, so the directional derivatives at $\mathbf{p}$ in the directions $\vec{\mathbf{e}}_1$, $\vec{\mathbf{e}}_2$, $\vec{\mathbf{e}}_3$, $\vec{\mathbf{v}}_1$, $\vec{\mathbf{v}}_2$ are

$$[0\ 2\ 1]\vec{\mathbf{e}}_1 = 0, \quad [0\ 2\ 1]\vec{\mathbf{e}}_2 = 2, \quad [0\ 2\ 1]\vec{\mathbf{e}}_3 = 1, \quad [0\ 2\ 1]\vec{\mathbf{v}}_1 = \sqrt{2}/2$$
$$[0\ 2\ 1]\vec{\mathbf{v}}_2 = 3\sqrt{2}/2.$$

So the function zy^2 increases most slowly in the direction $\vec{\mathbf{e}}_1$.

b. The derivative of the function $2x^2 - y^2$ at $\mathbf{p}$ is $[4\ -2\ 0]$, giving the directional derivatives at $\mathbf{p}$

$$[4\ \ -2\ \ \ 0]\vec{\mathbf{e}}_1 = 4, \qquad [4\ \ -2\ \ \ 0]\vec{\mathbf{e}}_2 = -2, \qquad [4\ \ -2\ \ \ 0]\vec{\mathbf{e}}_3 = 0$$
$$[4\ \ -2\ \ \ 0]\vec{\mathbf{v}}_1 = 2\sqrt{2}, \qquad [4\ \ -2\ \ \ 0]\vec{\mathbf{v}}_2 = -\sqrt{2}.$$

So to make $2x^2 - y^2$ increase as much as possible, also choose direction $\vec{\mathbf{e}}_1$.

3. To make $2x^2 - y^2$ decrease as much as possible, choose direction $\vec{\mathbf{e}}_2$.

1.21 a. This uses a trick:

$$\frac{x+y}{x^2-y^2} = \frac{x+y}{(x+y)(x-y)} = \frac{1}{x-y};$$

there is no limit as $\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, since when $\begin{pmatrix} x \\ y \end{pmatrix}$ is close to the origin, $x - y$ is also small (perhaps 0), so the quotient is big (or undefined).

b. Again the limit does not exist: on the line $y = kx$, this function is

$$\frac{x^4(1+k^2)^2}{x(1+k)} = x^3 \frac{(1+k^2)^2}{1+k}.$$

Choose $\epsilon > 0$, and set $1 + k = \epsilon^4$ and $x = \epsilon$; the function becomes

$$\frac{\epsilon^3}{\epsilon^4}(1+k^2) > \frac{1}{\epsilon}.$$

Thus there are points near the origin where the function is arbitrarily large. But there are also points where the function is arbitrarily close to 0, taking $x = y = \epsilon$.

c. Since $\lim_{u\to 0} u \ln|u| = 0$, the limit is 0.

d. Let us look at the function on the line $y = \epsilon$, for some $\epsilon > 0$. The function becomes

$$(x^2+\epsilon^2)(\ln|x| + \ln\epsilon) = x^2\ln|x| + \epsilon^2\ln|x| + x^2\ln\epsilon + \epsilon^2\ln\epsilon.$$

When $|x|$ is small, the first, third, and fourth terms are small. But the second is not. If $x = e^{-1/\epsilon^3}$, for instance, the second term is

$$-\frac{\epsilon^2}{\epsilon^3} = -\frac{1}{\epsilon},$$

which will become arbitrarily large as $\epsilon \to 0$. Thus along the curve given by $x = e^{-1/\epsilon^3}$ and $y = \epsilon$, the function tends to $-\infty$ as $\epsilon \to 0$. But along the curve $x = y = \epsilon$, the function tends to 0 as $\epsilon \to 0$, so it has no limit.

Solution 1.21, part c: The statement

$$\lim_{u\to 0} u\ln|u| = 0,$$

can be proved using l'Hôpital's rule applied to $\dfrac{\ln|u|}{1/u}$. We used this already in Solution 1.5.21.

1.23 Let π_n be π to n places (e.g., $\pi_2 = 3.14$). Then $\pi - \pi_n < 10^{-n}$. We can do the same with e. So:

$$\left|\mathbf{a}_n - \begin{bmatrix}\pi\\ e\end{bmatrix}\right| < \sqrt{2\cdot 10^{2(-n)}} = (\sqrt{2})10^{-n} < 10^{-n+1}.$$

So the best we can say in general is that to get $|\mathbf{a}_n - \begin{bmatrix}\pi\\ e\end{bmatrix}| < 10^{-m}$, we need $n = m+1$. When $m = 3$, we can get away with $n = 3$, since $(.59\dots)^2 + (.28\dots)^2 < 1$. But when $m = 4$, we really need $n = 5$, since $(.92\dots)^2 + (.81\dots)^2 > 1$.

Solution 1.23: For $m = 3$, we have

$$\left|\mathbf{a}_3 - \begin{bmatrix}\pi\\ e\end{bmatrix}\right| \approx \left|\begin{bmatrix}.00059\\ .00028\end{bmatrix}\right|$$
$$\approx \sqrt{(.0006)^2 + (.0003)^2}$$
$$= 10^{-3}\sqrt{.36+.09} < 10^{-3}.$$

1.25 We know by the fundamental theorem of algebra that $p(z)$ must have at least one root (and we know by Corollary 1.6.15 that, counted with multiplicity, it has exactly ten). We need to find R such that if $|z| \geq R$, then the leading term z^{10} has larger absolute value that the absolute value of the sum $2z^9 + 3z^8 + \cdots + 11$.

Being very sloppy with our inequalities, we see that

$$|2z^9 + 3z^8 + \cdots + 11| \leq 10\cdot 11|z|^9,$$

so if we take any $R > 110$, so that when $|z| \geq R$, then $|z|^{10} > 110|z^9|$, a root z_0 of the polynomial cannot satisfy $|z_0| \geq R$.

If we are less sloppy with our inequalities, we can observe that the successive terms of $2z^9 + \cdots + 11$ are decreasing in absolute value as soon as $|z| > 2$. So

$$|2z^9 + 3z^8 + \cdots + 11| \leq 2\cdot 10|z|^9,$$

and we see that $|z|^{10} > |2z^9 + 3z^8 + \cdots + 11|$ if $|z| > 20$.

1.27 We have $\sqrt{x^2} = |x|$, so

$$\lim_{h\to 0}\frac{1}{h}(f(h) - f(0)) = \lim_{h\to 0}\frac{|h|}{h},$$

which does not exist (since it is ± 1 depending on whether h is positive or negative).

For $\sqrt[3]{x^2}$, we have

$$\lim_{h\to 0}\frac{1}{h}(f(h)-f(0)) = \lim_{h\to 0}\frac{h^{2/3}}{h} = \lim_{h\to 0}\frac{1}{h^{1/3}},$$

which tends to $\pm\infty$. Of course, $\sqrt{x^4} = x^2$ is differentiable.

1.29 a. Both partials exist at $\begin{pmatrix}0\\0\end{pmatrix}$ and are 0, but the function is not differentiable. Indeed, if it were, the derivative would necessarily be the Jacobian matrix, i.e., the 0 matrix, and we would have

$$\lim_{\vec{\mathbf{h}}\to\mathbf{0}}\frac{f(\vec{\mathbf{h}})-f(\mathbf{0})-[0,0]\vec{\mathbf{h}}}{|\vec{\mathbf{h}}|} = 0.$$

But writing the definition out leads to

$$\lim_{\vec{\mathbf{h}}\to\mathbf{0}}\frac{h_1^2h_2}{(h_1^2+h_2^2)\sqrt{h_1^2+h_2^2}} = 0,$$

which isn't true: for instance, if you set $h_1 = h_2 = t$, the expression above becomes

$$\frac{t^3}{2\sqrt{2}\,|t|^3}, \quad \text{which does not become small as } t\to 0.$$

Solution 1.29, part b: If f were differentiable, then by the chain rule the composition would be differentiable, and it isn't.

b. This function is not differentiable. If you set $g(t) = \begin{pmatrix}t\\t\end{pmatrix}$, then $(f\circ g)(t) = 2|t|$ is not differentiable at $t=0$, but g is differentiable at $t=0$, so f is not differentiable at the origin, which is $g(0)$.

For x small, $\sin x \approx x$; see equation 3.4.6 for the Taylor polynomial of $\sin(x)$.

c. Here are two proofs of part c:

First proof This isn't even continuous at the origin, although both partials exist there and are 0. But if you set $x = t, y = t$, then

$$\frac{\sin(xy)}{x^2+y^2} = \frac{\sin t^2}{2t^2} \to \frac{1}{2}, \quad \text{as } t\to 0.$$

Second proof This function is not differentiable; although both partial derivatives exist at the origin, the function itself is not continuous at the origin. For example, along the diagonal,

$$\lim_{t\to 0}\frac{\sin t^2}{2t^2} = \frac{1}{2},$$

but the limit along the antidiagonal is $-1/2$:

$$\lim_{t\to 0}\frac{\sin(-t^2)}{2t^2} = -\frac{1}{2}.$$

1.31 a. By the chain rule this is

$$[\mathbf{D}f(t)] = \left[-\frac{1}{t+\sin t}, \frac{1}{t^2+\sin(t^2)}\right]\begin{bmatrix}1\\2t\end{bmatrix} = \frac{2t}{t^2+\sin(t^2)} - \frac{1}{t+\sin t}.$$

b. We defined f for $t>1$, but we will analyze it for all values of t. The function f is increasing for all $t>0$ and decreasing for all $t<0$. First let us see that f increases for $t > 0$, i.e., that the derivative is strictly positive for all $t > 0$. To see this, put the derivative on a common denominator:

$$[\mathbf{D}f(t)] = \frac{2t^2 + 2t\sin t - t^2 - \sin(t^2)}{(t^2 + \sin(t^2))(t + \sin t)}. \tag{1}$$

For $x > 0$, we have $x > \sin x$ (try graphing the functions x and $\sin x$), so for $t > 0$, the denominator is strictly positive. We need to show that the numerator is also strictly positive, i.e.,

$$t^2 + 2t\sin t - \sin(t^2) > 0. \tag{2}$$

For $-\pi < t \le \pi$, this is true: for $-\pi < t \le \pi$, we know that t and $\sin t$ are both positive or both negative, so $2t\sin t > 0$, and we have $t^2 > \sin(t^2)$ for $t \ne 0$. (Do not worry about $t^2 < t$ for small t; if you set $x = t^2$, the formula $x > \sin x$ still applies for $x > 0$.)

For $t > \pi$, equation 2 is also true: in that case,

$$t^2 + 2t\sin t - \sin(t^2) \ge t(t + 2\sin t) - 1 \ge t(\pi - 2) - 1 > \pi - 1.$$

To show that the function is decreasing for $t < 0$, we must show that the derivative is negative. For $t < 0$, the numerator is still strictly positive, by the argument above. But the denominator is negative: $t^2 + \sin(t^2)$ is positive, but $t + \sin t$ is negative.

Solution 1.31, part b: The function f tends to $-\infty$ as t tends to 0. To see this, note that if s is small, $\sin s$ is very close to s (see the margin note for Solution 1.29). So

$$\begin{aligned}\int_t^{t^2} \frac{ds}{s + \sin s} &\approx \int_t^{t^2} \frac{ds}{2s}\\ &= \frac{1}{2}\Big(\ln|t^2| - \ln|t|\Big)\\ &= \frac{1}{2}(2\ln|t| - \ln|t|\\ &= \frac{\ln|t|}{2}.\end{aligned}$$

1.33 Set $f(A) = A^{-1}$ and $g(A) = AA^\top + A^\top A$. Then $F = f \circ g$, and we wish to compute

$$\begin{aligned}[\mathbf{D}F(A)]H &= [\mathbf{D}f \circ g(A)]H = [\mathbf{D}f(g(A))][\mathbf{D}g(A)]H\\ &= [\mathbf{D}f(AA^\top + A^\top A)]\ \underbrace{[\mathbf{D}g(A)]H}_{\substack{\text{new increment}\\ \text{for } \mathbf{D}f}}\ .\end{aligned} \tag{1}$$

The linear terms in H of

$$g(A+H) - g(A) = (A+H)(A+H)^\top + (A+H)^\top(A+H) - AA^\top - A^\top A$$

are $AH^\top + HA^\top + A^\top H + H^\top A$; this is $[\mathbf{D}g(A)]H$, which is the new increment for $\mathbf{D}f$.

We know from Proposition 1.7.18 that $[\mathbf{D}f(A)]H = -A^{-1}HA^{-1}$, which we will rewrite as

$$[\mathbf{D}f(B)]K = -B^{-1}KB^{-1} \tag{2}$$

to avoid confusion. We substitute $AH^\top + HA^\top + A^\top H + H^\top A$ for the increment K in equation 2 and $g(A) = AA^\top + A^\top A$ for B. This gives

$$\begin{aligned}[\mathbf{D}F(A)]H &= [\mathbf{D}f(AA^\top + A^\top A)][\mathbf{D}g(A)]H\\ &= \underbrace{-\Big(AA^\top + A^\top A\Big)^{-1}}_{-B^{-1}}\underbrace{\Big(AH^\top + HA^\top + A^\top H + H^\top A\Big)}_{K}\underbrace{\Big(AA^\top + A^\top A\Big)^{-1}}_{B^{-1}}.\end{aligned}$$

There is no obvious way to simplify this expression.

Solution 1.33: Remember (see Example 1.7.17) that we cannot treat equation 1 as matrix multiplication. To treat the derivatives of f and g as matrices, we would have to identify Mat (n, n) with $\mathbb{R}^{n^2}$; the derivatives would be $n^2 \times n^2$ matrices. Instead we think of the derivatives as linear transformations.

1.35 a. All partial derivatives exist: away from the origin, they exist by Theorems 1.8.1 and 1.7.10; at the origin, they exist and are 0, since the function vanishes identically on all three coordinate axes.

b. By Theorem 1.8.1, the function is differentiable everywhere except at the origin. At the origin, the function is not differentiable. In fact, it isn't even continuous: the limit

$$\lim_{h\to 0} f\begin{pmatrix} t\\ t\\ t\end{pmatrix} = \lim_{h\to 0}\frac{h^3}{3h^4} = \lim_{h\to 0}\frac{1}{3h} \quad \text{does not exist.}$$

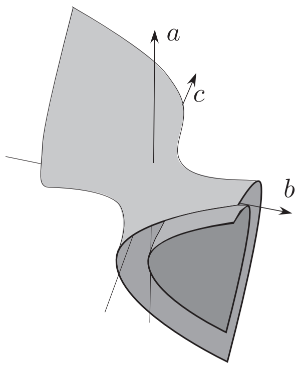


FIGURE FOR SOLUTION 1.37.

The outside surface has equation $a^2+bc=1$; it is a hyperboloid of one sheet and describes the solutions to part b. The inside surface has equation $a^2+bc=-1$; it is a hyperboloid of two sheets and describes the solutions to part c.

1.37 Compute

$$A^2 = \left(\begin{bmatrix} a & b\\ c & d\end{bmatrix}\right)^2 = \begin{bmatrix} a^2+bc & b(a+d)\\ c(a+d) & bc+d^2\end{bmatrix}$$

a. To get the diagonal entries of A^2 to be 0, we need $a^2=d^2=-bc$. Since $a^2=d^2$, we have $a=\pm d$; we will examine the cases $a=d$ and $a=-d$ separately. If $a=d\neq 0$, we get a contradiction: to get the off-diagonal terms to be 0 we must have either $b=0$ or $c=0$, but then a^2 must be 0, so a and d must be 0. If $a=d=0$, then either $b=0$ or $c=0$. Indeed, the matrices $\begin{bmatrix} 0 & b\\ 0 & 0\end{bmatrix}$ and $\begin{bmatrix} 0 & 0\\ c & 0\end{bmatrix}$ are matrices whose square is 0.

If $a=-d\neq 0$, then the off-diagonal terms of A^2 are 0. For the diagonal terms to be 0, we must have $bc=-a^2$, so neither b nor c can be 0 and we must have $c=\frac{-a^2}{b}$. Indeed, $A=\begin{bmatrix} a & b\\ -\frac{a^2}{b} & -a\end{bmatrix}$ satisfies $A^2=\begin{bmatrix} 0 & 0\\ 0 & 0\end{bmatrix}$.

b. If $A^2=I$, then

$$b(a+d)=c(a+d)=0$$

implies that either $b=c=0$ or $a+d=0$. If $b=c=0$, then $a^2=d^2=1$, so $a=\pm 1,\ d=\pm 1$, giving four solutions. If $a=-d$, then b and c must satisfy $bc=1-a^2$.

Solution 1.37, part b: The locus of equation $A^2=I$ is given by four equations in four unknowns, so you might expect it to be a union of finitely many points.

This is not the case; the locus is mainly the surface of equation $a^2+bc=1$: take any point of this surface and set $d=-a$ to find A such that $A^2=I$. There are two exceptional solutions not on this surface, namely $\pm I$.

c. As in part b, if $A^2=-I$, then either a+d=0, or $b=c=0$. The latter leads to $a^2=d^2=-1$, so there are no such solutions. So for all solutions $d=-a,\ a^2+bc=-1$.

1.39 a. This is a case where it is much easier to think of first rotating the telescope so that it is in the (x,z)-plane, then changing the elevation, then rotating back. This leads to the following product of matrices:

$$\begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & 0\\ \sin\theta_0 & \cos\theta_0 & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} \cos\varphi & 0 & -\sin\varphi\\ 0 & 1 & 0\\ \sin\varphi & 0 & \cos\varphi\end{bmatrix}\begin{bmatrix} \cos\theta_0 & \sin\theta_0 & 0\\ -\sin\theta_0 & \cos\theta_0 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta_0\cos\varphi-\sin^2\theta_0 & \cos\theta_0\sin\theta_0(\cos\varphi-1) & -\sin\varphi\cos\theta_0\\ cos\theta_0\sin\theta_0(\cos\varphi-1) & \cos\varphi\sin^2\theta_0+\cos^2\theta_0 & -\sin\theta_0\sin\varphi\\ \sin\varphi\cos\theta_0 & \sin\varphi\sin\theta_0 & \cos\varphi\end{bmatrix}$$

You may wonder about the signs of the $\sin\omega$ terms. Once the telescope is level, it is pointing in the direction of the x-axis. You, the astronomer rotating the telescope, are at the negative x end of the telescope. If you rotate it counterclockwise, as seen by you, the matrix is as we say. On the other hand, we are not absolutely sure that the problem is unambiguous as stated.

b. It is best to think of first rotating the telescope into the (x, z)-plane, then rotating it until it is horizontal (or vertical), then rotating it on its own axis, and then rotating it back (in two steps). This leads to the following product of five matrices:

$$\begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & 0 \\ \sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\varphi_0 & 0 & -\sin\varphi_0 \\ 0 & 1 & 0 \\ \sin\varphi_0 & 0 & \cos\varphi_0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega & \sin\omega \\ 0 & -\sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} \cos\varphi_0 & 0 & \sin\varphi_0 \\ 0 & 1 & 0 \\ -\sin\varphi_0 & 0 & \cos\varphi_0 \end{bmatrix} \begin{bmatrix} \cos\theta_0 & \sin\theta_0 & 0 \\ -\sin\theta_0 & \cos\theta_0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solutions for Chapter 2

2.1.1 a. $\begin{bmatrix} 3 & 1 & -4 \\ 0 & 2 & 1 \\ 1 & -3 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}\begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}$ b. $\begin{bmatrix} 3 & 1 & -4 & 0 \\ 0 & 2 & 1 & 4 \\ 1 & -3 & 0 & 1 \end{bmatrix}$ c. $\begin{bmatrix} 1 & -7 & 2 & 1 \\ 1 & -3 & 0 & 2 \\ 2 & -2 & 0 & -1 \end{bmatrix}$

2.1.3

a. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}$ b. $\begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 3 & 0 & -1 \\ 0 & 1 & 2 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 3 & -1 & 4 \\ 1 & 2 & 1 & 2 \\ 3 & 7 & 1 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 5 & -2 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

e. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & -3 & 3 & 3 \\ 1 & -4 & 2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 6/5 & 6/5 \\ 0 & 1 & -1/5 & -1/5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2.1.5 You can undo "multiplying row i by $m \neq 0$" by "multiplying row i by $1/m$" (which is possible because $m \neq 0$; see Definition 2.1.1).

You can undo "adding row i to row j" by "subtracting row i from row j," i.e., "adding ($-$row i) to row j".

You can undo "switching row i and row j" by "switching row i and row j" again.

2.1.7 Switching rows 2 and 3 of the matrix $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$ brings it to echelon form, giving $\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$.

The matrix $\begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ can be brought to echelon form by multiplying row 2 by 1/2 and subtracting row 3 from row 1, giving $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

The matrix $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ can be brought to echelon form by switching first the first and second rows, then the second and third rows:

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

The matrix $\begin{bmatrix} 0 & 1 & 0 & 3 & 0 & -3 \\ 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$ can be brought to echelon form by multiplying row 2 through by -1, then adding row 3 to row 2:

$$\begin{bmatrix} 0 & 1 & 0 & 3 & 0 & -3 \\ 0 & 0 & -1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 & 3 & 0 & -3 \\ 0 & 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 & 3 & 0 & -3 \\ 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}.$$

2.1.9 The first problem occurs when you subtract $2 \cdot 10^{10}$ from 1 to get from the second to the third matrix of equation 2.1.12 (second row, second entry). The 1 is "invisible" if computing only to 10 significant digits, and disappears in the subtraction: $1 - 20000000000 = -19999999999$, which to 10 significant digits is -20000000000. Another "invisible" 1 is found in the second row, third entry.

Solution 2.1.9: This is the main danger in numerical analysis: adding (or subtracting) numbers of very different sizes loses precision.

2.2.1 a. The augmented matrix $[A, \vec{\mathbf{b}}]$ corresponds to

$$\begin{aligned} 2x + y + 3z &= 1 \\ x - y &= 1 \\ x + y + 2z &= 1. \end{aligned}$$

Since $[A, \vec{\mathbf{b}}]$ row reduces to

$$\begin{bmatrix} \underline{1} & 0 & 1 & 0 \\ 0 & \underline{1} & 1 & 0 \\ 0 & 0 & 0 & \underline{1} \end{bmatrix},$$

x and y are pivotal unknowns, and z is a nonpivotal unknown.

b. If we list first the variable y, then z, then x, the system of equations becomes

$$\begin{array}{rcccccl} y & + & 3z & + & 2x & = 1 \\ -y & & & + & x & = 1 \\ y & + & 2z & + & x & = 1. \end{array}$$

The corresponding matrix is

$$\begin{bmatrix} 1 & 3 & 2 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 2 & 1 & 1 \end{bmatrix}, \text{ which row reduces to } \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This time y and z are the pivotal variables, and x is the nonpivotal variable.

2.2.3 a. Call the equations A, B, C, D. Adding A and B gives $2x + 4y - 2w = 0$; comparing this with C gives $-2w = z - 5w + v$, so

$$3w = z + v. \tag{1}$$

Comparing C and $2D$ gives $15w = 5z + 3v$, which is compatible with equation 1 only if $v = 0$. So equation 1 gives $3w = z$.

Substituting 0 for v and $3w$ for z in each of the four equations gives $z + 2y - w = 0$.

b. Since you can choose arbitrarily the value of y and w, and they determine the values of the other variables, the family of solutions depends on two parameters.

2.2.5 a. This system has a solution for every value of a. If you row reduce $\begin{bmatrix} a & 1 & 0 & 2 \\ 0 & a & 1 & 3 \end{bmatrix}$ you may seem to get

$$\begin{bmatrix} 1 & 0 & -1/a^2 & -\left(2/a + 3/a^2\right) \\ 0 & 1 & 1/a & 3/a \end{bmatrix},$$

which seems to indicate that there is a solution for any value of a except $a = 0$. However, obviously the system has a solution if $a = 0$; in that case, $y = 2$ and $z = 3$. The problem with the above row reduction is that if $a = 0$, then a can't be used for a pivotal 1. If $a = 0$, the matrix row reduces to $\begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}$.

b. We have two equations in three unknowns; there is no unique solution.

2.2.7 We can perform row operations to bring $\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & -1 & a & b \\ 2 & 0 & -b & 0 \end{bmatrix}$ to

$$\begin{bmatrix} 1 & 0 & (2+a)/2 & (1+b)/2 \\ 0 & 1 & (2-a)/2 & (1-b)/2 \\ 0 & 0 & 2+a+b & 1+b \end{bmatrix}.$$

a. There are then two possibilities. If $a + b + 2 \neq 0$, the first three columns row reduce to the identity, and the system of equations has the unique solution

$$x = \frac{b(b+1)}{2+a+b}, \quad y = \frac{-b^2 - 3b + 2a}{2(2+a+b)}, \quad z = \frac{1+b}{2+a+b}.$$

If $a+b+2 = 0$, then there are two possibilities to consider: either $b+1 = 0$ or $b+1 \neq 0$. If $b+1 = 0$, so that $a = b = -1$, the matrix row reduces to

$$\begin{bmatrix} 1 & 0 & 1/2 & 0 \\ 0 & 1 & 3/2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

In this case there are infinitely many solutions: the only nonpivotal variable is z, so we can choose its value arbitrarily; the others are $x = -z/2$ and

$y = 1 - (3z)/2$. If $a + b + 2 = 0$ and $b + 1 \neq 0$, then there is a pivotal 1 in the last column, and there are no solutions.

b. The first case, where $a + b + 2 \neq 0$, corresponds to an open subset of the (a, b)-plane. The second case, where $a = b = -1$, corresponds to a closed set. The third is neither open nor closed.

2.2.9 Row reducing

$$\begin{bmatrix} 1 & -1 & -1 & -3 & 1 & 1 \\ 1 & 1 & -5 & -1 & 7 & 2 \\ -1 & 2 & 2 & 2 & 1 & 0 \\ -2 & 5 & -4 & 9 & 7 & \beta \end{bmatrix} \text{ gives } \begin{bmatrix} 1 & 0 & 0 & -4 & 3 & 2 \\ 0 & 1 & 0 & -1/3 & 7/3 & 5/6 \\ 0 & 0 & 1 & -2/3 & -1/3 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & \beta + 1/2 \end{bmatrix}.$$

There are then two possibilities: either $\beta \neq -1/2$ or $\beta = -1/2$. If $\beta \neq -1/2$, there will be a pivotal 1 in the last column (once we have divided by $\beta + 1/2$), so there is no solution. But if $\beta = -1/2$, there are infinitely many solutions: x_4 and x_5 are nonpivotal, so their values can be chosen arbitrarily, and then the values of x_1, x_2 and x_3 are given by

$$\begin{aligned} x_1 &= 2 + 4x_4 - 3x_5 \\ x_2 &= 5/6 + x_4/3 - 7x_5/3 \\ x_3 &= 1/6 + 2x_4/3 + x_5/3. \end{aligned}$$

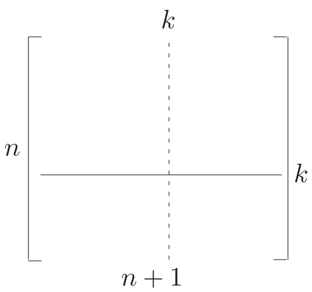


FIGURE FOR SOLUTION 2.2.11, part b. This $n \times (n+1)$ matrix represents a system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ of n equations in n unknowns. By the time we are ready to obtain a pivotal 1 at the intersection of the kth column (dotted) and kth row, all the entries on the kth row to the left of the kth column are 0, so we only need to place a 1 in position k, k and then justify that act by dividing all the entries on the kth row to the right of kth column by the (k, k) entry. There are $n + 1 - k$ such entries.

If the (k, k) entry is 0, we go down the kth column until we find a nonzero entry. In computing the total number of computations, we are assuming the worse case scenario, where all entries of the kth column are nonzero.

2.2.11 a. $R(1) = 1 + 1/2 - 1/2 = 1, \qquad R(2) = 8 + 2 - 1 = 9.$

If we have one equation in one unknown, we need to perform one division. If we have two equations in two unknowns, we need two divisions to get a pivotal 1 in the first row (the 1 is free), followed by two multiplications and two additions to get a 0 in the first element of the second row (the 0 is free). One more division, multiplication and addition get us a pivotal 1 in the second row and a 0 for the second element of the first row, for a total of nine.

b. As illustrated by the figure in the margin, we need $n+1-k$ divisions to obtain a pivotal 1 in the column k. To obtain a 0 in another entry of column k requires $n + 1 - k$ multiplications and $n + 1 - k$ additions. We need to do this for $n - 1$ entries of column k. So our total is

$$(n + 1 - k) + 2(n - 1)(n + 1 - k) = (2n - 1)(n - k + 1).$$

c. Using $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$, compute

$$\begin{aligned} \sum_{k=1}^{n} (2n-1)(n-k+1) &= \sum_{k=1}^{n} (2n-1)(n+1) - \sum_{k=1}^{n} (2n-1)k \\ &= n(2n-1)(n+1) - (2n-1)\frac{n(n+1)}{2} \\ &= n^3 + \frac{n^2}{2} - \frac{n}{2}. \end{aligned}$$

d.
$$Q(1) = \frac{2}{3} + \frac{3}{2} - \frac{7}{6} = 1,$$
$$Q(2) = \frac{2}{3}8 + \frac{3}{2}4 - \frac{7}{6}2 = 9,$$
$$Q(3) = \frac{2}{3}27 + \frac{3}{2}9 - \frac{7}{6}3 = 28.$$

The function $R(n) - Q(n)$ is the cubic polynomial $\frac{1}{3}n^3 - n^2 + \frac{2}{3}n$, with a root at $n = 2$. Its derivative, $n^2 - 2n + \frac{2}{3}$, has roots at $1 \pm \sqrt{1/3}$; it is strictly positive for $n \geq 2$. So the function $R(n) - Q(n)$ is increasing as a function of n for $n \geq 2$, and hence is strictly positive for $n \geq 3$.

Another way to show that $Q(n) < R(n)$ is to note that for $n \geq 3$,
$$R(n) - Q(n) = \frac{1}{3}n^3 - n^2 + \frac{2}{3}n \geq \frac{1}{3}(3n^2) - n^2 + \frac{2}{3}n = \frac{2}{3}n > 0.$$

e. For partial row reduction for a single column, the operations needed are like those for full row reduction (part b) except that we are just putting zeros below the diagonal, so we can replace $n - 1$ in the total for full row reduction by $n - k$, to get
$$(n + 1 - k) + 2(n - k)(n + 1 - k) = (n - k + 1)(2n - 2k + 1)$$
total operations (divisions, multiplications, and additions).

f. Denote by $P(n)$ the total computations needed for partial row reduction. By part e, we have
$$P(n) = \sum_{k=1}^{n} (n - k + 1)(2n - 2k + 1).$$
Let
$$P_1(n) = \frac{2}{3}n^3 + \frac{1}{2}n^2 - \frac{1}{6}n.$$
We will show by induction that $P = P_1$.

Clearly, $P(1) = P_1(1) = 1$. If $P(n) = P_1(n)$, we get:

In line 4, we get the next-to-last term using
$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}.$$

$$\begin{aligned} P(n+1) &= \sum_{k=1}^{n+1}(n-k+2)(2n-2k+3) = 1 + \sum_{k=1}^{n}(n-k+2)(2n-2k+3) \\ &= 1 + \sum_{k=1}^{n}\Big((n-k+1)+1\Big)\Big((2n-2k+1)+2\Big) \\ &= 1 + \underbrace{\sum_{k=1}^{n}(n-k+1)(2n-2k+1)}_{P(n)} + \sum_{k=1}^{n}(4n-4k+5) \\ &= 1 + \underbrace{\frac{2}{3}n^3 + \frac{1}{2}n^2 - \frac{1}{6}n}_{P_1(n) \text{ by inductive hypothesis}} + 4n^2 - 4\frac{n^2+n}{2} + 5n \\ &= \frac{2}{3}(n+1)^3 + \frac{1}{2}(n+1)^2 - \frac{1}{6}(n+1) = P_1(n+1). \end{aligned}$$

So the relation is true for all $n \geq 1$.

This can also be solved by direct computation (as in part c), using

$$\sum_{k=1}^{n} k = \frac{n(n_+1)}{2} \quad \text{and} \quad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

g. We need $n-k$ multiplications and $n-k$ additions for the row k, so the total number of operations for back substitution is $B(n) = n^2 - n$.

h. So the total number of operations for n equations in n unknowns is

$$Q(n) = P(n) + B(n) = \frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n \quad \text{for all } n \geq 1.$$

2.3.1 The inverse of A is $A^{-1} = \begin{bmatrix} 3 & -1 & -4 \\ 1 & -1 & -1 \\ -2 & 1 & 3 \end{bmatrix}$. Now compute

$$\begin{bmatrix} 3 & -1 & -4 \\ 1 & -1 & -1 \\ -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 & -5 \\ -1 & 1 & -2 \\ 2 & -1 & 4 \end{bmatrix}.$$

The columns of the product are the solutions to the three systems we were trying to solve.

2.3.3 For instance,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \text{ but } \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

2.3.5 a. Since $A = \begin{bmatrix} 3 & -1 & 3 & 1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ row reduces to $\begin{bmatrix} 1 & 0 & 0 & 3/8 \\ 0 & 1 & 0 & 1/2 \\ 0 & 0 & 1 & 1/8 \end{bmatrix}$, the solution is $x = 3/8$, $y = 1/2$, $z = 1/8$.

b. Since $\begin{bmatrix} 3 & -1 & 3 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$ row reduces to

$$\begin{bmatrix} 1 & 0 & 0 & 3/16 & 1/4 & -1/16 \\ 0 & 1 & 0 & -1/4 & 0 & 3/4 \\ 0 & 0 & 1 & 1/16 & -1/4 & 5/16 \end{bmatrix}, \quad \text{we have}$$

$$A^{-1} = \begin{bmatrix} 3/16 & 1/4 & -1/16 \\ -1/4 & 0 & 3/4 \\ 1/16 & -1/4 & 5/16 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3/16 & 1/4 & -1/16 \\ -1/4 & 0 & 3/4 \\ 1/16 & -1/4 & 5/16 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/8 \\ 1/2 \\ 1/8 \end{bmatrix}.$$

2.3.7 It just so happens that $A = A^{-1}$:

$$\begin{bmatrix} 1 & -6 & 3 \\ 2 & -7 & 3 \\ 4 & -12 & 5 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \text{So by Proposition 2.3.1, we have}$$

$$\vec{\mathbf{x}} = A^{-1}\begin{bmatrix}5\\7\\11\end{bmatrix} = A\begin{bmatrix}5\\7\\11\end{bmatrix} = \begin{bmatrix}-4\\-6\\-9\end{bmatrix}.$$

2.3.9 a. The products will be

$$\begin{bmatrix}-2 & 3 & -14\\0 & 2 & 3\\1 & 0 & 4\end{bmatrix} \quad \text{3 times the third row is subtracted from the first}$$

$$\begin{bmatrix}1 & 3 & -2\\0 & 4 & 6\\1 & 0 & 4\end{bmatrix} \quad \text{the second row is multiplied by 2}$$

$$\begin{bmatrix}1 & 3 & -2\\1 & 0 & 4\\0 & 2 & 3\end{bmatrix} \quad \text{the second and third rows are switched.}$$

b.

$$\begin{array}{ccc} & \begin{bmatrix}1 & 3 & -2\\0 & 2 & 3\\1 & 0 & 4\end{bmatrix} & & \begin{bmatrix}1 & 3 & -2\\0 & 2 & 3\\1 & 0 & 4\end{bmatrix} & & \begin{bmatrix}1 & 3 & -2\\0 & 2 & 3\\1 & 0 & 4\end{bmatrix}\\ \begin{bmatrix}1 & 0 & -3\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} & \begin{bmatrix}-2 & 3 & -14\\0 & 2 & 3\\1 & 0 & 4\end{bmatrix}, & \begin{bmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 1\end{bmatrix} & \begin{bmatrix}1 & 3 & -2\\0 & 4 & 6\\1 & 0 & 4\end{bmatrix}, & \begin{bmatrix}1 & 0 & 0\\0 & 0 & 1\\0 & 1 & 0\end{bmatrix} & \begin{bmatrix}1 & 3 & -2\\1 & 0 & 4\\0 & 2 & 3\end{bmatrix}.\end{array}$$

2.3.11 Let A be an $n \times m$ matrix. Then

$AE_1(i, x)$ has the same columns as A, except the ith, which is multiplied by x.

$AE_2(i, j, x)$ has the same columns as A except the jth, which is the sum of the jth column of A (contributed by the 1 in the (j, j)th position), and x times the ith column (contributed by the x in the (i, j)th position).

$AE_3(i, j)$ has the same columns as A, except for the ith and jth, which are switched.

2.3.13 Here is one way to show this. Denote by a the ith row and by b the jth row of our matrix. Assume we wish to switch the ith and the jth rows. Then multiplication on the left by $E_2(i, j, 1)$ turns the ith row into $a + b$. Multiplication on the left by $E_2(j, i, -1)$ then by $E_1(j, -1)$ turns the jth row into a. Finally, we multiply on the left by $E_2(i, j, -1)$ to subtract a from the ith row, making that row b. So we can switch rows by multiplying with only the first two types of elementary matrices.

Here is a different explanation of the same argument: Compute the product

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

This certainly shows that the 2×2 elementary matrix $E_3(1, 2)$ can be written as a product of elementary matrices of type 1 and 2.

More generally,

$$E_2(i,j,-1)E_1(j,-1)E_2(j,i,-1)E_2(i,j,1) = E_3(i,j).$$

2.3.15 1. Suppose that $[A|I]$ row reduces to $[I|B]$. This can be expressed as multiplication on the left by elementary matrices:

$$E_k \dots E_1[A|I] = [I|B]. \tag{1}$$

The left and right sides of equation (1) give

$$E_k \dots E_1 A = I \quad \text{and} \quad E_k \dots E_1 I = B. \tag{2}$$

Thus B is a product of elementary matrices, which are invertible, so (by Proposition 1.2.15) B is invertible: $B^{-1} = E_1^{-1} \dots E_k^{-1}$. Substituting the right equation of (2) into the left equation gives $BA = I$, so B is a left inverse of A. We don't need to check that it is also a right inverse, but doing so is straightforward: multiplying $BA = I$ by B^{-1} on the left and B on the right gives

$$I = B^{-1}IB = B^{-1}(BA)B = (B^{-1}B)AB = AB.$$

So B is also a right inverse of A.

Going from equation 1 to equation 2 in Solution 2.3.15: If

$$[E][A|I] = [I|B], \quad \text{then}$$
$$[E][A] = I \quad \text{and} \quad [E][I] = B,$$

since the multiplication of each column of A and of I by $[E]$ occurs independently of all other columns:

$$\begin{array}{cc} & [A|I] \\ [E] & [EA|EI]. \end{array}$$

2. If $[A|I]$ row reduces to $[A'|A'']$, where A' is *not* the identity, then by Theorem 2.2.6 the equation $A\vec{\mathbf{x}}_i = \vec{\mathbf{e}}_i$ either has no solution or has infinitely many solutions for each $i = 1, \dots, n$. In either case, A is not invertible.

2.4.1 The only way you can write

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + a_k \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix},$$

is if $a_1 = a_2 = \dots = a_k = 0$.

Solution 2.4.3: The vectors in equation 1 are orthogonal, by Corollary 1.4.8; they are orthonormal because they are orthogonal and each vector has length 1.

2.4.3 To make the basis orthonormal, each vector needs to be normalized to give it length 1. This is done by dividing each vector by its length (see equation 1.4.6). So the orthonormal basis is $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$. These vectors form a basis of $\mathbb{R}^2$ because they are two linearly independent vectors in $\mathbb{R}^2$; they are orthogonal because

$$\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = 0. \tag{1}$$

Solution 2.4.5: Recall that only a very few, very special subsets of $\mathbb{R}^n$ are subspaces of $\mathbb{R}^n$; see Definition 1.1.5. Roughly, a subspace is a flat subset that goes through the origin: to be closed under multiplication, a subspace must contain the zero vector, so that $0\vec{\mathbf{v}} = \vec{\mathbf{0}}$.

2.4.5 To show that Span $(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k)$ is a subspace of $\mathbb{R}^n$, we need to show that it is closed under addition and under multiplication by scalars. This follows from the computations

$$\begin{aligned} &c(a_1\vec{\mathbf{v}}_1 + \dots + a_k\vec{\mathbf{v}}_k) = ca_1\vec{\mathbf{v}}_1 + \dots + ca_k\vec{\mathbf{v}}_k; \\ &(a_1\vec{\mathbf{v}}_1 + \dots + a_k\vec{\mathbf{v}}_k) + (b_1\vec{\mathbf{v}}_1 + \dots + b_k\vec{\mathbf{v}}_k) \\ &\qquad\qquad = (a_1+b_1)\vec{\mathbf{v}}_1 + \dots + (a_k+b_k)\vec{\mathbf{v}}_k. \end{aligned}$$

To see that it is the smallest subspace that contains the $\vec{\mathbf{v}}_i$, note that any subspace that contains the $\vec{\mathbf{v}}_i$ must contain their linear combinations, hence the smallest such subspace is Span $(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k)$.

2.4.7 Both parts depend on the definition of an orthonormal basis:

$$\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j \end{cases}$$

Proof of part 1: Since $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ form a basis, they span $\mathbb{R}^n$: any $\vec{\mathbf{x}} \in \mathbb{R}^n$ can be written

$$\vec{\mathbf{x}} = a_1\vec{\mathbf{v}}_1 + \dots + a_n\vec{\mathbf{v}}_n.$$

Take the dot product of both sides with $\vec{\mathbf{v}}_i$:

$$\vec{\mathbf{x}} \cdot \vec{\mathbf{v}}_i = a_1\vec{\mathbf{v}}_1 \cdot \vec{\mathbf{v}}_i + \dots + a_n\vec{\mathbf{v}}_n \cdot \vec{\mathbf{v}}_i = a_i\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_i = a_i.$$

Proof of part 2:

$$|\vec{\mathbf{x}}|^2 = \left(\sum_{i=1}^n a_i\vec{\mathbf{v}}_i\right) \cdot \left(\sum_{j=1}^n a_j\vec{\mathbf{v}}_j\right) = \sum_{i=1}^n\sum_{j=1}^n a_i a_j(\vec{\mathbf{v}}_i \cdot \vec{\mathbf{v}}_j) = \sum_{i=1}^n a_i^2.$$

2.4.9 To see that condition 2 implies condition 3, first note that $2 \implies 3$ is logically equivalent to (not 3) $\implies$ (not 2). Now suppose $\{\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k\}$ is a *linearly dependent* set spanning V, so by Definition 2.4.2, there exists a nontrivial solution to

$$a_1\vec{\mathbf{v}}_1 + \dots + a_k\vec{\mathbf{v}}_k = \mathbf{0}.$$

Without loss of generality, we may assume that a_k is nonzero (if it isn't, renumber the vectors so that a_k is nonzero). Using the above relation, we can solve for $\vec{\mathbf{v}}_k$ in terms of the other vectors:

$$\vec{\mathbf{v}}_k = -\left(\frac{a_1}{a_k}\vec{\mathbf{v}}_1 + \dots + \frac{a_{k-1}}{a_k}\vec{\mathbf{v}}_{k-1}\right).$$

This implies that $\{\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k\}$ cannot be a minimal spanning set, because if we were to drop $\vec{\mathbf{v}}_k$ we could still form all the linear combinations as before. So (not 3) $\implies$ (not 2), and we are finished.

To show that $3 \implies 1$:

The vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k$ span V, so for any vector $\vec{\mathbf{w}} \in V$, there exist some numbers $a_1, \dots, a_n$ such that

$$a_1\vec{\mathbf{v}}_1 + \dots + a_n\vec{\mathbf{v}}_n = \vec{\mathbf{w}}.$$

Thus, if we add this vector to $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k$, we will have a linearly dependent set because

$$a_1\vec{\mathbf{v}}_1 + \dots + a_n\vec{\mathbf{v}}_n - \vec{\mathbf{w}} = \mathbf{0}$$

is a nontrivial linear combination of the vectors that equals $\mathbf{0}$. Since $\vec{\mathbf{w}}$ can be any vector in V, the set $\{\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k\}$ is a maximal linearly independent set.

2.4.11 a. For any n, we have $n+1$ linear equations for the $n+1$ unknowns $a_{0,n},\ a_{1,n}, \dots, a_{n,n}$, which say

$$a_{0,n}\left(\frac{0}{n}\right)^k + a_{1,n}\left(\frac{1}{n}\right)^k + a_{2,n}\left(\frac{2}{n}\right)^k + \cdots + a_{n,n}\left(\frac{n}{n}\right)^k = \int_0^1 x^k\,dx = \frac{1}{k+1},$$

one for each $k = 0, 1, \dots, n$. These systems of linear equations are:

- When $n = 1$

$$\begin{aligned} a_{0,1}1 + a_{1,1}1 &= 1 \\ a_{0,1}0 + a_{1,1}1 &= 1/2 \end{aligned}$$

- When $n = 2$

$$\begin{aligned} a_{0,2}1 + a_{1,2}1 + a_{2,2}1 &= 1 \\ a_{0,2}0 + a_{1,2}(1/2) + a_{2,2}1 &= 1/2 \\ a_{0,2}0 + a_{1,2}(1/4) + a_{2,2}1 &= 1/3 \end{aligned}$$

- When $n = 3$

$$\begin{aligned} a_{0,3}1 + a_{1,3}1 + a_{2,3}1 + a_{3,3}1 &= 1 \\ a_{0,3}0 + a_{1,3}(1/3) + a_{2,3}(2/3) + a_{3,3}1 &= 1/2 \\ a_{0,3}0 + a_{1,3}(1/9) + a_{2,3}(4/9) + a_{3,3}1 &= 1/3 \\ a_{0,3}0 + a_{1,3}(1/27) + a_{2,3}(8/27) + a_{3,3}1 &= 1/4. \end{aligned}$$

The system of equations for $n = 3$ could be written as the augmented matrix $[A|\vec{\mathbf{b}}]$:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1/3 & 2/3 & 1 & 1/2 \\ 0 & 1/9 & 4/9 & 1 & 1/3 \\ 0 & 1/27 & 8/27 & 1 & 1/4 \end{bmatrix}.$$

b. These wouldn't be too bad to solve by hand (although already the last would be distinctly unpleasant). We wrote a little MATLAB m-file to do it systematically:

```
function [N,b,c] = EqSp(n)

N = zeros(n+1); % make an n+1 × n+1 matrix of zeros
c=linspace(1,n+1,n+1); % make a place holder for the right side
for i=1:n+1
    for j=1:n+1
    N(i,j)= ((j-1)/n)^(i-1); % put the right coefficients in the matrix
end
c(i)=1/c(i); % put the right entries in the right side
end
b=c'; % our c was a row vector, take its transpose
c=N% this solves the system of linear equations
```

If you write and save this file as 'EqSp.m', and then type

[A,b,c]=EqSp(5), for the case when $n = 5$,

you will get

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1/5 & 2/5 & 3/5 & 4/5 & 1 \\ 0 & 1/25 & 4/25 & 9/25 & 16/25 & 1 \\ 0 & 1/125 & 8/125 & 27/125 & 64/125 & 1 \\ 0 & 1/625 & 16/625 & 81/625 & 256/625 & 1 \\ 0 & 1/3125 & 32/3125 & 243/3125 & 541/1651 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \\ 1/4 \\ 1/5 \\ 1/6 \end{bmatrix}, \quad c = \begin{bmatrix} 19/288 \\ 25/96 \\ 25/144 \\ 25/144 \\ 25/96 \\ 19/288 \end{bmatrix}.$$

This corresponds to the equation $Ac = b$, where the matrix A is the matrix of coefficients for $n = 5$, and the vector c is the desired set of coefficients – the solutions when $n = 5$.

When $n = 1, 2, 3$, the coefficients – i.e., the solutions to the systems of equations in part a – are

For instance, for $n = 2$, we have
$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}.$$

$$\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \quad \begin{bmatrix} 1/6 \\ 2/3 \\ 1/6 \end{bmatrix}, \quad \begin{bmatrix} 1/8 \\ 3/8 \\ 3/8 \\ 1/8 \end{bmatrix}.$$

The approximations to $\int_0^1 \frac{dx}{1+x} = \log 2 = 0.69314718055995\ldots$ obtained with these coefficients are .75 for $n = 1$, $\frac{25}{36} = .6944\ldots$ for $n = 2$, and $\frac{111}{160} = .69375$ for $n = 3$.

c. If you compute

$$\sum_{i=0}^{5} a_{i,5} \frac{1}{(i/5)+1} \approx \int_0^1 \frac{dx}{1+x} = \log 2 = 0.69314718055995\ldots$$

you will find 0.69316302910053, which is a pretty good approximation for a Riemann sum with six terms. For instance, the midpoint Riemann sum gives

$$\frac{1}{5} \sum_{i=1}^{5} \frac{1}{\big((2i-1)/10\big)} \approx 0.69190788571594,$$

which is a much worse approximation. But this scheme runs into trouble. All the coefficients are positive up to $n = 7$, but for $n = 8$ they are

$$\begin{bmatrix} 248/7109 \\ 578/2783 \\ -111/3391 \\ 97/262 \\ -454/2835 \\ 97/262 \\ -111/3391 \\ 578/2783 \\ 248/7109 \end{bmatrix} \approx \begin{bmatrix} 0.0349 \\ 0.2077 \\ -0.0327 \\ 0.3702 \\ -0.1601 \\ 0.3702 \\ -0.0327 \\ 0.2077 \\ 0.0349 \end{bmatrix}$$

and the approximation scheme starts depending on cancellations. This is much worse when $n = 20$, where the coefficients are as shown in the margin.

$$\begin{bmatrix} 0.0118 \\ 0.1141 \\ -0.2362 \\ 1.2044 \\ -3.7636 \\ 10.3135 \\ -22.6521 \\ 41.7176 \\ -63.9006 \\ 82.5706 \\ -89.7629 \\ 82.5829 \\ -63.9189 \\ 41.7345 \\ -22.6633 \\ 10.3191 \\ -3.7656 \\ 1.2050 \\ -0.2363 \\ 0.1141 \\ 0.0118 \end{bmatrix}.$$

Coefficients when $n = 20$.

Despite these bad sign variations, the Riemann sum works pretty well: the approximation to the integral above gives 0.69314718055995, which is $\ln 2$ to the precision of the machine.

2.4.13 In the process of row reducing $A = \begin{bmatrix} 1 & a & a & a \\ 1 & 1 & a & a \\ 1 & 1 & 1 & a \\ 1 & 1 & 1 & 1 \end{bmatrix}$, you will come to the matrix

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

If $a = 1$, the matrix will not row reduce to the identity, because you can't choose a pivotal 1 in the second column, so one necessary condition for A to be invertible is that $a \neq 1$. Let us suppose that this is the case. We can now row reduce two steps further to find

$$\begin{bmatrix} 1 & 0 & a & a \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1-a & 0 \\ 0 & 0 & 1-a & 1-a \end{bmatrix}, \quad \text{and then} \quad \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1-a \end{bmatrix}$$

The next step row reduces the matrix to the identity, so the matrix is invertible if and only if $a \neq 1$.

2.4.15 Suppose $\vec{\mathbf{w}}$ can be written as a linear combination of the $\vec{\mathbf{v}}_i$ in two ways: $\vec{\mathbf{w}} = \sum_{i=1}^k x_i \vec{\mathbf{v}}_i$ and $\vec{\mathbf{w}} = \sum_{i=1}^k y_i \vec{\mathbf{v}}_i$, with $x_i \neq y_i$ for at least one i. Then we can write

$$\sum_{i=1}^k x_i \vec{\mathbf{v}}_i - \sum_{i=1}^k y_i \vec{\mathbf{v}}_i = \vec{\mathbf{0}}, \quad \text{i.e.,} \quad \sum_{i=1}^k (x_i - y_i)\vec{\mathbf{v}}_i = \vec{\mathbf{0}},$$

and it follows from Definition 2.4.2 that the $\vec{\mathbf{v}}_i$ are not linearly independent. In the other direction, if any vector can be written as a linear combination of the $\vec{\mathbf{v}}_i$ in at most one way, then $\vec{\mathbf{0}} = \sum_{i=1}^k 0\vec{\mathbf{v}}_i = \sum_{i=1}^k x_i \vec{\mathbf{v}}_i$ implies that all the x_i are 0, so the $\vec{\mathbf{v}}_i$ are linearly independent.

2.5.1 a. The vectors $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_3$ are in the kernel of A, since $A\vec{\mathbf{v}}_1 = \vec{\mathbf{0}}$ and $A\vec{\mathbf{v}}_3 = \vec{\mathbf{0}}$. But $\vec{\mathbf{v}}_2$ is not, since $A\vec{\mathbf{v}}_2 = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$. The vector $\begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$ is in the image of A.

b. The matrix T represents a transformation from $\mathbb{R}^5$ to $\mathbb{R}^3$; it takes a vector in $\mathbb{R}^5$ and gives a vector in $\mathbb{R}^3$. Therefore, $\vec{\mathbf{w}}_4$ has the right height to be in the kernel (although it isn't), and $\vec{\mathbf{w}}_1$ and $\vec{\mathbf{w}}_3$ have the right height to be in its image.

Since the sum of the second and fifth columns of T is $\vec{\mathbf{0}}$, one element of the kernel is $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

2.5.3 nullity $T = \dim\ker T =$ number of nonpivotal columns of T;

$$\begin{aligned}\text{rank of } T &= \text{dim image } T\\ &= \text{number of linearly independent columns of } T\\ &= \text{number of pivotal columns of } T.\end{aligned}$$

rank $T+$ nullity $T =$ dim domain T

2.5.5 By Definition 1.1.5 of a subspace, we need to show that the kernel and the image of a linear transformation T are closed under addition and multiplication by scalars. These are straightforward computations, using the linearity of T.

The kernel of T: If $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in \ker T$, i.e., if $T(\vec{\mathbf{v}}) = \vec{\mathbf{0}}$ and $T(\vec{\mathbf{w}}) = \vec{\mathbf{0}}$, then

$$T(\vec{\mathbf{v}}+\vec{\mathbf{w}}) = T(\vec{\mathbf{v}})+T(\vec{\mathbf{w}}) = \vec{\mathbf{0}}+\vec{\mathbf{0}} = \vec{\mathbf{0}} \quad \text{and} \quad T(a\vec{\mathbf{v}}) = aT(\vec{\mathbf{v}}) = a\vec{\mathbf{0}} = \vec{\mathbf{0}},$$

so $\vec{\mathbf{v}}+\vec{\mathbf{w}} \in \ker T$ and $a\vec{\mathbf{v}} \in \ker T$.

The image of T: If $\vec{\mathbf{v}} = T(\vec{\mathbf{v}}_1),\ \vec{\mathbf{w}} = T(\vec{\mathbf{w}}_1)$, then

$$\vec{\mathbf{v}}+\vec{\mathbf{w}} = T(\vec{\mathbf{w}}_1)+T(\vec{\mathbf{v}}_1) = T(\vec{\mathbf{w}}_1+\vec{\mathbf{v}}_1) \quad \text{and} \quad a\vec{\mathbf{v}} = aT(\vec{\mathbf{v}}_1) = T(a\vec{\mathbf{v}}_1).$$

So the image is also closed under addition and multiplication by scalars.

Solution 2.5.7: All triples of columns of A are linearly independent except columns 3, 4, and 6.

2.5.7 a. $n = 3$. The last three columns of the matrix are clearly linearly independent, so the matrix has rank at least 3, and it has rank at most 3 because there can be at most three linearly independent vectors in $\mathbb{R}^3$.

b. Yes. For example, the first three columns are linearly independent, since the matrix composed of just those columns row reduces to the identity.

c. The 3rd, 4th, and 6th columns are linearly dependent.

d. You cannot choose freely the values of x_1, x_2, x_5. Since the rank of the matrix is 3, three variables must correspond to pivotal (linearly independent) columns. For the variables x_1, x_2, x_5 to be freely chosen, i.e., nonpivotal, x_3, x_4, x_6 would have to correspond to linearly independent columns.

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

2.5.11 The sketch is shown in the margin.

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. By row operations, we can bring the matrix B to

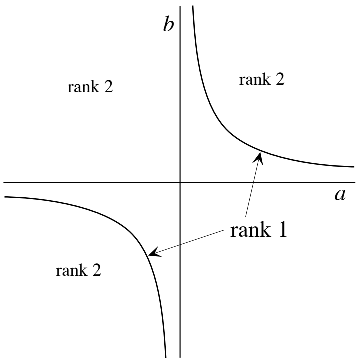
$$\begin{bmatrix} 1 & 2 & a \\ 0 & b & ab-a \\ 0 & 2a-b & a^2-a \end{bmatrix}.$$

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

- If $b = 0$, the matrix is $\begin{bmatrix} 1 & 2 & a \\ 0 & 0 & -a \\ 0 & 2a & a^2-a \end{bmatrix}$, which has rank 3 unless $a = 0$, in which case it has rank 1. (Of course, since $n = 3$, rank 3 corresponds to $\dim\ker = 0$, and rank 1 corresponds to $\dim\ker = 2$.)
- If $b \neq 0$, then we can do further row operations to bring the matrix to the form

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

The third entry in the third row factors as $\frac{a}{b}(a-b)(2-b)$, so we have rank 2 (and $\dim\ker = 1$) if $a = 0$ or $a = b$ or $b = 2$. Otherwise we have rank 3 (and $\dim\ker = 0$).



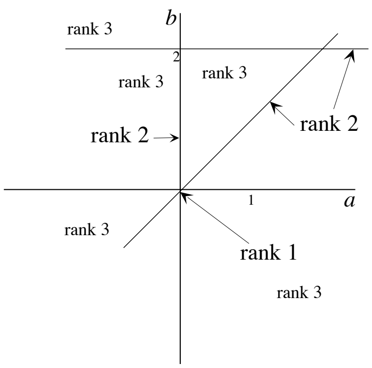


FIGURE FOR SOLUTION 2.5.11

TOP: On the hyperbola, 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 a-axis, the line $a = b$, and the line $b = 2$, the rank of B is 2, so 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.

2.5.13 a. As in Example 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 2.5.13 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.

You could avoid this by using a different factorization:

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

2.5.15 We give two solutions.

Solution 1

Since A and B are square, $C = (AB)^{-1}$ is square. We have $CAB = I$ and $ABC = I$, so (using the associativity of matrix multiplication) $(CA)B = I$, so CA is a left inverse of B, and $A(BC) = I$, so BC is a right inverse of A. Since (see the discussion after Corollary 2.2.7) a square matrix with a one-sided inverse (left or right) is invertible, and the one-sided inverse is the inverse, it follows that A and B are invertible.

Solution 2

The second solution reproves the fact that if a square matrix has a left inverse it has a right inverse, and the two inverses are equal.

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.

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 multiplication (Proposition 1.2.9).

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

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 $T_1 \circ T_2$.

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

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.

***2.5.17** 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 $b = 5$.

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.

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 there exists a solution of $M_{\mathbf{x}}(p) = \begin{bmatrix} a_0 \\ \vdots \\ a_n \end{bmatrix}$ and that it is unique.

c. Take the linear transformation $M'_{\mathbf{x}}$ from P_k to $\mathbb{R}^{2n+2}$ defined by $p \mapsto \begin{bmatrix} p(x_0) \\ \vdots \\ p(x_n) \\ p'(x_0) \\ \vdots \\ p'(x_n) \end{bmatrix}$. If ker $(M'_{\mathbf{x}}) = \{0\}$, 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.

2.5.19 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, there exist numbers $a_1, \dots, a_n$ not all zero such that

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

But we can write

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

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

***2.5.21** 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 $T_2(\vec{\mathbf{v}}) = \vec{\mathbf{0}}$, then (since S is linear), $(S \circ T_2)(\vec{\mathbf{v}}) = T_1(\vec{\mathbf{v}}) = \vec{\mathbf{0}}$.

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

For any $\vec{\mathbf{v}} \in \operatorname{img} T_2$, choose $\vec{\mathbf{v}}' \in \mathbb{R}^n$ such that $T_2(\vec{\mathbf{v}}') = \vec{\mathbf{v}}$, and set $S(\vec{\mathbf{v}}) = T_1(\vec{\mathbf{v}}')$. We need to show that this does not depend on the choice of $\vec{\mathbf{v}}'$. If $\vec{\mathbf{v}}'_1$ also satisfies $T_2(\vec{\mathbf{v}}'_1) = \mathbf{v}$, then $\vec{\mathbf{v}}' - \vec{\mathbf{v}}'_1 \in \ker T_2 \subset \ker T_1$, so $T_1(\vec{\mathbf{v}}') = T_1(\vec{\mathbf{v}}'_1)$, 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 the definition of image). If $T_1 = T_2 \circ S$, then $T_2(S(\vec{\mathbf{v}})) = \vec{\mathbf{w}}$, 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 there exists a linear transformation $S : \mathbb{R}^n \to \mathbb{R}^n$ 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 $\mathbf{v}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. We have $\begin{bmatrix} 2 & 1 \\ 5 & 4 \end{bmatrix} = 2\mathbf{v}_1 + \mathbf{v}_2 + 5\mathbf{v}_3 + 4\mathbf{v}_4$.

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

2.6.3

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

2.6.5 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

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

2.6.7 a. This is not a subspace, since 0 (the zero function) is not in it. (It is an affine subspace.)

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 subspace: 0 is in it, but that is not enough to make it a subspace. 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.

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

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

any vectors $\mathbf{w}_i$ 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 Definition 2.6.11.

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

2.7.1 The successive vectors $\vec{\mathbf{e}}_1$, $A\vec{\mathbf{e}}_1$, $A^2\vec{\mathbf{e}}_1, \dots$ are

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \dots.$$

The vectors $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}1\\1\end{bmatrix}$ are linearly independent, and

$$\begin{bmatrix}1\\2\end{bmatrix} = -\begin{bmatrix}1\\0\end{bmatrix} + 2\begin{bmatrix}1\\1\end{bmatrix}, \quad \text{or better,} \quad A^2\vec{\mathbf{e}}_1 - 2A\vec{\mathbf{e}}_1 + \vec{\mathbf{e}}_1 = \vec{\mathbf{0}}.$$

Thus the polynomial p_1 is $p_1(t) = t^2 - 2t + 1 = (t-1)^2$.

Solution 2.7.3: The notation $T(p) = p + xp''$ is standard, but it may be confusing, because when applied (for example) to the polynomial x, we have x playing a double role – it is both the polynomial and the name of the variable. It would be heavier but perhaps clearer to write

$$(T(p))(x) = p(x) + xp''(x).$$

Then, applied to the basis vectors 1, x, x^2, x^3, we have

$$\begin{aligned}(T(1))(x) &= 1 + x \cdot 0 = 1,\\ (T(x))(x) &= x + x \cdot 0 = x,\\ (T(x^2))(x) &= x^2 + x \cdot 2\\ (T(x^3))(x) &= x^3 + x \cdot 6x.\end{aligned}$$

2.7.3 a. We will give two solutions.

First solution: In the basis $\{\underline{\mathbf{x}}\} \overset{\text{def}}{=} (1, x, x^2, x^3)$, the matrix of T is

$$A = \begin{bmatrix}1&0&0&0\\0&1&2&0\\0&0&1&6\\0&0&0&1\end{bmatrix},$$

since

$$T(1) = 1,\ T(x) = x,\ T(x^2) = x^2 + 2x,\ T(x^3) = x^3 + 6x^2.$$

Denote the basis $p_1, \dots, p_4$ by $\{\underline{\mathbf{p}}\}$. The change of basis matrix $[P_{\underline{\mathbf{p}} \to \underline{\mathbf{x}}}]$ is

$$S = \begin{bmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{bmatrix}, \quad \text{with inverse} \quad S^{-1} = \begin{bmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{bmatrix},$$

giving for our desired matrix

$$S^{-1}AS = \begin{bmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&0&0&0\\0&1&2&0\\0&0&1&6\\0&0&0&1\end{bmatrix}\begin{bmatrix}1&1&1&1\\0&1&1&1\\0&0&1&1\\0&0&0&1\end{bmatrix} = \begin{bmatrix}1&0&-2&-2\\0&1&2&-4\\0&0&1&6\\0&0&0&1\end{bmatrix}.$$

Second solution: The other approach is to compute directly:

$$\begin{aligned}T(1) &= 1\\ T(1+x) &= 1+x\\ T(1+x+x^2) &= 1+x+x^2+2x = (1+x+x^2) + 2(x+1) - 2\\ T(1+x+x^2+x^3) &= 1+x+x^2+x^3+2x+6x^2\\ &= (1+x+x^2+x^3) + 6(x^2+x+1) - 4(x+1) - 2.\end{aligned}$$

This expresses the images of the basis vectors p_i under T as linear combinations of the p_i, so the coefficients are the columns of the desired matrix, giving again

$$\begin{bmatrix}1&0&-2&-2\\0&1&2&-4\\0&0&1&6\\0&0&0&1\end{bmatrix}.$$

b. We give a solution analogous to the first above. In the "standard basis" $1, x, \dots, x^k$ of P_k, the matrix of T is the diagonal matrix

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k+1 \end{bmatrix}.$$

The change of basis matrix is analogous to the above, and

$$\underbrace{\begin{bmatrix} 1 & -1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}}_{S^{-1}} \underbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k+1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}}_{S} = \begin{bmatrix} 1 & -1 & -1 & \dots & -1 \\ 0 & 2 & -1 & \dots & -1 \\ 0 & 0 & 3 & \dots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & k+1 \end{bmatrix},$$

where S has 1's on and above the diagonal and 0's elsewhere, and S^{-1} has 1's on the diagonal, -1's on the "superdiagonal" immediately above the main diagonal, and 0's elsewhere. The product has -1's everywhere above the diagonal, and $1, \dots, k+1$ on the diagonal.

2.7.5 a. The recursive formula for the b_n can be written

$$\begin{bmatrix} b_n \\ b_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} b_{n-1} \\ b_n \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}^n \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

To compute the power of the matrix, we use eigenvalues and eigenvectors. The eigenvalues are $1 \pm \sqrt{2}$, and a basis of eigenvectors is

$$\begin{bmatrix} 1 \\ 1+\sqrt{2} \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 1-\sqrt{2} \end{bmatrix}.$$

This leads to the change of basis matrix

$$S = \begin{bmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{bmatrix}, \quad \text{with inverse} \quad S^{-1} = \frac{\sqrt{2}}{4} \begin{bmatrix} \sqrt{2}-1 & 1 \\ \sqrt{2}+1 & -1 \end{bmatrix}.$$

Indeed, you can check by multiplication that

$$\frac{\sqrt{2}}{4} \begin{bmatrix} \sqrt{2}-1 & 1 \\ \sqrt{2}+1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1+\sqrt{2} & 1-\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}.$$

Thus

$$S^{-1} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}^n S = \begin{bmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{bmatrix}^n = \begin{bmatrix} (1+\sqrt{2})^n & 0 \\ 0 & (1-\sqrt{2})^n \end{bmatrix},$$

and finally

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}^n &= S \begin{bmatrix} (1+\sqrt{2})^n & 0 \\ 0 & (1-\sqrt{2})^n \end{bmatrix} S^{-1} \\ &= \frac{\sqrt{2}}{4} \begin{bmatrix} (\sqrt{2}-1)(1+\sqrt{2})^n + (\sqrt{2}+1)(1-\sqrt{2})^n & (1+\sqrt{2})^n - (1-\sqrt{2})^n \\ (\sqrt{2}-1)(1+\sqrt{2})^{n+1} + (\sqrt{2}+1)(1-\sqrt{2})^{n+1} & (1+\sqrt{2})^{n+1} - (1-\sqrt{2})^{n+1} \end{bmatrix}. \end{aligned}$$

Multiplying this by $\begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ we find

$$b_n = \frac{(1+\sqrt{2})^n + (1-\sqrt{2})^n}{2}$$

b. Since $|1-\sqrt{2}| < 1$, it will contribute practically nothing to b_{1000}, and $\frac{1}{2}(1+\sqrt{2})^{1000}$ has the same number of digits and the same leading digits as b_{1000}. You will find that your calculator will refuse to evaluate this, but using logarithms base 10 for a change, you find

$$\log_{10} b_{1000} \approx 382.475,$$

so b_{1000} has 382 digits, starting with 2983.

We have

$$A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 2 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 2 & 3 & 4 \end{bmatrix}$$

c. This time the relevant matrix description of the recursion is

$$\begin{bmatrix} c_n \\ c_{n+1} \\ c_{n+2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_{n-1} \\ c_n \\ c_{n+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}^n \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Set $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. One can easily check that

$$A^3 - A^2 - A = I,$$

so the roots of the polynomial $\lambda^3 - \lambda^2 - \lambda - 1$ are eigenvalues of the matrix.[1] These eigenvalues are approximately

$$\lambda_1 \approx 1.83926\ldots \quad \text{and} \quad \lambda_{2,3} \approx -.419643 \pm .6062907.$$

A basis of eigenvectors is

$$S = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix}$$

and we can compute the high powers as above by

$$A^n = S \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix} S^{-1}.$$

[1] How did we find the coefficients $x = y = z = -1$ for the equation

$$A^3 + xA^2 + yA + zI = 0?$$

This equation corresponds to

$$\begin{bmatrix} 1+z & 1+y & 1+x \\ 1+x & 2+x+z & 2+x+y \\ 2+x+y & 3+2x+y & 4+2x+y+z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

A glance just at the top line immediately gives the answer. Normally one would not expect a system of nine equations in three unknowns to have a solution! The Cayley-Hamilton theorem (Theorem 4.8.27) says that every $n \times n$ matrix satisfies a polynomial equation of degree at most n, whereas you would expect that it would only satisfy an equation of degree n^2.

Computing S^{-1} is what computers are for, but when the dust has settled, we find

$$S^{-1}\begin{bmatrix}1\\1\\1\end{bmatrix} \approx \begin{bmatrix}.4356\\.28219-.359i\\.28219+.359i\end{bmatrix},$$

and, as above,

$$c_n \approx .4356\cdot\lambda_1^n + (.28219-.359i)\cdot\lambda_2^n + (.28219+.359i)\cdot\lambda_3^n.$$

Finally, to find the number of digits of c_{1000}, only the term involving λ_1 is relevant, and we resort to logarithms base 10:

$$\log .4356 + 1000\log 1.83926 \approx 264.282.$$

Thus c_{1000} has 265 digits.

Inequality 1 follows from the second paragraph of the proof of Theorem 2.7.10. Inequality 2 is true because A and B have only positive entries, $A \geq B$, and $\vec{\mathbf{v}}_B$ has only positive entries.

2.7.7 a. Let A and B be square matrices with λ_A the leading eigenvalue of A and λ_B the leading eigenvalue of B. First let us see that if $A \geq B > \mathbf{0}$, then $\lambda_A \geq \lambda_B$. Choose strictly positive unit eigenvectors $\vec{\mathbf{v}}_A$ and $\vec{\mathbf{v}}_B$ with eigenvalues λ_A and λ_B; then $\lambda_A \geq \lambda_B$:

$$\lambda_A = |\lambda_A\vec{\mathbf{v}}_A| = |A\vec{\mathbf{v}}_A| \overbrace{\geq}^{1} |A\vec{\mathbf{v}}_B| \overbrace{\geq}^{2} |B\vec{\mathbf{v}}_B| = |\lambda_B\vec{\mathbf{v}}_B| = \lambda_B.$$

(This result is of interest in its own right.)

Let A_n be the matrix obtained from A by adding $1/n$ to every entry. Then $A_n > \mathbf{0}$ and by Theorem 2.7.10, A_n has an eigenvector $\mathbf{v}_n \in \mathring{\Delta}$, where Δ is the set of unit vectors in the positive quadrant Q. Moreover, by the argument above, the corresponding eigenvalues λ_n are nonincreasing and positive, so they have a limit $\lambda \geq 0$.

Since Δ is compact, the sequence $n \mapsto \mathbf{v}_n$ has a convergent subsequence that converges, say to $\mathbf{v} \in \Delta$ (not necessarily in $\mathring{\Delta}$). Passing to the limit in the equation $A_n\mathbf{v}_n = \lambda_n\mathbf{v}_n$, we see that $A\mathbf{v} = \lambda\mathbf{v}$.

b. By part a, the matrix A has an eigenvector $\mathbf{v} \in \Delta$ with eigenvalue $\lambda \geq 0$. The vector $\mathbf{v}$ is still an eigenvector for A^n, this time with eigenvalue λ^n. Thus in this case where $A^n > 0$, we must have $\mathbf{v} \in \mathring{\Delta}$, i.e., $\mathbf{v} > 0$, and $\lambda^n > 0$, hence $\lambda > 0$. Moreover, if μ is any other eigenvalue of A, then μ^n is an eigenvalue of A^n, so $|\mu|^n < \lambda^n$, hence $|\mu| < \lambda$.

2.8.1 The partial derivatives are

$$\begin{aligned} &D_1F_1 = -\sin(x-y),\ D_2F_1 = \sin(x-y) - 1,\\ &D_1F_2 = \cos(x+y) - 1,\ D_2F_2 = \cos(x+y), \end{aligned}$$

so

$$\begin{aligned} &D_{1,1}F_1 = D_{2,2}F_1 = -\cos(x-y); \qquad D_{1,2}F_1 = D_{2,1}F_1 = \cos(x-y)\\ &D_{1,1}F_2 = D_{2,2}F_2 = D_{1,2}F_2 = D_{2,1}F_2 = -\sin(x+y). \end{aligned}$$

Therefore, the absolute value of each second partial is bounded by 1, and there are eight second partials, so

$$\big|[\mathbf{DF}(\mathbf{u})] - [\mathbf{DF}(\mathbf{v})]\big| \le \sqrt{8 \cdot 1^2}|\mathbf{u} - \mathbf{v}|;$$

using this method we get $M = 2\sqrt{2}$, the same result as equation 2.8.62.

2.8.3 a. Use the triangle inequality:

$$\begin{matrix} |x| = |x-y+y| \le |x-y| + |y| \\ |y| = |y-x+x| \le |y-x| + |x| \end{matrix} \quad \text{imply} \quad \begin{matrix} |x| - |y| \le |x-y| \\ |y| - |x| \le |x-y| \end{matrix} \quad \text{imply} \quad \big||x| - |y|\big| \le |x-y|.$$

b. Choose $C > 0$, and find x such that $1/(2\sqrt{x}) > C$, which will happen for $x > 0$ sufficiently small. From the definition of the derivative,

$$\frac{1}{2\sqrt{x}} = \lim_{y \to x} \frac{\sqrt{y} - \sqrt{x}}{y - x}.$$

Thus taking $y > x$ sufficiently close to x, we will find

$$|\sqrt{y} - \sqrt{x}| \ge (C-1)|y-x|.$$

Since this is possible for every $C > 0$, we see that there cannot exist a number M such that $|\sqrt{y} - \sqrt{x}| \le M|y-x|$ for all $x, y > 0$.

2.8.5 a. Newton's method gives $x_{n+1} = \dfrac{2x_n^3 + 9}{3x_n^2}$, which leads to

$$x_0 = 2,\ x_1 = \frac{25}{12} = 2.08\overline{3},\ x_2 = \frac{46802}{22500} = 2.0800\overline{8},$$
$$x_3 = \frac{307548373003216}{147853836270000} \approx 2.08008382306424.$$

b. In this case,

$$h_0 = -\frac{2^3 - 9}{2 \cdot 2^2} = \frac{1}{12}, \quad U_1 = \left[2, 2 + \frac{1}{6}\right].$$

We have $f(x_0) = 2^3 - 9 = -1$, $f'(x_0) = 3 \cdot 2^2 = 12$, and finally

$$M = \sup_{x \in U_1} |f''(x)| = 6\left(2 + \frac{1}{6}\right) = 13.$$

c. Just compute:

$$\frac{|f(x_0)|M}{|f'(x_0)|^2} = \frac{1 \cdot 13}{12^2} = 0.090\overline{27} < \frac{1}{2},$$

so we know that Newton's method will converge – not that there was much doubt after part a.

2.8.7 a. The MATLAB command

```
newton([cos(x1)+x2-1.1; x1+cos(x1+x2)-.9], [0; 0], 5)
```

produces the output

$$\begin{pmatrix} -0.10000000000000 \\ 0.10000000000000 \end{pmatrix}, \begin{pmatrix} -0.10000000000000 \\ 0.10499583472197 \end{pmatrix}, \begin{pmatrix} -0.09998746447058 \\ 0.10499458325724 \end{pmatrix},$$
$$\begin{pmatrix} -0.09998746440624 \\ 0.10499458332900 \end{pmatrix}, \begin{pmatrix} -0.09998746440624 \\ 0.10499458332900 \end{pmatrix}.$$

We see that the first nine significant digits are unchanged from the third to the fourth step, and that all 15 digits are preserved from the fourth to the fifth step.

b. Set

$$\mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} \cos x + y - 1.1 \\ x + \cos(x+y) - .9 \end{bmatrix}.$$

Then

$$\left[\mathbf{Df}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{bmatrix} -\sin x & 1 \\ 1 - \sin(x+y) & -\sin(x+y) \end{bmatrix},$$

so

$$\left[\mathbf{Df}\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{h}_0 = \mathbf{a}_1 = \begin{bmatrix} 0.1 \\ -0.1 \end{bmatrix}.$$

Now we estimate the Lipschitz ratio:

$$\left|\begin{bmatrix} -\sin u_1 & 1 \\ 1 - \sin(u_1+v_1) & -\sin(u_1+v_1) \end{bmatrix} - \begin{bmatrix} -\sin u_2 & 1 \\ 1 - \sin(u_2+v_2) & -\sin(u_2+v_2) \end{bmatrix}\right|$$
$$\leq \left|\begin{bmatrix} |u_1 - u_2| & 0 \\ |(u_1-u_2)+(v_1-v_2)| & |(u_1-u_2)+(v_1-v_2)| \end{bmatrix}\right|$$
$$= \left((u_1-u_2)^2 + 2((u_1-u_2)+(v_1-v_2))^2\right)^{1/2} \leq \sqrt{5}\left((u_1-u_2)^2 + (v_1-v_2)^2\right)^{1/2},$$

so we may take $M = \sqrt{5}$.

The quantity

$$M|\mathbf{f}(\mathbf{a}_0)|\left|[\mathbf{D}\vec{\mathbf{f}}(\mathbf{a}_0)]^{-1}\right|^2 = \sqrt{5}\frac{\sqrt{2}}{10}(\sqrt{2})^2 = \frac{2\sqrt{10}}{10} \approx .632\ldots$$

is not smaller than 1/2, so we cannot simply invoke Kantorovich.

Using the norm rather than the length is justified by Theorem 2.9.8. The norm is defined in Definition 2.9.6.

There are at least two ways around this. One is to use the norm of the derivative rather than the length: the norm of $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is 1, not $\sqrt{2}$, so this improves the estimate by a factor of 2, which is enough.

Another is to look at the next step of Newton's method, which gives $|\mathbf{f}(\mathbf{a}_1)| \approx .005$. We can use the same Lipschitz ratio. The derivative $L = [\mathbf{Df}(\mathbf{a}_1)]$ is, by the same Lipschitz estimate, of the form

$$\begin{bmatrix} a & 1+b \\ 1+c & d \end{bmatrix}$$

with $|a|$, $|b|$, $|c|$, $|d|$ all less than 0.1. So $|\det L| > .75$, and we have $|L^{-1}| < 2.2/.75 < 3$. This gives

$$M|\mathbf{f}(\mathbf{a}_1)|\left|L^{-1}\right|^2 < \sqrt{5}\cdot(.005)\cdot 9,$$

which is much less than $1/2$.

2.8.9 We will start Newton's method at the origin, and try to solve the system of equations

$$F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y^2-a \\ y+z^2-b \\ z+x^2-c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

First, a bit of computation will show you that $M=2$ is a global Lipschitz ratio for $[\mathbf{D}F(\mathbf{x})]$. Next, the derivative at the origin is the identity, so its inverse is also the identity, with length $\sqrt{3}$. So we find that the condition for Kantorovich's theorem to work is

$$\sqrt{a^2+b^2+c^2}\cdot(\sqrt{3})^2\cdot 2 \le \frac{1}{2}, \quad \text{i.e.,} \quad \sqrt{a^2+b^2+c^2} \le \frac{1}{12}.$$

Solution 2.8.9: This is an exceptionally easy case in which to apply Kantorovich's theorem.

What would we get if we used the norm (discussed in Section 2.9) instead of the length? Since the norm of the identity is 1, we would get $\sqrt{a^2+b^2+c^2} \le \frac{1}{4}$. We don't know how sharp this is, but $1/2$ is certainly too big.

2.8.11 a. Set $p(x)=x^5-x-6$. One step of Newton's method to solve $p(x)=0$ is

$$N(a) = a - \frac{a^5-a-6}{5a^4-1} = \frac{4a^5+6}{5a^4-1}.$$

In particular, $x_1 = N(2) = \frac{134}{79}$, with a first Newton step $h_0 = -24/79$.

b. To apply Kantorovitch's theorem, we need to show that

$$\frac{Mp(2)}{(p'(2))^2} < 1/2, \quad \text{where } M = \sup_{x\in[110/79,\,2]} |p''(x)|. \tag{1}$$

The second derivative of p is $20x^3$; since $20x^3$ is increasing, the supremum M is $20\cdot 2^3 = 160$. We have $p(2)=24$ and $p'(2)=79$. Unfortunately,

$$\frac{24\cdot 160}{79^2} \approx .6152860119 > \frac{1}{2},$$

so Kantorovitch's theorem does not apply. But it does apply at x_1. Using decimal approximations, we find $p(x_1)\approx 1.09777$ and $p'(x_1)\approx 27.0579$. The supremum M_1 of $p''(x)$ over $\overline{U}_2$ is

$$M_1 = p''(x_1) = p''\left(\frac{134}{79}\right) = 20\left(\frac{134}{79}\right)^3 \approx 98, \quad \text{giving}$$

$$\frac{M_1 p(x_1)}{(p'(x_1))^2} \approx \frac{98\cdot 1.1}{(27)^2} \approx .148 < 1/2.$$

Solution 2.8.11, part b: Since $U_1 = B_{|\vec{\mathbf{h}}_0|}(x_1)$, the sup in equation (1) for M is over

$$\begin{aligned}\overline{U}_1 &= [x_1 - |\vec{\mathbf{h}}_0|, x_1 + |\vec{\mathbf{h}}_0|] \\ &= \left[\frac{134}{79} - \frac{24}{79}, \frac{134}{79} + \frac{24}{79}\right] \\ &= \left[\frac{110}{79}, \frac{158}{79}\right] = \left[\frac{110}{79}, 2\right].\end{aligned}$$

For an initial guess x_0, Kantorovich's theorem requires a Lipschitz ratio in $\overline{U}_1$; for an initial guess x_1, it requires a Lipschitz ratio M_1 in $\overline{U}_2$.

The original M works also:

$$\begin{aligned}\frac{Mp(x_1)}{(p'(x_1))^2} &\approx \frac{160\cdot 1.097}{724} \\ &\approx .24 < 1/2.\end{aligned}$$

2.8.13 a. Direct calculation gives

$$\begin{aligned}|[\mathbf{DF}(\mathbf{u})] - [\mathbf{DF}(\mathbf{v})]| &= \left|\begin{bmatrix} 2u_1-2 & 2u_2 \\ u_1-1 & u_2-1 \end{bmatrix} - \begin{bmatrix} 2v_1-2 & 2v_2 \\ v_1-1 & v_2-1 \end{bmatrix}\right| \\ &= \left|\begin{bmatrix} 2(u_1-v_1) & 2(u_2-v_2) \\ u_1-v_1 & u_2-v_2 \end{bmatrix}\right| = \sqrt{5}|\mathbf{u}-\mathbf{v}|.\end{aligned}$$

So $\sqrt{5}$ is a global Lipschitz ratio. The second derivative approach gives $\sqrt{10}$.

b. We have

$$\mathbf{F}\begin{pmatrix}5\\1\end{pmatrix}=\begin{bmatrix}1\\-1\end{bmatrix},\quad \left[\mathbf{DF}\begin{pmatrix}5\\1\end{pmatrix}\right]=\begin{bmatrix}8&2\\0&4\end{bmatrix},\quad \left[\mathbf{DF}\begin{pmatrix}5\\1\end{pmatrix}\right]^{-1}=\frac{1}{32}\begin{bmatrix}4&-2\\0&8\end{bmatrix},$$

so

$$\vec{\mathbf{h}}_0=-\frac{1}{32}\begin{bmatrix}4&-2\\0&8\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}-3/16\\1/4\end{bmatrix}$$

$$\mathbf{x}_1=\begin{pmatrix}5\\1\end{pmatrix}+\begin{bmatrix}-3/16\\1/4\end{bmatrix}=\begin{pmatrix}77/16\\5/4\end{pmatrix},$$

and $|\vec{\mathbf{h}}_0|=5/16$. For Kantorovich's theorem to guarantee convergence, we must have

$$\sqrt{2}\frac{84}{32^2}\sqrt{5}\le\frac{1}{2},$$

which is true (and is still true if we use $\sqrt{10}$ instead of $\sqrt{5}$).

c. By part b, there is a unique root in the disc of radius 5/16 around

$$\begin{pmatrix}77/16\\5/4\end{pmatrix}\approx\begin{pmatrix}4.8\\1.25\end{pmatrix}.$$

See the picture in the margin.

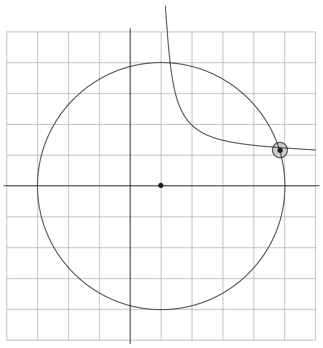


FIGURE FOR SOLUTION 2.8.13. The curve of equation

$$x^2+y^2-2x-15=0$$

is the circle of radius 4 centered at $\binom{1}{0}$. The curve of equation

$$xy-x-y=0$$

is a hyperbola with asymptotes the lines $x=1$ and $y=1$. Exercise 2.8.13 uses Newton's method to find an intersection of these curves. There are four such intersections, but starting at $\binom{5}{1}$ leads to a sequence that rapidly converges to the intersection in the small shaded circle.

2.8.15 The system of equations to be solved is

$$\mathbf{F}\left(\begin{pmatrix}y\\z\end{pmatrix},\lambda\right)=\begin{aligned}1+a-\lambda+a^2y&=0\\a+(2-\lambda)y+az&=0\\-a+(3-\lambda)z&=0\end{aligned}$$

for the unknowns $(\begin{pmatrix}y\\z\end{pmatrix},\lambda)$, starting at $(\begin{pmatrix}0\\0\end{pmatrix},1)$; the number a is a parameter. We have

$$\mathbf{F}\left(\begin{pmatrix}0\\0\end{pmatrix},1\right)=\begin{pmatrix}a\\a\\-a\end{pmatrix},\quad \left|\mathbf{F}\left(\begin{pmatrix}0\\0\end{pmatrix},1\right)\right|=|a|\sqrt{3}.$$

Now compute

$$\left[\mathbf{DF}\left(\begin{pmatrix}y\\z\end{pmatrix},\lambda\right)\right]=\begin{bmatrix}a^2&0&-1\\2-\lambda&a&-y\\0&(3-\lambda)&-z\end{bmatrix},\left[\mathbf{DF}\left(\begin{pmatrix}0\\0\end{pmatrix},1\right)\right]=\begin{bmatrix}a^2&0&-1\\1&a&0\\0&2&0\end{bmatrix}.$$

The inverse $\left[\mathbf{DF}\left(\begin{pmatrix}0\\0\end{pmatrix},1\right)\right]^{-1}$ is computed by row reducing

$$\begin{bmatrix}a^2&0&-1&1&0&0\\1&a&0&0&1&0\\0&2&0&0&0&1\end{bmatrix}\quad\text{to get}\quad\begin{bmatrix}1&0&0&0&1&-\frac{a}{2}\\0&1&0&0&0&\frac{1}{2}\\0&0&1&-1&a^2&-\frac{a^3}{2}\end{bmatrix}$$

This gives

$$\left|\left[\mathbf{DF}\left(\begin{pmatrix}0\\0\end{pmatrix},1\right)\right]^{-1}\right|^2=1+\frac{a^2}{4}+\frac{1}{4}+1+a^4+\frac{a^6}{4}=\frac{9}{4}+\frac{a^2}{4}+a^4+\frac{a^6}{4}.$$

Finally,

$$\begin{aligned}
&\left|\left[\mathbf{DF}\left(\begin{pmatrix} y_1 \\ z_1 \end{pmatrix}, \lambda_1\right)\right] - \left[\mathbf{DF}\left(\begin{pmatrix} y_2 \\ z_2 \end{pmatrix}, \lambda_2\right)\right]\right| \\
&\qquad = \left|\begin{bmatrix} 0 & 0 & 0 \\ -\lambda_1+\lambda_2 & 0 & y_2-y_1 \\ 0 & -\lambda_1+\lambda_2 & -z_2+z_1 \end{bmatrix}\right| \\
&\qquad = \sqrt{2(\lambda_1-\lambda_2)^2+(y_1-y_2)^2+(z_1-z_2)^2} \\
&\qquad \le \sqrt{2}\left|\left(\begin{pmatrix} y_1 \\ z_1 \end{pmatrix}, \lambda_1\right) - \left(\begin{pmatrix} y_2 \\ z_2 \end{pmatrix}, \lambda_2\right)\right|
\end{aligned}$$

So the inequality in Kantorovich's theorem will be satisfied if

$$|a|\sqrt{3}\left(\frac{9}{4}+\frac{a^2}{4}+a^4+\frac{a^6}{4}\right)\sqrt{2} \le \frac{1}{2}.$$

Since the expression on the left is an increasing function of $|a|$, if it is satisfied for some particular $a_0 > 0$, it will be satisfied for $|a| < a_0$. The best a_0 is the solution of this 7th degree equation, $a_0 \approx .0906$; arithmetic will show that $a_0 = .09$ does satisfy the inequality.

2.9.1 For the polynomial $f(x) = (x-1)^2$, it is always true that

$$\left|f'(a) - f'(b)\right| = \left|2(a-1) - 2(b-1)\right| = 2|a-b|,$$

so $M = 2$ is a Lipschitz ratio, and certainly the best ratio, realized at every pair of points.

Since

$$\frac{f(0)M}{f'(0)^2} = \frac{1\cdot 2}{(-2)^2} = \frac{1}{2},$$

inequality 2.8.53 of Kantorovich's theorem is satisfied as an equality.

One step of Newton's method is

$$a_{n+1} = a_n + \frac{(1-a_n)^2}{2(1-a_n)} = \frac{a_n+1}{2},$$

so if we assume by induction that $a_n = 1 - 1/2^n$, which is true for $n = 0$, we find

$$a_{n+1} = \frac{1-1/2^n+1}{2} = 1 - 1/2^{n+1},$$

and $h_n = 1/2^{n+1}$, which is exactly the worst case in Kantorovich's theorem.

Solution 2.9.3: In the second inequality, we choose one unit $\vec{\mathbf{v}}$ to make $|A\vec{\mathbf{v}}|$ as large as possible, and another to make $|B\vec{\mathbf{v}}|$ as large as possible. The sum we get this way is certainly at least as large as the sum we would get if we chose the same $\vec{\mathbf{v}}$ for both terms.

2.9.3 Just apply the definition of the norm, and the triangle inequality in the codomain:

$$\begin{aligned}
\|A+B\| &= \sup_{|\vec{\mathbf{v}}|=1} |(A+B)\vec{\mathbf{v}}| = \sup_{|\vec{\mathbf{v}}|=1} |A\vec{\mathbf{v}} + B\vec{\mathbf{v}}| \\
&\le \sup_{|\vec{\mathbf{v}}|=1} |A\vec{\mathbf{v}}| + |B\vec{\mathbf{v}}| \le \sup_{|\vec{\mathbf{v}}|=1} |A\vec{\mathbf{v}}| + \sup_{|\vec{\mathbf{v}}|=1} |B\vec{\mathbf{v}}| = \|A\| + \|B\|.
\end{aligned}$$

2.9.5 a. In part b, we will solve this equation by Newton's method. For now, you might guess that there exists a solution of the form $\begin{bmatrix} x & x \\ x & x \end{bmatrix}$. Substitute that into the equation, to find

$$\begin{bmatrix} 2x^2 & 2x^2 \\ 2x^2 & 2x^2 \end{bmatrix} + \begin{bmatrix} x & x \\ x & x \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \text{i.e.,} \quad 2x^2 + x - 1 = 0.$$

This has two solutions, -1 and $1/2$, which lead to two solutions of the original problem:

$$\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

b. We will use the *length* of matrices, but the *norm* for elements of $\mathcal{L}(\text{Mat}\,(2,2), \text{Mat}\,(2,2))$: i.e., for linear maps from Mat $(2,2)$ to Mat $(2,2)$.

The equation we want to solve is $X^2 + X - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, so the relevant mapping is the (nonlinear) mapping $F : \text{Mat}\,(2,2) \to \text{Mat}\,(2,2)$ given by

Solution 2.9.5, part b: If we wanted to consider $[\mathbf{D}F(I)]$ and $[\mathbf{D}F(I)]^{-1}$ as matrices, we would have to consider an element of Mat $(2,2)$ as an element of $\mathbb{R}^4$ and the matrices in question would be 4×4, with 16 entries, making computing the length cumbersome.

$$F(X) = X^2 + X - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Clearly $|F(I)| = \left| \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \right| = 2$. Since

$$[\mathbf{D}F(I)](H) = HI + IH + H = 3H,$$

we have $[\mathbf{D}F(I)]^{-1}H = H/3$, i.e., $[\mathbf{D}F(I)]^{-1}$ is $1/3$ times the identity function $\text{id}: \text{Mat}\,(2,2) \to \text{Mat}\,(2,2)$. Thus

$$\left\| [\mathbf{D}F(I)]^{-1} \right\| = \frac{1}{3},$$

which squared gives $1/9$.

Moreover, $\|[\mathbf{D}F(A)] - [\mathbf{D}F(B)]\| \le 2|A - B|$:

Inequality 1: The first inequality is the triangle inequality; the second inequality is Schwarz. Remember that we are taking the sup of H such that $|H| = 1$.

$$\begin{aligned} \Big\| [\mathbf{D}F(A)] - [\mathbf{D}F(B)] \Big\| &= \sup_{|H|=1} \Big| (AH + HA + H) - (BH + HB + H) \Big| \\ &= \sup_{|H|=1} \Big| H(A - B) + (A - B)H \Big| \\ &\le \sup_{|H|=1} \Big(\big|H(A - B)\big| + \big|(A - B)H\big| \Big) \qquad (1) \\ &\le \sup_{|H|=1} \Big(|H||A - B| + |A - B||H| \Big) = 2|A - B|. \end{aligned}$$

So the Kantorovich inequality is satisfied:

$$2 \cdot 1/9 \cdot 2 < 1/2.$$

2.9.7 First, let us compute the norm of $f(A_0) = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$:

$$\begin{aligned} \|f(A_0)\| &= \sup_{|\vec{\mathbf{x}}|=1} \left| \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right| = \sup_{|\vec{\mathbf{x}}|=1} \left| \begin{bmatrix} x_1 - x_2 \\ x_1 - x_2 \end{bmatrix} \right| \\ &= \sup_{|\vec{\mathbf{x}}|=1} \sqrt{2(x_1 - x_2)^2} \\ &= \sup_{|\vec{\mathbf{x}}|=1} \sqrt{2(x_1^2 + x_2^2) - 4x_1x_2}. \end{aligned}$$

If we substitute $\cos\theta$ for x_1 and $\sin\theta$ for x_2 (since $x_1^2 + x_2^2 = 1$) this gives

$$\sup \sqrt{2(\cos^2\theta + \sin^2\theta) - 4\cos\theta\sin\theta} = \sup \sqrt{2 - 2\sin 2\theta} = \sqrt{4} = 2.$$

The following (virtually identical to equation 2.9.23) shows that the derivative of f is Lipschitz with Lipschitz ratio 2:

$$\begin{aligned} \|[\mathbf{D}f(A_1)] - [\mathbf{D}f(A_2)]\| &= \sup_{|B|=1} \left| \Big([\mathbf{D}f(A_1)] - [\mathbf{D}f(A_2)] \Big) B \right| \\ &= \sup_{|B|=1} \left| A_1B + BA_1 + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - A_2B - BA_2 - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right| \\ &= \sup_{|B|=1} |(A_1 - A_2)B + B(A_1 - A_2)| \\ &\le \sup_{|B|=1} |A_1 - A_2||B| + |B||A_1 - A_2| \\ &\le \sup_{|B|=1} 2|B||A_1 - A_2| \\ &= 2|A_1 - A_2|. \end{aligned}$$

Finally we need to compute the norm of $[\mathbf{D}f(A_0)]^{-1} = \frac{1}{63}\begin{bmatrix} 8 & -1 \\ -1 & 8 \end{bmatrix}$:

$$\begin{aligned} \left\| \frac{1}{63} \begin{bmatrix} 8 & -1 \\ -1 & 8 \end{bmatrix} \right\| &= \sup_{|\vec{\mathbf{x}}|=1} \left| \begin{bmatrix} \frac{8x_1}{63} - \frac{x_2}{63} \\ -\frac{x_1}{63} + \frac{8x_2}{63} \end{bmatrix} \right| \\ &= \sup_{|\vec{\mathbf{x}}|=1} \sqrt{\left(\frac{8x_1 - x_2}{63} \right)^2 + \left(\frac{8x_2 - x_1}{63} \right)^2} \\ &= \sup_{|\vec{\mathbf{x}}|=1} \frac{1}{63} \sqrt{64x_1^2 - 16x_1x_2 + x_2^2 + 64x_2^2 - 16x_2x_1 + x_1^2} \\ &= \sup_{|\vec{\mathbf{x}}|=1} \frac{1}{63} \sqrt{(x_1^2 + x_2^2) + 64(x_1^2 + x_2^2) - 32x_1x_2}. \end{aligned}$$

Switching again to sines and cosines, we get

$$\sup \frac{1}{63} \sqrt{65(\cos^2\theta + \sin^2\theta) - 16\sin 2\theta} = \frac{1}{63}\sqrt{65 + 16} = \frac{9}{63} = \frac{1}{7}.$$

Squaring this norm gives $\|[\mathbf{D}f(A_0)]^{-1}\|^2 \approx 0.0204$.

Thus instead of

$$\underbrace{2}_{|f(A_0)|} \cdot \underbrace{2.8}_{M} \cdot \underbrace{(.256)^2}_{|[\mathbf{D}f(A_0)]^{-1}|^2} = .367 < .5,$$

obtained using the length (see equation 2.8.73 in the text), we have

$$\underbrace{2}_{|f(A_0)|} \cdot \underbrace{2}_{M} \cdot \underbrace{0.0204}_{\|[\mathbf{D}f(A_0)]^{-1}\|^2} = 0.08166.$$

Solution 2.10.1: Remember, the inverse function theorem says that *if the derivative of a mapping is invertible, the mapping is locally invertible.*

2.10.1 Parts a, c, and d: The theorem does not guarantee that these functions are locally invertible with differentiable inverse, since the derivative is not invertible: for the first two, because the derivative is not square; for the third, because all entries of the derivative are 0. (In fact, Exercise 2.10.4 shows that the functions are definitely not locally invertible with differentiable inverse.)

b. Invertible: the derivative at $\begin{pmatrix}1\\1\end{pmatrix}$ is $\begin{bmatrix}2 & 1\\-2 & 0\end{bmatrix}$, which is invertible; hence the mapping is invertible.

e. The same mapping as in part d is invertible at $\begin{pmatrix}1\\1\\1\end{pmatrix}$; the derivative at that point is $\begin{bmatrix}1 & 1 & 1\\2 & 0 & 0\\0 & 0 & 2\end{bmatrix}$, which is invertible.

2.10.3 a. Because $\sin 1/x$ isn't defined at $x = 0$, to see that this function is differentiable at 0, you *must* apply the definition of the derivative:

$$f'(0) = \lim_{h\to 0} \frac{h/2 + h^2 \sin(1/h)}{h} = \frac{1}{2} + \lim_{h\to 0} h \sin\frac{1}{h} = \frac{1}{2}.$$

b. Away from 0, you can apply the ordinary rules of differentiation, to find

$$f'(x) = \frac{1}{2} + 2x\sin\frac{1}{x} + x^2\cos\frac{1}{x}\left(-\frac{1}{x^2}\right) = \frac{1}{2} + 2x\sin\frac{1}{x} - \cos\frac{1}{x}.$$

As x approaches 0, the term $2x\sin\frac{1}{x}$ approaches 0, but the term $\cos\frac{1}{x}$ oscillates infinitely many times between 1 and -1. Even if you add 1/2, there still are x arbitrarily close to 0 where $f'(x) < 0$. Since f is decreasing at these points, the function f is not monotone in any neighborhood of 0, and thus it cannot have an inverse on any neighborhood of 0.

c. It doesn't contradict Theorem 2.10.2 because f is not increasing or decreasing. More important, it doesn't contradict Theorem 2.10.7 (the inverse function theorem), because that theorem requires that the derivative be continuous, and although f is differentiable, the derivative is not continuous.

2.10.5 a. Two different implicit functions exist, both for $x \le -1/4$. We find them by computing

$$y^2 + y + 3x + 1 = (y^2 + y + 1/4) + 3x + 3/4 = 0, \quad \text{so that}$$
$$(y + 1/2)^2 = -3(x + 1/4),$$

giving the implicit functions $y = \pm\sqrt{-3x - 3/4} - 1/2$. Since $-3x - 3/4$ can't be negative, we have $x \le -1/4$.

Solution 2.10.5, part b: Does this seem contradictory? After all, we know that given $x = -1/4$ we can use

$$4y = \pm\sqrt{-3x - 3/4} - 1/2$$

to compute $y = -1/2$, and here is the implicit function theorem telling us there's no implicit function there! But the implicit function theorem doesn't talk about implicit functions at points, it talks about neighborhoods. At $x = -1/4 + \epsilon$, with $\epsilon > 0$, there is no implicit function.

b. To check that this result agrees with the implicit function theorem, we compute the derivative of $f\begin{pmatrix} x \\ y \end{pmatrix} = y^2 + y + 3x + 1$, getting

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \end{pmatrix}\right] = [3 \quad 2y + 1].$$

Defining y implicitly as a function of x means considering y the pivotal variable, so the matrix corresponding to $[D_1\mathbf{F}(\mathbf{c}), \dots, D_{n-k}\mathbf{F}(\mathbf{c})]$ is $[2y+1]$, which is invertible if $y \ne -1/2$. This value for y gives $x \ne -1/4$. Given any point $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ satisfying $y_0^2 + y_0 + 3x_0 + 1 = 0$ *except* the point $\begin{pmatrix} -1/4 \\ -1/2 \end{pmatrix}$, there is a neighborhood of $f\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ in which $f\begin{pmatrix} x \\ y \end{pmatrix} = 0$ expresses y implicitly as a function of x.

c. Since y is the pivotal unknown, we write

$$L = \begin{bmatrix} 2y+1 & 3 \\ 0 & 1 \end{bmatrix}, \quad L^{-1} = \frac{1}{2y+1}\begin{bmatrix} 1 & -3 \\ 0 & 2y+1 \end{bmatrix}.$$

So at the point with coordinates $x = -\frac{1}{2}, y = \frac{\sqrt{3}-1}{2}$, we have

$$L^{-1} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & -3 \\ 0 & \sqrt{3} \end{bmatrix}, \qquad |L^{-1}| = \frac{1}{\sqrt{3}}\sqrt{13}, \quad \text{giving } |L^{-1}|^2 = \frac{13}{3}.$$

The derivative of f is Lipschitz with Lipschitz ratio 2, so the largest radius R of a ball on which the implicit function g is guaranteed to exist by the implicit function theorem satisfies

$$2 = \frac{1}{2R|L^{-1}|^2}, \quad \text{so } R = \frac{3}{52}. \tag{1}$$

Equations 1: In the first equation we are using equation 2.10.26.

d. If we choose the point with coordinates $x = -\frac{13}{4}$, $y = \frac{5}{2}$, we do a bit better; in this case

$$L^{-1} = \frac{1}{6}\begin{bmatrix} 1 & -3 \\ 0 & 6 \end{bmatrix}, \text{ giving } |L^{-1}|^2 = \frac{23}{18} \text{ and } R = \frac{18}{92}.$$

2.10.7 Yes. Set $\mathbf{F}\begin{pmatrix} x \\ y \\ t \end{pmatrix} = \begin{pmatrix} x^2 \\ y - t \end{pmatrix}$. The function $\mathbf{g}(t) = \begin{pmatrix} 0 \\ t \end{pmatrix}$ is an appropriate implicit function, i.e., $\mathbf{F}\begin{pmatrix} \mathbf{g}(t) \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in a neighborhood of $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. But the implicit function theorem does not say that an implicit function exists: at $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ the derivative $\left[\mathbf{DF}\begin{pmatrix} x \\ y \\ t \end{pmatrix}\right] = \begin{bmatrix} 2x & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}$ is $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix}$; this matrix is not onto $\mathbb{R}^2$.

2.10.9 The function $\mathbf{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y + \sin(xy) \\ \sin(x^2 + y) \end{pmatrix}$ maps the origin to the origin, and $[\mathbf{DF}(\mathbf{0})] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Certainly F is differentiable with Lipschitz derivative. So it is locally invertible near the origin, and the point $\begin{pmatrix} a \\ 2a \end{pmatrix}$ will be in the domain of the inverse for $|a|$ sufficiently small.

2.10.11 a. If you set up the two multiplications $AH_{i,j}$ and $H_{i,j}A$, you will see that

$$T(H_{i,j}) = (a_i + a_j)H_{i,j}.$$

b. The $H_{i,j}$ form a basis for Mat (n, n). Their images are all multiples of themselves, so they are linearly independent unless any multiple is 0, i.e., unless $a_i = -a_j$ for some i and j.

c. False. If there were such a differentiable map, its derivative would be the inverse of $T : \text{Mat}\,(n, n) \to \text{Mat}\,(n, n)$, where $T(H) = BH + HB$, and

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

But T is not invertible, since $a_1 = -1$ and $a_2 = 1$, so $a_1 = -a_2$.

It follows from the chain rule that the derivative of the inverse is the inverse of the derivative.

2.10.13 False. For example, the function $f\begin{pmatrix} x \\ y \\ x \end{pmatrix} = y + z$ satisfies the hypotheses but does not express x in any way. For the implicit function theorem to apply, we would need $D_1 f \neq 0$.

2.10.15 a. The derivative of $\mathbf{F}$ is given by $\left[\mathbf{DF}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{bmatrix} e^x & e^y \\ e^x & -e^{-y} \end{bmatrix}$.

The columns of this matrix are linearly independent, so the inverse function theorem guarantees that F is locally invertible.

b. Differentiate $\mathbf{F}^{-1} \circ \mathbf{F} = I$ using the chain rule:

$$[\mathbf{D}(\mathbf{F}^{-1} \circ \mathbf{F})(\mathbf{a})] = [\mathbf{DF}^{-1}(\mathbf{F}(\mathbf{a}))][\mathbf{DF}(\mathbf{a})] = [\mathbf{DF}^{-1}(\mathbf{b})][\mathbf{DF}(\mathbf{a})] = I,$$

$$\text{so} \qquad [\mathbf{DF}^{-1}(\mathbf{b})] = [\mathbf{DF}(\mathbf{a})]^{-1} = \frac{-1}{e^{a_1}(e^{a_2} + e^{-a_2})}\begin{bmatrix} -e^{-a_2} & -e^{a_2} \\ -e^{a_1} & e^{a_1} \end{bmatrix}.$$

Solution 2.10.15, part a: We can see that the columns of the derivative are linearly independent because the two entries of the first column have the same sign and the two entries of the second column have opposite signs, or because the determinant is $-e^x(e^y + e^{-y})$, which never vanishes.

2.10.17 We will assume that the function $f : [a, b] \to [c, d]$ is increasing. If it isn't, replace f by $-f$.

1. Pick a number $y \in [c,d]$, and consider the function $f_y : x \mapsto f(x) - y$, which is continuous, since f is. Since

$$f_y(a) = c - y \le 0 \quad \text{and} \quad f_y(b) = d - y \ge 0,$$

by the intermediate value theorem there must exist a number $x \in [a,b]$ such that $f_y(x) = 0$, i.e., $f(x) = y$.

Moreover, this number x is unique: if $x_1 < x_2$ and both x_1 and x_2 solve the equation, then $f(x_1) = f(x_2) = y$, but $f(x_1) < f(x_2)$ since f is increasing on $[a,b]$. So set $g(y) \in [a,b]$ to be the unique solution of $f_y(x) = 0$. Then by definition we have

$$0 = f_y\big(g(y)\big) = f\big(g(y)\big) - y, \quad \text{so} \quad f\big(g(y)\big) = y.$$

This says that

$$f\Big(g\big(f(x)\big)\Big) = f(x),$$

for any $x \in [a,b]$. But any $y \in [c,d]$ can be written $y = f(x)$, so $f(g(y)) = y$ for any $y \in [c,d]$.

We need to check that g is continuous. Choose $y \in (c,d)$ (the endpoints c and d are slightly different and a bit easier) and $\epsilon > 0$. Consider the sequences $n \mapsto x'_n$ and $n \mapsto x''_n$ defined by

$$x'_n = g(y - 1/n) \quad \text{and} \quad x''_n = g(y + 1/n).$$

These sequences are respectively increasing and decreasing, and both are bounded, so they have limits x' and x''. Applying f (which we know to be continuous) gives

$$\begin{aligned} f(x') &= \lim_{n\to\infty} f(x'_n) = \lim_{n\to\infty}\left(y - \frac{1}{n}\right) = y \\ f(x'') &= \lim_{n\to\infty} f(x''_n) = \lim_{n\to\infty}\left(y + \frac{1}{n}\right) = y. \end{aligned}$$

Thus $x' = x'' = g(y)$, and there exists N such that $x''_N - x'_N < \epsilon$. Now set $\delta = 1/N$; if $|z - y| < \delta$, then $g(z) \in [x'_N, x''_N]$, and in particular, $|g(z) - g(y)| < \epsilon$. This is the definition of continuity.

2. Choose $y \in [c,d]$, and define two sequences $n \mapsto x'_n$, $n \mapsto x''_n$ in $[a,b]$ by induction, setting $x'_0 = a$, $x''_0 = b$, and

$$\begin{cases} x'_{n+1} = \dfrac{x'_n + x''_n}{2}, \; x''_{n+1} = x''_n & \text{if } y \ge f\left(\dfrac{x'_n + x''_n}{2}\right); \\[2ex] x''_{n+1} = \dfrac{x'_n + x''_n}{2}, \; x'_{n+1} = x'_n & \text{if } y < f\left(\dfrac{x'_n + x''_n}{2}\right). \end{cases}$$

Then both sequences converge to $g(y)$. Indeed, the sequence $n \mapsto x'_n$ is nondecreasing and the sequence $n \mapsto x''_n$ is nonincreasing; and both are bounded, so they both converge to something. Moreover,

$$x''_n - x'_n = \frac{b-a}{2^n},$$

so they converge to the same limit. Finally,

$$f(x'_n) \leq y \quad \text{and} \quad f(x''_n) \geq y$$

for all n; passing to the limit, this means that

$$y \geq \lim_{n\to\infty} f(x'_n) = \lim_{n\to\infty} f(x''_n) \geq y,$$

so all the inequalities are equalities.

3. Suppose $y = f(x)$, and that f is differentiable at x, with $f'(x) \neq 0$. Define the function $k(h)$ by $g(y+h) = x + k(h)$. Since g is continuous, $\lim_{h\to 0} k(h) = 0$. Now

$$\frac{g(y+h) - g(y)}{h} = \frac{x + k(h) - x}{h} = \frac{k(h)}{f(x + k(h)) - f(x)}.$$

Take the limit as $h \to 0$. Since

$$\lim_{h\to 0} \frac{k(h)}{f(x+k(h)) - f(x)} = \frac{1}{f'(x)},$$

we see that

$$\lim_{h\to 0} \frac{g(y+h) - g(y)}{h} = g'(y)$$

also exists, and the limits are equal.

Solutions for review exercises, chapter 2

2.1 a. Row reduction gives

$$\begin{bmatrix} 1 & 1 & -1 & a \\ 1 & 0 & 2 & b \\ 1 & a & 1 & b \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & -1 & a \\ 0 & -1 & 3 & b-a \\ 0 & a-1 & 2 & b-a \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2 & b \\ 0 & 1 & -3 & a-b \\ 0 & 0 & 3a-1 & a(b-a) \end{bmatrix}.$$

We will consider the cases $a = 1/3$ and $a \neq 1/3$. If $a = 1/3$, we get

$$\begin{bmatrix} 1 & 0 & 2 & b \\ 0 & 1 & -3 & \frac{1}{3}-b \\ 0 & 0 & 0 & \frac{1}{3}(b-\frac{1}{3}) \end{bmatrix},$$

so if $a = 1/3$ and $b = 1/3$, there are infinitely many solutions. If $a = 1/3$ and $b \neq 1/3$, then there are no solutions.

If $a = 1/3$ and $b = 1/3$, there is no pivotal 1 in the 4th column or the 3rd column. If $a = 1/3$ and $b \neq 1/3$, then by further row reducing we can get a pivotal 1 in the 4th column.

If $a \neq 1/3$, we can further row reduce our original matrix to get a pivotal 1 in the third column. In that case, the system of equations has a unique solution.

b. We have already done all the work: the matrix of coefficients (i.e., the matrix consisting of the first three columns) is invertible if and only if $a \neq 1/3$.

2.3 a. The first statement is true; the others are false.

b. Statements 1, 2, and 4 are true if $n = m$.

2.5 a. Saying that a matrix is invertible is equivalent to saying that it can be row reduced to the identity. Thus $[A,\ C]$ can be reduced to $[I_{n-k}, \widetilde{C}]$ if and only if A is invertible; likewise, $[[0],\ B]$ can be reduced to $[[0],\ I_k]$ if and only if B is invertible. So

$$M = \begin{bmatrix} A & C \\ [0] & B \end{bmatrix} \text{can be transformed into} \begin{bmatrix} I_{n-k} & \widetilde{C} \\ [0] & I_k \end{bmatrix}$$

by row operations. We can now add multiples of the last k rows to the first $n-k$ to cancel the entries of $\widetilde{C}$, which row reduces the whole matrix to the identity.

b. If you guess that there is an inverse of the form $\begin{bmatrix} A^{-1} & Y \\ [0] & B^{-1} \end{bmatrix}$ and multiply out, you find

$$\begin{array}{cc} & \begin{bmatrix} A & C \\ [0] & B \end{bmatrix} \\ \begin{bmatrix} A^{-1} & Y \\ [0] & B^{-1} \end{bmatrix} & \begin{bmatrix} I & A^{-1}C + YB \\ [0] & I \end{bmatrix} \end{array}.$$

So if $A^{-1}C + YB = [0]$, you will have found an inverse; this occurs when $Y = -A^{-1}CB^{-1}$.

Solution 2.5, part b: If you did not assume that the inverse is upper triangular, you could argue that it is of the form $\begin{bmatrix} X & Y \\ Z & W \end{bmatrix}$, where X, like A, is $(n-k)\times(n-k)$, and W, like B, is $k \times k$. Then the multiplication

$$\begin{bmatrix} A & C \\ [0] & B \end{bmatrix}\begin{bmatrix} X & Y \\ Z & W \end{bmatrix} = \begin{bmatrix} I & [0] \\ [0] & I \end{bmatrix}$$

says that $BZ = [0]$; since B is invertible, this says that $Z = [0]$. It also says that $BW = I$, i.e., that $W = B^{-1}$. So now we have

$$\begin{bmatrix} A & C \\ [0] & B \end{bmatrix}\begin{bmatrix} X & Y \\ [0] & B^{-1} \end{bmatrix} = \begin{bmatrix} AX & AY + CB^{-1} \\ [0] & I \end{bmatrix},$$

where

$$AX = I \quad \text{and} \quad AY + CB^{-1} = [0],$$

giving $Y = -A^{-1}CB^{-1}$.

2.7 a. A couple of row operations will bring

$$\begin{bmatrix} 1 & -1 & -1 \\ 0 & a & 1 \\ 2 & a+2 & a+2 \end{bmatrix} \quad \text{to} \quad \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & (a+3)/4 \\ 0 & 0 & (1-a)(a+4)/4 \end{bmatrix}.$$

Thus the matrix is invertible if and only if $(1-a)(a+4) \neq 0$, i.e., if $a \neq 1$ and $a \neq -4$.

b. This is an unpleasant row reduction of

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

The final result is

$$\frac{1}{(a-1)(a+4)} \begin{bmatrix} (a+2)(a-1) & 0 & a-1 \\ 2 & a+4 & -1 \\ -2a & -(a+4) & a \end{bmatrix}.$$

2.9 a. The vectors $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ and $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ are orthogonal since

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \cdot \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} = 0$$

(Corollary 1.4.8). By Proposition 2.4.16, they are linearly independent, and two linearly independent vectors in $\mathbb{R}^2$ form a basis (Corollary 2.4.20). The length of each basis vector is 1:

$$\sqrt{\cos^2\theta + \sin^2\theta} = 1 \quad \text{and} \quad \sqrt{\cos^2\theta + (-\sin\theta)^2} = 1,$$

so the basis is orthonormal, not just orthogonal.

b. The argument in part a applies.

Recall that an $n \times n$ matrix whose columns form an orthonormal basis of $\mathbb{R}^n$ is called an orthogonal matrix.

c. Let $A = [\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2]$ be any orthogonal matrix. Then $\vec{\mathbf{v}}_1$ is a unit vector, so it can be written $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ for some angle θ. The vector $\vec{\mathbf{v}}_2$ must be orthogonal to $\vec{\mathbf{v}}_1$ and of length 1. There are only two choices: $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$ and $\begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$. The first gives positive determinant and corresponds to rotation. Note that Proposition 1.4.14 says that to have positive determinant, $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ must be clockwise from $\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$, which it always will be.

We saw in Example 1.3.9 that

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

corresponds to rotation by θ counterclockwise around the origin.

The second choice gives a negative determinant and corresponds to reflection with respect to a line through the origin that forms angle $\theta/2$ with the x-axis (see Example 1.3.5).

2.11 a. Write $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ as $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$; the question is whether the columns of $\begin{bmatrix} 1 & 1 & 5 & 13 \\ 0 & 2 & 4 & 14 \\ 0 & 2 & 4 & 14 \\ 1 & 1 & 5 & 13 \end{bmatrix}$ are linearly independent. Since this matrix row reduces to $\begin{bmatrix} 1 & 0 & 3 & 6 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, it follows that I and A are linearly independent, and

$$A^2 = 2A + 3I, \quad A^3 = 7A + 6I.$$

The subspace spanned by I, A, A^2, A^3 has dimension 2.

Here we write

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ as } \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \text{ as } \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix},$$

$$A^2 = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} \text{ as } \begin{bmatrix} 5 \\ 4 \\ 4 \\ 5 \end{bmatrix},$$

and so on.

b. Suppose B_1 and B_2 satisfy $AB_1 = B_1A$ and $AB_2 = B_2A$. Then

$$(xB_1 + yB_2)A = xB_1A + yB_2A = xAB_1 + yAB_2 = A(xB_1 + yB_2).$$

So W is a subspace of $\operatorname{Mat}(2,2)$. If you set $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and write out $AB = BA$ as equations in a, b, c, d, you find

$$a+2c = a+2b \quad , \quad b+2d = 2a+b \quad , \quad 2a+c = c+2d \quad , \quad 2b+d = 2c+d.$$

These equations pretty obviously describe a 2-dimensional subspace of $\operatorname{Mat}(2,2)$, but to make it clear, write the equations

$$b - c = 0 \quad , \quad a - d = 0 \quad , \quad a - d = 0 \quad , \quad b - c = 0.$$

In other words, W is the kernel of the matrix

$$\begin{bmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}. \quad \text{This matrix row reduces to} \quad \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

so W has dimension 2, the number of nonpivotal columns.

c. If $B = aI + bA$, then

$$BA = (aI + bA)A = aA + bA^2 = A(aI + bA) = AB,$$

so clearly $V \subset W$. But they have the same dimension, so they are equal.

Solution 2.13: The pivotal columns of T form a basis for $\operatorname{Im} T$ (see Theorem 2.5.4). By definition 2.2.5, a column of T is pivotal if the corresponding column of the row-reduced matrix $\widetilde{T}$ contains a pivotal 1.

2.13 a. The matrix

$$\begin{bmatrix} 1 & 1 & 3 & 6 & 2 \\ 2 & -1 & 0 & 4 & 1 \\ 4 & 1 & 6 & 16 & 5 \end{bmatrix} \quad \text{row reduces to} \quad \begin{bmatrix} 1 & 0 & 1 & 10/3 & 1 \\ 0 & 1 & 2 & 8/3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The first two columns are a basis for the image, and the vectors

$$\begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -10/3 \\ -8/3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for the kernel (see Theorem 2.5.6). Indeed we find

$$\dim(\ker\ ) + \dim(\mathrm{img}) = 3 + 2 = 5.$$

b. The matrix

$$\begin{bmatrix} 2 & 1 & 3 & 6 & 2 \\ 2 & -1 & 0 & 4 & 1 \end{bmatrix} \quad \text{row reduces to} \quad \begin{bmatrix} 1 & 0 & 3/4 & 5/2 & 3/4 \\ 0 & 1 & 3/2 & 1 & 1/2 \end{bmatrix}.$$

If we had followed to the letter the procedure outlined in Example 2.5.7, and had not "cleared denominators", we would have 1 instead of -4 in the first and third vector, and 1 instead of -2 in the second vector. The other nonzero entries would then be fractions.

So the image is all of $\mathbb{R}^2$, and a basis for the kernel (we cleared denominators) is

$$\begin{bmatrix} 3 \\ 6 \\ -4 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \\ 0 \\ -4 \end{bmatrix}, \quad \text{which has dim} = 3.$$

Again, we find $\dim(\ker\ ) + \dim(\mathrm{img}) = 3 + 2 = 5$.

2.15 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. The equation $A^\top A = I$ gives

$$a^2 + c^2 = 1, \quad b^2 + d^2 = 1, \quad ab + cd = 0.$$

So $\begin{bmatrix} a \\ c \end{bmatrix}$ is a unit vector and can be written $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ for some θ. The vector $\begin{bmatrix} b \\ d \end{bmatrix}$ is a unit vector orthogonal to $\begin{bmatrix} a \\ c \end{bmatrix}$, so it is either

$$\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}.$$

In the first case, A is rotation by angle θ (Example 1.3.9); in the second it is reflection in the line with polar angle $\theta/2$ (see equation 1.3.11).

2.17 Let P_{2k-1} be the space of polynomials of degree at most $2k-1$; it can be identified with $\mathbb{R}^{2k}$, by identifying $c_0 + c_1 x + \cdots + c_{2k-1}x^{2k-1}$ with $\begin{bmatrix} c_0 \\ \vdots \\ c_{2k-1} \end{bmatrix}$. We will apply the dimension formula to the linear mapping $T : P_{2k-1} \to \mathbb{R}^{2k}$ defined in the margin. The kernel of T is $\{0\}$: indeed, a polynomial in the kernel can be factored as

$$T(p) = \begin{bmatrix} p(1) \\ \vdots \\ p(k) \\ p'(1) \\ \vdots \\ p'(k) \end{bmatrix}$$

Map T for Solution 2.17

$$p(x) = q(x)(x-1)^2 \dots (x-k)^2,$$

and such a polynomial has degree at least $2k$ unless $q = 0$. Since p has degree at most $2k-1$, it follows that q must be 0, hence p must be 0, so the kernel of T is $\{0\}$. So from the dimension formula, the image of T has dimension $2k$, hence is all of $\mathbb{R}^{2k}$. So there is a unique solution to the equations $p(1) = a_1, \dots, p(k) = a_k$ and $p'(1) = b_1, \dots, p'(k) = b_k$.

2.19 Just compute away:

$$\begin{aligned}\big(T(af+bg)\big)(x) &= (x^2+1)(af+bg)''(x) - x(af+bg)'(x) + 2(af+bg)(x)\\ &= (x^2+1)af''(x) + (x^2+1)bg''(x) - xaf'(x) - xbg'(x) + 2af(x) + 2bg(x)\\ &= (x^2+1)af''(x) - xaf'(x) + 2af(x) + (x^2+1)bg''(x) - xbg'(x) + 2bg(x)\\ &= \big(aT(f) + bT(g)\big)(x).\end{aligned}$$

2.21 Suppose that the vectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis of V, and that $\mathbf{w}_1, \dots, \mathbf{w}_m$ are linearly independent. Then (by Proposition 2.6.15), $\Phi_{\{\mathbf{v}\}}$ is one to one and onto, hence invertible, and $\Phi_{\{\mathbf{w}\}}$ is one to one, so $\Phi_{\{\mathbf{v}\}}^{-1} \circ \Phi_{\{\mathbf{w}\}} : \mathbb{R}^m \to \mathbb{R}^n$ is a one-to-one linear transformation, so the columns of its matrix are linearly independent. But these are m vectors in $\mathbb{R}^n$, so $m \le n$.

Similarly, if $\mathbf{v}_1, \dots, \mathbf{v}_n$ form a basis of V, and $\mathbf{w}_1, \dots, \mathbf{w}_m$ span V, then $\Phi_{\{\mathbf{v}\}}^{-1} \circ \Phi_{\{\mathbf{w}\}} : \mathbb{R}^m \to \mathbb{R}^n$ is onto, so the columns of its matrix are m vectors that span $\mathbb{R}^n$. Then $m \ge n$.

Solution 2.23: If your copy of the textbook is the first printing, note that the statement of the exercise has been changed to the following:

Let A be an $n \times m$ matrix. Show that if you row reduce the augmented matrix $[A^\top | I_m]$ to get $[\widetilde{A}^\top \widetilde{B}^\top]$, the nonzero columns of $\widetilde{A}$ form a basis for the image of A, and the columns of $\widetilde{B}$ corresponding to the zero columns of $\widetilde{A}$ form a basis for the kernel of A.

2.23 Let A be an $n \times m$ matrix. Column reducing a matrix is the same as row reducing the transpose, then transposing back. Moreover, row reducing A can be done by multiplying A on the left by a product of elementary matrices, which is invertible. Further, the matrix $[A^\top \mid I]$ is onto, so after row reduction there will be a leading 1 in every row.

So for any matrix A we can find an invertible $m \times m$ matrix which we will call $H^\top$ such that

$$H^\top [A^\top \mid I] = [\widetilde{A}^\top \mid \widetilde{B}^\top].$$

Transposing gives the $(n+m) \times m$ matrix

$$\begin{bmatrix} A \\ I \end{bmatrix} H = \begin{bmatrix} AH \\ H \end{bmatrix} = \begin{bmatrix} \widetilde{A} \\ \widetilde{B} \end{bmatrix} = \begin{bmatrix} \vec{\mathbf{a}}_1 \cdots \vec{\mathbf{a}}_k & \vec{\mathbf{0}} \\ * & \vec{\mathbf{b}}_1 \cdots \vec{\mathbf{b}}_l \end{bmatrix},$$

where $k + l = m$ and by definition k is the index of the last column whose pivotal 1 is in some row whose index is at most n.

The vectors $\vec{\mathbf{a}}_1, \dots, \vec{\mathbf{a}}_k$ are in the image of A, since they are columns of AH, hence linear combinations of the columns of A. Moreover they are linearly independent since each contains a pivotal 1 in a different row, with 0's above, and they span the image since H is invertible. So they form a basis for the image of A.

The vectors $\vec{\mathbf{b}}_1, \dots, \vec{\mathbf{b}}_l$ (the lower right "corner") are the last l columns of H. The upper right "corner" says that each of the last l columns of AH is $\vec{\mathbf{0}}$. So $\vec{\mathbf{b}}_1, \dots, \vec{\mathbf{b}}_l$ are in $\ker A$. These vectors are linearly independent (again because they contain pivotal 1's in different rows, with 0's above). They span the kernel by the dimension formula:

$$m = k + l \quad \text{and} \quad m = \dim \operatorname{img} A + \dim \ker A = k + \dim \ker A,$$

so $\dim \ker A = l$. Thus $\vec{\mathbf{b}}_1, \dots, \vec{\mathbf{b}}_l$ form a basis of $\ker A$.

2.25 It takes a bit of patience entering this problem into MATLAB, much helped by using the computer to compute A^3 symbolically. The answer ends up being

$$\begin{bmatrix} 2.08008382305190 & -0.00698499417571 & 0.08008382305190 \\ 0 & 1.91293118277239 & 0 \\ 0 & 0.17413763445522 & 2 \end{bmatrix}.$$

Note that one step of Newton's method gives

$$\begin{bmatrix} 2+1/12 & 0 & 1/12 \\ 0 & 2-1/12 & 0 \\ 0 & 1/6 & 2 \end{bmatrix} = \begin{bmatrix} 2.08\overline{3} & 0 & 0.08\overline{3} \\ 0 & 1.91\overline{6} & 0 \\ 0 & 0.1\overline{6} & 2 \end{bmatrix}$$

which is very close.

2.27 a. Since $\left| \begin{bmatrix} 2x_1 & -1 \\ -1 & 2y_1 \end{bmatrix} - \begin{bmatrix} 2x_2 & -1 \\ -1 & 2y_2 \end{bmatrix} \right| = 2|\mathbf{x}-\mathbf{y}|$, we see that 2 is a Lipschitz constant for the derivative of $\mathbf{f}$.

b. The inverse of $[\mathbf{Df}(\mathbf{x})]$ is

$$[\mathbf{Df}(\mathbf{x})]^{-1} = \frac{1}{4xy-1}\begin{bmatrix} 2y & 1 \\ 1 & 2x \end{bmatrix} \quad \text{which at } \mathbf{x}_0 = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \text{ is } \frac{1}{23}\begin{bmatrix} 6 & 1 \\ 1 & 4 \end{bmatrix},$$

so we have

$$\vec{\mathbf{h}}_0 = -\frac{1}{23}\begin{bmatrix} 6 & 1 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} +5/23 \\ -3/23 \end{bmatrix}, \quad \text{so} \quad \mathbf{x}_1 = \begin{bmatrix} 2+5/23 \\ 3-3/23 \end{bmatrix}.$$

The length of $[\mathbf{Df}(\mathbf{x}_0)]^{-1}$ is $\sqrt{54}/23$, so

$$|\mathbf{f}(\mathbf{x}_0)| \ \left|[\mathbf{Df}(\mathbf{x}_0)]^{-1}\right|^2 \ M = \sqrt{2} \cdot \frac{54}{23^2} \cdot 2 \approx 0.289 < \frac{1}{2},$$

so Newton's method converges, to a point of the disc of radius

$$|\vec{\mathbf{h}}_0| = \frac{\sqrt{34}}{23} \approx .25 \quad \text{around } \mathbf{x}_1.$$

c. The disc centered at $\begin{pmatrix} 51/23 \\ 66/23 \end{pmatrix}$ and with radius $\dfrac{\sqrt{34}}{23}$ is guaranteed to contain a unique root; see the figure in the margin.

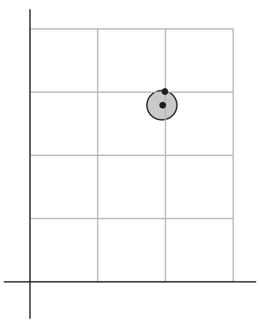


FIGURE FOR SOLUTION 2.27.

The shaded region, described in part c, is guaranteed to contain a unique root.

2.29 First, note that

$$\|A\| = \sup_{|\vec{\mathbf{v}}|=1} |A\vec{\mathbf{v}}| = \sup_{|\vec{\mathbf{w}}|=1, |\vec{\mathbf{v}}|=1} |\vec{\mathbf{w}}^\top||A\vec{\mathbf{v}}| \geq \sup_{|\vec{\mathbf{w}}|=1, |\vec{\mathbf{v}}|=1} |\vec{\mathbf{w}}^\top A\vec{\mathbf{v}}|; \tag{1}$$

similarly,

$$\|A^\top\| = \sup_{|\vec{\mathbf{w}}|=1} |A^\top\vec{\mathbf{w}}| = \sup_{|\vec{\mathbf{v}}|=1, |\vec{\mathbf{w}}|=1} |\vec{\mathbf{v}}^\top||A^\top\vec{\mathbf{w}}| \geq \sup_{|\vec{\mathbf{w}}|=1, |\vec{\mathbf{v}}|=1} |\vec{\mathbf{v}}^\top A^\top\vec{\mathbf{w}}|. \tag{2}$$

Equation 3: Since $(A\vec{\mathbf{v}})^\top\vec{\mathbf{w}}$ is a number, it equals its transpose.

Note also that

$$|\vec{\mathbf{w}}^\top A\vec{\mathbf{v}}| = \left|\left((A\vec{\mathbf{v}})^\top\vec{\mathbf{w}}\right)^\top\right| = |(A\vec{\mathbf{v}})^\top\vec{\mathbf{w}}| = |\vec{\mathbf{v}}^\top A^\top\vec{\mathbf{w}}|. \tag{3}$$

Therefore, if we can show that the inequalities (1) and (2) are equalities for appropriate $\vec{\mathbf{v}}, \vec{\mathbf{w}}$, we will have shown that $\|A\| = \|A^\top\|$.

By definition, $\|A\| = \sup_{|\vec{\mathbf{v}}|=1} |A\vec{\mathbf{v}}|$, and since the set $\{\, \vec{\mathbf{v}} \in \mathbb{R}^n \mid |\vec{\mathbf{v}}| = 1 \,\}$ is compact, the maximum is realized: there exists a unit vector $\vec{\mathbf{v}}_0$ such that $|A\vec{\mathbf{v}}_0| = \|A\|$. Let $\vec{\mathbf{w}}_0 = \dfrac{A\vec{\mathbf{v}}_0}{|A\vec{\mathbf{v}}_0|}$. Then

$$\|A\| = |A\vec{\mathbf{v}}_0| = \underbrace{\vec{\mathbf{w}}_0^\top \vec{\mathbf{w}}_0}_{1} |A\vec{\mathbf{v}}_0| = \vec{\mathbf{w}}_0^\top \frac{A\vec{\mathbf{v}}_0}{|A\vec{\mathbf{v}}_0|} |A\vec{\mathbf{v}}_0| = \vec{\mathbf{w}}_0^\top A\vec{\mathbf{v}}_0 = |\vec{\mathbf{w}}_0^\top A\vec{\mathbf{v}}_0|.$$

Therefore,

$$\|A\| = \sup_{|\vec{\mathbf{w}}|=1, |\vec{\mathbf{v}}|=1} |\vec{\mathbf{w}}^\top A\vec{\mathbf{v}}|$$

and by a similar argument,

$$\|A^\top\| = \sup_{|\vec{\mathbf{w}}|=1, |\vec{\mathbf{v}}|=1} |\vec{\mathbf{v}}^\top A^\top \vec{\mathbf{w}}|.$$

It follows that $\|A\| = \|A^\top\|$.

Solution 2.31: We cannot simply invoke the implicit function theorem in part a. Since

$$D_1F(\mathbf{a}) = 0,$$

the theorem does not guarantee existence of a differentiable implicit function expressing x and in terms of y and z, but it does not guarantee nonexistence either.

Since $F \circ G$ goes from $\mathbb{R}^3$ to $\mathbb{R}$, its derivative $[\mathbf{D}F \circ G(\mathbf{a})]$ is a 1×3 (row) matrix.

2.31 a. False. Suppose $\sin(xyz) - z = 0$ expresses x implicitly as $g\begin{pmatrix} y \\ z \end{pmatrix}$ near $\mathbf{a} = \begin{pmatrix} \pi/2 \\ 1 \\ 1 \end{pmatrix}$. Set

$$F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sin(xyz) - z \quad \text{and} \quad G\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} g\begin{pmatrix} y \\ z \end{pmatrix} \\ y \\ z \end{pmatrix}.$$

Then $F \circ G = 0$, so, by the chain rule, $[\mathbf{D}F(G(\mathbf{a}))][\mathbf{D}G(\mathbf{a})] = [0 \quad 0 \quad 0]$. Writing this out in terms of matrices gives

$$\Big[\ \underbrace{1 \cdot 1 \cdot \cos\frac{\pi}{2}}_{D_1F = yz\cos xyz}\ \ \underbrace{\frac{\pi}{2} \cdot 1 \cdot \cos\frac{\pi}{2}}_{D_2F = xz\cos xyz}\ \ \underbrace{\frac{\pi}{2}\cos\frac{\pi}{2} - 1}_{D_3F = xy\cos xyz - 1}\ \Big] \begin{bmatrix} 0 & D_1g\begin{pmatrix} 1 \\ 1 \end{pmatrix} & D_2g\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [0 \quad 0 \quad 0].$$

This is a contradiction, since the third entry of the product is -1.

b. True. Since the partial derivative with respect to z at $\mathbf{a}$ is $D_3F(\mathbf{a}) = -1$, which is invertible, the implicit function theorem guarantees that $F(\mathbf{x}) = 0$ expresses z implicitly as a function g of x and y near $\mathbf{a}$, and that it is differentiable, with derivative given by equation 2.10.29.

2.33 False. This is a problem most easily solved using the chain rule, not the inverse function theorem. Call S the map $S(A) = A^2$. If such a mapping g exists, then g is an inverse function of S, so $S \circ g = I$, and since

g is differentiable, we can apply the chain rule:

$$\overbrace{\left[\mathbf{D}S\left(g\left(\begin{bmatrix}-3 & 0\\ 0 & -3\end{bmatrix}\right)\right)\right]}^{\left[\mathbf{D}S\left(\begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix}\right)\right]}\left[\mathbf{D}g\left(\begin{bmatrix}-3 & 0\\ 0 & -3\end{bmatrix}\right)\right] = \left[\mathbf{D}(S\circ g)\left(\begin{bmatrix}-3 & 0\\ 0 & -3\end{bmatrix}\right)\right]$$
$$= \left[\mathbf{D}I\left(\begin{bmatrix}-3 & 0\\ 0 & -3\end{bmatrix}\right)\right] = I.$$

The derivative of the identity is the identity everywhere; this uses Theorem 1.8.1, part 2.

Thus if the desired function g exists, then $\left[\mathbf{D}S\left(\begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix}\right)\right]$ is invertible; if the derivative is not invertible, either g does not exist, or it exists but is not differentiable. From Example 1.7.17, we know that $[\mathbf{D}S(A)]H = AH + HA$, so

$$\left[\mathbf{D}S\left(\begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix}\right)\right]B = \begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix}B + B\begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix}.$$

Written out in coordinates, setting $B = \begin{bmatrix}a & b\\ c & d\end{bmatrix}$, this gives

$$\begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} + \begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & 2\\ -2 & -1\end{bmatrix} = 2\begin{bmatrix}a-b+c & a+d\\ -a-d & -b+c-d\end{bmatrix}.$$

The transformation

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix} \mapsto \begin{bmatrix}a-b+c & a+d\\ -a-d & -b+c-d\end{bmatrix} \tag{1}$$

is not invertible, since the entries are not linearly independent.[2] Therefore g does not exist.

[2]Note that the matrix of the linear transformation (1) is *not* the matrix $\begin{bmatrix}a-b+c & a+d\\ -a-d & -b+c-d\end{bmatrix}$; it is $\begin{bmatrix}1 & -1 & 1 & 0\\ 1 & 0 & 0 & 1\\ -1 & 0 & 0 & -1\\ 0 & -1 & 1 & -1\end{bmatrix}$, since

$$\begin{bmatrix}1 & -1 & 1 & 0\\ 1 & 0 & 0 & 1\\ -1 & 0 & 0 & -1\\ 0 & -1 & 1 & -1\end{bmatrix}\begin{bmatrix}a\\ b\\ c\\ d\end{bmatrix} = \begin{bmatrix}a-b+c\\ a+d\\ -a-d\\ -b+c-d\end{bmatrix}.$$

This matrix is clearly noninvertible, since the third column is a multiple of the second column.

Alternatively, you could show that the derivative $\left[\mathbf{D}S\left(\begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}\right)\right]$ is not invertible by noticing that

$$\begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

so that

$$\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \in \ker \left[\mathbf{D}f\left(\begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}\right)\right].$$

If you want to find out whether a map $\mathbf{f}$ is locally invertible, the first thing to do is to see whether the derivative $[\mathbf{Df}]$ (here, $[\mathbf{D}S]$) is invertible. If it is not, then the chain rule tells you that there are two possibilities. Either there is no implicit function (the usual case), or there is one and it is not differentiable.

If $[\mathbf{Df}]$ is invertible, the next step is to show that $\mathbf{f}$ is continuously differentiable, so that the inverse function theorem applies. In this case, it is an easy consequence of Theorem 1.8.1 (rules for computing derivatives) that the squaring function S is continuously differentiable.

2.35 If the matrix $[\mathbf{D}f(x_n)]$ is not invertible, one can pick any invertible matrix A, and use it instead of $[\mathbf{D}f(x_n)]^{-1}$ at the troublesome iterate x_n:

$$x_{n+1} = x_n - Af(x_n).$$

This procedure cannot produce a false root x^*, since then we would have

$$x^* = x^* - Af(x^*),$$

which means that $Af(x^*) = 0$. The requirement that A be invertible then implies that $f(x^*) = 0$, so x^* is indeed a true root. Of course, we can not longer hope to have superconvergence, but the fact that our algorithm doesn't break down is probably worth the sacrifice.

2.37 a. Suppose first that p_1 and p_2 have no common factors, and that

$$p_1(x)q_1(x) + p_2(x)q_2(x) = 0;$$

we need to show that then q_1 and q_2 are the zero polynomial.

Write

$$p_1(x) = (x-a_1)\dots(x-a_{k_1}) \quad \text{and} \quad p_2(x) = (x-b_1)\dots(x-b_{k_2}).$$

By the fundamental theorem of algebra (Theorem 1.6.14), this factoring is possible, but the roots may be complex.

Since

$$p_1(x)q_1(x) + p_2(x)q_2(x) = 0 \text{ and } p_1(b_i) \neq 0 \quad \text{for all } i = 1, \dots, k_2,$$

we have $q_1(b_1) = \dots = q_1(b_{k_2}) = 0$. Similarly, $q_2(a_1) = \dots = q_2(a_{k_1}) = 0$.

If all the a_i are distinct, this shows that q_2 is a polynomial of degree $< k_1$ that vanishes at k_1 points; so it must be the zero polynomial. The same argument applies to q_1 if the b_i are distinct.

Actually, this still holds if the roots have multiplicities. Suppose that q_2 is not the zero polynomial. In the equation

$$(x-a_1)\dots(x-a_{k_1})q_1(x) + (x-b_1)\dots(x-b_{k_2})q_2(x) = 0,$$

cancel all the powers of $(x-a_i)$ that appear in both $p_1(x)$ and $q_2(x)$. There will be some such term $(x-a_j)$ left from $p_1(x)$ after all cancellations, since p_1 has higher degree than q_2, so if after cancellation we evaluate what is left at a_j, the first summand will give 0 and the second will not. This is a contradiction and shows that $q_2 = 0$. The same argument shows that $q_1 = 0$.

This proves the "if" part of (a). For the "only if" part, suppose that

$$p_1(x) = (x-c)\tilde{p}_1(x) \quad \text{and} \quad p_2(x) = (x-c)\tilde{p}_2(x),$$

i.e., that p_1 and p_2 have the common factor $(x-c)$. Then

$$p_1(x)\tilde{p}_2(x) - p_2(x)\tilde{p}_1(x) = (x-c)(\tilde{p}_1(x)\tilde{p}_2(x) - \tilde{p}_2(x)\tilde{p}_1(x)) = 0,$$

so $q_1 = \tilde{p}_2$ and $q_2 = -\tilde{p}_1$ provide a nonzero element of the kernel of T.

b. If p_1 and p_2 are relatively prime, then T is an injective (one to one) linear transformation between spaces of the same dimension, hence by the dimension formula (theorem 2.5.8) it is onto. In particular, the polynomial 1 is in the image. But if p_1 and p_2 have the common factor $(x-c)$, i.e.,

$$p_1(x) = (x-c)\tilde{p}_1(x) \quad \text{and} \quad p_2(x) = (x-c)\tilde{p}_2(x),$$

then any polynomial of the form $p_1(x)q_1(x) + p_2(x)q_2(x)$ will vanish at c, and hence will not be a nonzero constant.

Nathaniel Schenker suggested a different solution to Exercise 2.39: Row reduce A by multiplying by elementary matrices on the left. In the matrix BA, where B is a product of elementary matrices, there is a pivotal 1 in every row that is not all zeros, so by column operations (multiplication by elementary matrices D on the right; see Exercise 2.3.11), we can eliminate all the nonzero entries of nonpivotal columns. Now we have a matrix whose only entries are 0's and pivotal 1's. By further multiplication on left and right by elementary matrices we can move all the pivotal 1's to make the matrix of the form J_k. Thus $J_k = \widetilde{B}A\widetilde{D}$, where $\widetilde{B}$ and $\widetilde{D}$ are products of elementary matrices. This gives $A = \widetilde{B}^{-1}J_k\widetilde{D}^{-1}$, or $A = QJ_kP^{-1}$, with $Q = \widetilde{B}^{-1}$ and $P = \widetilde{D}$.

2.39 In the domain of A, choose vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-k}$ that span the kernel of A (and hence are linearly independent, since the dimension of the kernel is $n-k$, by the dimension formula (Theorem 2.5.8). Add further linearly independent vectors to make a basis $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-k}, \vec{\mathbf{v}}_{n-k+1}, \dots, \vec{\mathbf{v}}_n$. (The theorem of the incomplete basis justifies our adding vectors to make a basis; see Exercise 2.6.9.) In the codomain, define the vectors

$$\vec{\mathbf{w}}_{n-k+1} = A\vec{\mathbf{v}}_{n-k+1}, \dots, \vec{\mathbf{w}}_n = A\vec{\mathbf{v}}_n.$$

These vectors are linearly independent, because

$$\mathbf{0} = \sum_{i=n-k+1}^{n} a_i\vec{\mathbf{w}}_i = \sum_{i=n-k+1}^{n} a_iA\vec{\mathbf{v}}_i = A\left(\sum_{i=n-k+1}^{n} a_i\vec{\mathbf{v}}_i\right)$$

implies that $\sum_{i=n-k+1}^{n} a_i\vec{\mathbf{v}}_i$ is in ker A, hence it is in the span of the vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-k}$. So we can write

$$\sum_{i=n-k+1}^{n} a_i\vec{\mathbf{v}}_i = \sum_{i=1}^{n-k} b_i\vec{\mathbf{v}}_i,$$

which implies that all a_i (and all b_i) vanish.

Hence $\vec{\mathbf{w}}_{n-k+1}, \dots, \vec{\mathbf{w}}_n$ can also be completed to form a basis

$$\vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_{n-k}, \vec{\mathbf{w}}_{n-k+1}, \dots, \vec{\mathbf{w}}_n.$$

Set

$$P = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n] \quad \text{and} \quad Q = [\vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_n].$$

Both these matrices are invertible, and we have

The ith column of the matrix $Q = [\vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_n]$ is $Q\vec{\mathbf{e}}_i = \vec{\mathbf{w}}_i$, so

$$Q^{-1}(\vec{\mathbf{w}}_i) = Q^{-1}Q\vec{\mathbf{e}}_i = \vec{\mathbf{e}}_i.$$

$$Q^{-1}AP(\vec{\mathbf{e}}_i) = Q^{-1}A\vec{\mathbf{v}}_i = \begin{cases} 0 & \text{if } i \le n-k \\ Q^{-1}(\vec{\mathbf{w}}_i) = \vec{\mathbf{e}}_i & \text{if } i > n-k. \end{cases}$$

So $Q^{-1}AP = J_k$, or equivalently, $A = QJ_kP^{-1}$.

2.41 Start with $[A \,|\, I]$ and suppose that you have performed row operations so that there are pivotal 1's in the first m columns. The matrix then has the form

$$\begin{array}{cc} & \begin{array}{cccc} m & n-m & m & n-m \end{array} \\ \begin{array}{c} m \\ \\ n-m \end{array} & \left[\begin{array}{c|c|c|c} \begin{matrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{matrix} & * & * & 0 \\ \hline O & * & * & \begin{matrix}1 & & 0\\ & \ddots & \\ 0 & & 1\end{matrix} \end{array}\right] \end{array}.$$

All the entries marked * are likely to be nonzero; any zeros make the computation shorter. In the general case, to get from the mth step to the $(m+1)$th step, you need to perform

– n divisions
– $n(n-1)$ multiplications
– $(n-1)(n-1)$ additions

We have n divisions because there are $n+1$ possibly nonzero entries in the $(m+1)$th line; the pivotal 1 requires no division. We have $n(n-1)$ multiplications because each of the n rows except the $(m+1)$th requires n multiplications. You might think that by the same reasoning we have $n(n-1)$ additions, but the $m+1$st column of the future inverse does not require an addition because we are adding a number to 0, so we actually have $(n-1)(n-1)$ additions.

So to compute A^{-1} by bringing $[A|I]$ to row echelon form you need

$$n\Big(n + n(n-1) + (n-1)(n-1)\Big) = 2n^3 - 2n^2 + n \quad \text{operations.}$$

There are at least two ways to compute $A^{-1}\vec{\mathbf{b}}$. If you first compute A^{-1} and then $A^{-1}\vec{\mathbf{b}}$, you will use

$$2n^3 - 2n^2 + n + 2n^2 - n = 2n^3$$

operations. (The $2n^2 - n$ is for the matrix multiplication $A^{-1}\vec{\mathbf{b}}$.)

But if you row reduce $[A|\vec{\mathbf{b}}]$ and consider the last column, we have seen (Exercise 2.2.11) that we need $n^3 + \frac{1}{2}n^2 - \frac{1}{2}n$ operations, or

$$\frac{2}{3}n^3 + \frac{3}{2}n^2 - \frac{7}{6}n$$

operations if we use partial row reduction and substitution. Thus computing $A^{-1}\vec{\mathbf{b}}$ by first computing A^{-1} and then performing the matrix multiplication requires almost 3 times as many operations as partial row reduction and substitution.

Solutions for Chapter 3

3.1.1 The derivative of $F\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 + x_2^2 + x_3^2 - 1$ is $[2x_1 \quad 2x_2 \quad 2x_3]$, which is $[0 \quad 0 \quad 0]$ only at the origin, which does not satisfy $F(\mathbf{x}) = 0$. So the unit sphere given by $x_1^2 + x_2^2 + x_3^2 = 1$ is a smooth surface.

3.1.3 If the straight line is not vertical, it is the graph of the function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = a + bx$, i.e., $y = a + bx$. Since b is a constant, every value of x gives one and only one value of y. If the straight line is vertical, it is the graph of the function $g : \mathbb{R} \to \mathbb{R}$ given by $g(y) = a$, i.e., $x = a$; every value of y gives one and only one value of x.

3.1.5 a. The derivative of x^2+y^3 is $[2x, \; 3y^2]$. Since this is a transformation $\mathbb{R}^2 \to \mathbb{R}$, the only way it can fail to be onto is for it to vanish, which happens only if $x = y = 0$. This point is on X_0, so all X_c are smooth curves for $c \neq 0$. But X_0 is not a smooth curve near the origin. The equation $x^2+y^3 = 0$ can be "solved" for x: $x = \pm\sqrt{-y^3}$, and this is not a function near $y = 0$, for instance it isn't defined for $y > 0$. It can also be solved for y: $y = -x^{2/3}$. This is a function, but it is not differentiable at 0. Thus X_0 is not a smooth curve.

b. These curves are shown in the figure at left.

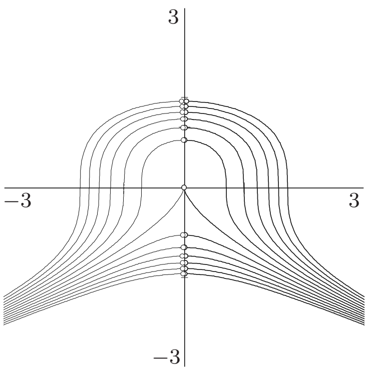


Figure for Solution 3.1.5. The curves X_c for

$$c = -3, -2.5, \dots, 3.$$

Note that X_0 has a cusp at the origin, which is why it is not a smooth curve.

3.1.7 a. Both X_a and Y_b are graphs of smooth functions representing z as functions of $\begin{pmatrix} x \\ y \end{pmatrix}$, so they are smooth surfaces.

b. The intersection $X_a \cap Y_b$ is a smooth curve if the derivative of

$$\mathbf{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2 + y^3 + z - a \\ x + y + z - b \end{pmatrix}$$

is onto (i.e., has rank 2) at every point of $X_a \cap Y_b$. The derivative is

$$[\mathbf{DF}] = \begin{bmatrix} 2x & 3y^2 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

and the only way it can have rank < 2 is if all three columns are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, i.e., if $2x = 3y^2 = 1$. So all the intersections $X_a \cap Y_b$ are smooth curves except the ones that go through the two vertical lines $x = 1/2, \; y = \pm 1/\sqrt{3}$. The values of a and b for the surfaces X_a and Y_b that contain such a point

$$\begin{pmatrix} 1/2 \\ \pm 1/\sqrt{3} \\ z \end{pmatrix} \qquad \text{are} \qquad \begin{aligned} a &= z + \frac{1}{4} \pm \left(\frac{1}{\sqrt{3}}\right)^3 \\ b &= z + \frac{1}{2} \pm \left(\frac{1}{\sqrt{3}}\right) \end{aligned},$$

which occurs precisely when $a - b = -1/4 \pm 2/(3\sqrt{3})$.

"Precisely when" means "if and only if".

At the points $\begin{pmatrix} 1/2 \\ +1/\sqrt{3} \\ z \end{pmatrix}$, $\begin{pmatrix} 1/2 \\ -1/\sqrt{3} \\ z \end{pmatrix}$, the plane Y_b is the tangent plane to the surface X_a.

3.1.9 a. We have

$$x^4 + x^2 + y^4 - y^2 = \left(x^2 + \frac{1}{2}\right)^2 + \left(y^2 - \frac{1}{2}\right)^2 - \frac{1}{2}.$$

So we can take

$$p(x) = x^2 + \frac{1}{2} \quad \text{and} \quad q(y) = y^2 - \frac{1}{2}.$$

b. The graphs of p, p^2, q, and q^2 are shown at top and middle of the figure in the margin. The graph of the function F is shown at the bottom. You can see that the slices at x constant look like the graph of q^2, raised or lowered, and similarly, slices at y constant look like the graph of p^2, raised or lowered, all with a single minimum. (Looking ahead to section 3.6, we see that the critical point $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is a saddle point, where p goes up and q goes down. The other critical points are $x = 0,\ y = \pm 1/\sqrt{2}$, which are minima.)

If you slice this graph parallel to the (x, y)-plane, you will get the curves of Figure 3.1.10 in the text.

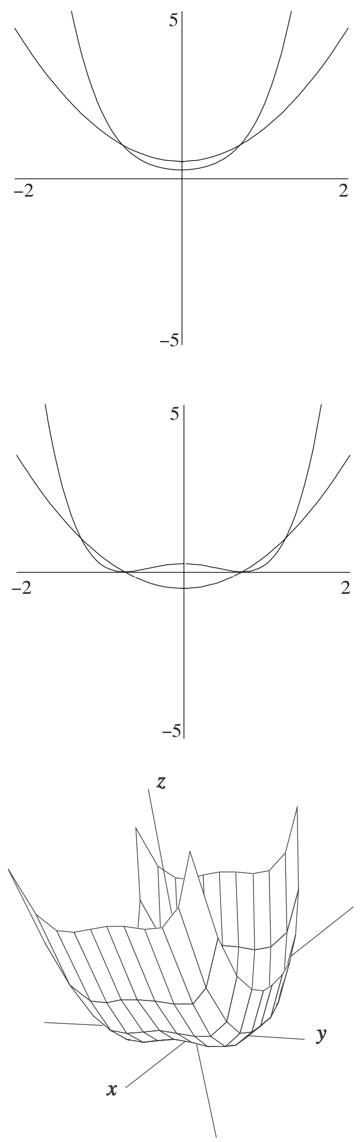


FIGURE FOR SOLUTION 3.1.9, part b. TOP: The graphs of p and p^2. MIDDLE: The graphs of q and q^2. BOTTOM: Graph of F.

3.1.11 a. The union is parametrized by $\begin{pmatrix} s \\ t \end{pmatrix} \mapsto \begin{pmatrix} st \\ st^2 \\ st^3 \end{pmatrix}$.

b. This means eliminating s and t from the equations $x = st$, $y = st^2$, and $z = st^3$.

In this case, it is easy to see that $t = y/x$, $s = x^2/y$, so that an equation for X is

$$x = \frac{y^2}{z} \quad \text{or} \quad xz = y^2.$$

The set defined by this equation contains two lines that are not in X: the x-axis and the z-axis.

c. The derivative of

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = xz - y^2 \quad \text{is} \quad \left[\mathbf{D}f\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = [z,\ -2y,\ x],$$

which evidently never vanishes except at the origin.

d. Just substitute into the equation $xz = y^2$:

$$r(1 + \sin\theta)r(1 - \sin\theta) = r^2(\cos\theta)^2.$$

The figure below shows the two parametrizations. The parametrization of part a, shown to the right, doesn't look like much. The left side shows the parametrization of part d; it certainly looks like a cone of revolution. Indeed, if you switch to the coordinates $x = u + v,\ z = u - v$, the equation becomes $u^2 - v^2 = y^2$. In the form $u^2 = y^2 + v^2$, you should recognize it as a cone of revolution around the u-axis.

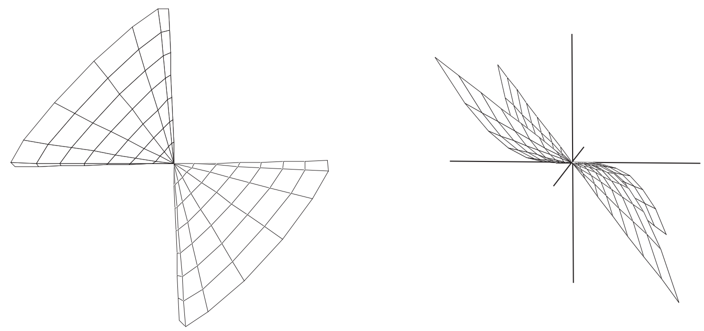


FIGURE FOR SOLUTION 3.1.11. LEFT: The parametrization for part d. RIGHT: The parametrization for part a.

e. The set of noninvertible symmetric matrices $\begin{bmatrix} a & b \\ b & d \end{bmatrix}$ has equation $ad - b^2 = 0$, so it is the same cone.

3.1.13 a. The equation is $(\mathbf{x} - \mathbf{a}) \cdot \vec{\mathbf{v}} = 0$.

b. The equation of P_t is

$$\left[\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}\right] \cdot \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix} = 0.$$

c. Writing the above equations in full, we find that the equations of $P_{t_1} \cap P_{t_2}$ are

$$\begin{aligned} x + 2t_1 y + 3t_1^2 z &= t_1 + 2t_1^3 + 3t_1^5 \\ x + 2t_2 y + 3t_2^2 z &= t_2 + 2t_2^3 + 3t_2^5. \end{aligned}$$

Two planes $x + a_1 y + b_1 z = c_1$ and $x + a_2 y + b_2 z = c_2$ fail to intersect if they coincide or are parallel. They coincide if the equations are the same; they are parallel if $x + a_1 y + b_1 z$ is a multiple of $x + a_2 y + b_2 z$. In either case, the matrix of coefficients

$$\begin{bmatrix} 1 & a_1 & b_1 \\ 1 & a_2 & b_2 \end{bmatrix}$$

has row rank 1, so by Proposition 2.5.11 it has rank 1.

To be sure that the planes intersect in a line, we must check that the matrix of coefficients

$$\begin{bmatrix} 1 & 2t_1 & 3t_1^2 \\ 1 & 2t_2 & 3t_2^2 \end{bmatrix}$$

has exactly two pivotal columns (i.e., rank 2). But if $t_1 \neq t_2$, the first two columns are linearly independent, so that is true.

d. The intersection $P_1 \cap P_{1+h}$ has equations

$$\begin{aligned} x + 2y + 3z &= 6 \\ x + 2(1+h)y + 3(1+h)^2 z &= (1+h) + 2(1+h)^3 + 3(1+h)^5 \\ &= 6 + 22h + 36h^2 + 32h^3 + 15h^4 + 3h^5. \end{aligned}$$

To solve this system, we row reduce the matrix

$$\begin{bmatrix} 1 & 2 & 3 & 6 \\ 1 & 2+2h & 3+6h+3h^2 & 6+22h+36h^2+32h^3+15h^4+3h^5 \end{bmatrix}.$$

Note that when row reducing this matrix, we can get rid of one power of h (subtract the second row from the first and divide by h). This leads to

$$\begin{bmatrix} 1 & 0 & -3-3h & -16-36h-32h^2-15h^3-3h^4 \\ 0 & 1 & 3+\frac{3}{2}h & 11+18h+16h^2+\frac{15}{2}h^3+\frac{3}{2}h^4 \end{bmatrix}.$$

The equations encoded by this matrix have a perfectly good limit when $h \to 0$:

$$x = 3z - 16 \quad \text{and} \quad y = -3z + 11,$$

which are a parametric representation of the required limiting line.

3.1.15 Clearly if $l_1 = l_2 + l_3 + l_4$, the only positions are those where the vertices are aligned. In that case, the position of $\mathbf{x}_1$ and the polar angle of the linkage determine the position in X_2; thus X_2 is a 3-dimensional manifold, which looks like $\mathbb{R}^2 \times S^1$. Similarly, X_3 is a 5-dimensional manifold, which looks like $\mathbb{R}^3 \times S^2$.

3.1.17 Two circles with radii r_1 and r_2, whose centers are a distance d apart, intersect at two points exactly if

$$|r_1 - r_2| < d, \quad |r_1 + r_2| > d.$$

If one of these inequalities is satisfied and the other is an equality, the circles are tangent and intersect in one point.

The point $\mathbf{x}_2$ is on the circle of radius l_1 around $\mathbf{x}_1$ and on the circle of radius l_2 around $\mathbf{x}_3$. These two circles therefore intersect at two points precisely if

$$|l_1 - l_2| < |\mathbf{x}_1 - \mathbf{x}_3|, \quad |l_1 + l_2| > |\mathbf{x}_1 - \mathbf{x}_3|.$$

Similarly, the point $\mathbf{x}_4$ is on the circle of radius l_3 around $\mathbf{x}_3$ and the circle of radius l_4 around $\mathbf{x}_1$. These two circles intersect in two points precisely if

$$|l_3 - l_4| < |\mathbf{x}_1 - \mathbf{x}_3|, \quad |l_3 + l_4| > |\mathbf{x}_1 - \mathbf{x}_3|.$$

Under these conditions, there are two choices for $\mathbf{x}_2$, and two choices for $\mathbf{x}_4$, leading to four positions in all.

3.1.19 a. The space of 2×2 matrices of rank 1 is precisely the set of matrices $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $\det A = ad - bc = 0$, but $A \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Compute the derivative:

$$\left[\mathbf{D} \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\right] = [\, d \quad -c \quad -b \quad a \,],$$

so that the derivative is nonzero on $M_1(2,2)$.

b. This is similar but longer. The space $M_2(3,3)$ is the space of 3×3 matrices whose determinant vanishes, but which have two linearly independent columns. If two vectors

$$\vec{\mathbf{a}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \quad \text{and} \quad \vec{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

are linearly independent, then $\vec{\mathbf{a}} \times \vec{\mathbf{b}} \neq \vec{\mathbf{0}}$ (see Exercise 1.4.13). Now we have

$$\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = a_1b_2c_3 - a_1c_2b_3 - b_1a_2c_3 + b_1c_2a_3 + c_1a_2b_3 - c_1b_2a_3,$$

so taking the variables in the order $a_1, a_2, \dots, c_3$, the derivative of the determinant is

$$[b_2c_3 - c_3b_2,\ -b_1c_3 + c_1b_3,\ b_1c_2 - c_1b_2,\ \dots,\ a_1b_2 - b_1a_2],$$

precisely the determinants of all the 2×2 submatrices you can make from the 3×3 matrix, or alternatively, precisely the coordinates of all the cross products of the columns. We just saw that at least one coordinate of such a cross product must be nonzero at a point in $M_2(3,3)$.

Solution 3.1.21, part b: Recall that we cannot use Theorem 3.1.10 to determine that a locus is *not* a smooth manifold; a locus defined by $\mathbf{F}(\mathbf{z}) = \mathbf{0}$ may be a smooth manifold even though $[\mathbf{DF}(\mathbf{z})]$ is not onto. But in this case, at these values of a and b the locus $X_{a,b}$ is a single point, which is not a smooth curve.

3.1.21 a. Yes, the theorem tells us that they are smooth surfaces. We define $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = e^x + 2e^y + 3e^z - a$, so that X_a is defined by $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$. This has solutions only for $a > 0$; for $a \le 0$, the set X_a is the empty set and hence is a smooth surface by our definition. If $a > 0$, we have to check that the derivative is onto at all points:

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = [e^x\ 2e^y\ 3e^z] \neq [0\ 0\ 0] \quad \text{for all} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

Thus X_a is a smooth surface.

b. No, Theorem 3.1.10 does not guarantee that they are smooth curves. We define $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^2$ as

$$\mathbf{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} e^x + 2e^y + 3e^z - a \\ x + y + z - b \end{pmatrix},$$

so that the subsets in question are given by $\mathbf{F}(\mathbf{x}) = \mathbf{0}$. Then

$$[\mathbf{DF}(\mathbf{x})] = \begin{bmatrix} e^x & 2e^y & 3e^z \\ 1 & 1 & 1 \end{bmatrix},$$

which has a single pivotal column if and only if $e^x = 2e^y = 3e^z$, i.e., $x = y + \ln(2) = z + \ln(3)$. A triple $\begin{pmatrix} x \\ x - \ln(2) \\ x - \ln(3) \end{pmatrix}$ satisfies both equations of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ when

$$e^x = a/3 \text{ and } 3x - \ln(2) - \ln(3) = b, \quad \text{i.e.,}$$

$$x = \ln(a/3) \quad \text{and} \quad x = \frac{b + \ln(6)}{3}.$$

Setting these two equations equal, we find that the subsets $X_{a,b}$ may not be smooth curves when $a = 3e^{(b+\ln(6))/3}$.

3.1.23 Manifolds are invariant under rotations (see Corollary 3.1.17), so we may assume that our linkage is on the x-axis, say at points

$$\mathbf{a}_1 \stackrel{\text{def}}{=} \begin{pmatrix} a_1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 \stackrel{\text{def}}{=} \begin{pmatrix} a_2 \\ 0 \end{pmatrix}, \quad \mathbf{a}_3 \stackrel{\text{def}}{=} \begin{pmatrix} a_3 \\ 0 \end{pmatrix}, \quad \mathbf{a}_4 \stackrel{\text{def}}{=} \begin{pmatrix} a_4 \\ 0 \end{pmatrix},$$

with $a_1 < a_2 < a_4 < a_3$ and

$$a_2 - a_1 = l_1, \; a_3 - a_2 = l_2, \; a_4 - a_3 = l_3, \; a_1 - a_4 = l_4.$$

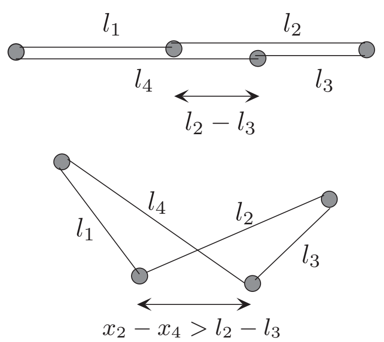


FIGURE FOR SOLUTION 3.1.23.

The case where the candidate coordinates consist of x_2 and x_4 and two y-variables.

Recall that $X_2 \subset (\mathbb{R}^2)^4$ is the set defined by the equations

$$|\mathbf{x}_1 - \mathbf{x}_2| = l_1, \quad |\mathbf{x}_2 - \mathbf{x}_3| = l_2, \quad |\mathbf{x}_3 - \mathbf{x}_4| = l_3, \quad |\mathbf{x}_4 - \mathbf{x}_1| = l_4.$$

We will denote by $A \in X_2$ the position where $\mathbf{x}_i = \mathbf{a}_i$, for $i = 1, 2, 3, 4$. Let $X_2^\epsilon \subset X_2$ be the subset where $|\mathbf{x}_i - \mathbf{a}_i| < \epsilon$, for $i = 1, 2, 3, 4$. We will call our variables $x_1, \dots, x_4, y_1, \dots, y_4$, where $\mathbf{x}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$.

The question is whether there exists $\epsilon > 0$ such that X_2^ϵ is the graph of a C^1 map representing four of the eight variables in terms of the other four. We will show that this is not the case: indeed, X_2^ϵ is not the graph of *any* map (C^1 or otherwise) representing four of the eight variables in terms of the other four.

There are many (8 choose 4, i.e., 70) possible choices of candidate "domain variables" (the active variables), which we will call *candidate coordinates*; they will be *coordinates* if the other variables are indeed functions of these. Let us make some preliminary simplifications, which will eliminate many of the candidates. Clearly y_1, y_2, y_3, y_4 are not a possible set of coordinates, since we can add any fixed constant to x_1, x_2, x_3, x_4 and stay in X_2. Also, no two linked x_i can belong to any set of possible coordinates, since $|x_{i+1} - x_i| \le l_i$. Any choice of three x_i-variables will contain a pair of linked ones, so any set of coordinates must include either two x-variables and two y-variables, or one x-variable and three y-variables.

We can choose the values of the coordinate functions to be whatever we like; in particular, they can be 0. At the base position A, all the y_i vanish.

The two cases are different. If our candidate coordinates consist of two x-variables (easily seen to be necessarily x_2 and x_4), and two y-variables, then in every neighborhood of the position A, there exists a position X where the two coordinate y-variables are 0, and the two non-coordinate y-variables are nonzero. Indeed, as shown by the figure in the margin, if you can choose x_2 and x_4 so that $|x_2 - x_4| > l_2 - l_3$, and then the linkage will not lie on a line. If

$$X = \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ y_3 \end{pmatrix}, \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} \right) \in X_2^\epsilon$$

then

$$X' = \left(\begin{pmatrix} x_1 \\ -y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ -y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ -y_3 \end{pmatrix}, \begin{pmatrix} x_4 \\ -y_4 \end{pmatrix} \right) \in X_2^\epsilon$$

also and has the same values for the candidate coordinates. Thus these candidate coordinates fail: their values do not determine the other four variables.

When the candidate consists of one x-coordinate and three y-coordinates, two y-coordinates must belong to non-linked points, say y_1 and y_3 (y_2 and y_4 would do just as well), and without loss of generality, we may assume that the third coordinate y-variable is y_2. There is then a position X where $y_1 = y_3 = 0,\ y_2 \neq 0$; it then follows that $y_4 \neq 0$ also. If

$$X = \left(\begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ 0 \end{pmatrix}, \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} \right) \in X_2^\epsilon$$

then

$$X' = \left(\begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}, \begin{pmatrix} x_3 \\ 0 \end{pmatrix}, \begin{pmatrix} x_4 \\ -y_4 \end{pmatrix} \right) \in X_2^\epsilon$$

also and has the same values for the candidate coordinates. Thus the candidate y_1, y_2, y_3, and x_j are not coordinates for any j, as they do not determine y_4.

The upshot is that there is no successful candidate for coordinate variables, so X_2 is not a manifold near A.

3.1.25 a. Set $F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = z - (x - \sin z)(\sin z - x + y)$. We need to check that the equation $F\left(\mathbf{g}\begin{pmatrix} u \\ v \end{pmatrix}\right) = 0$ is satisfied for all $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2$. This is straightforward:

$$\begin{aligned} F\left(\mathbf{g}\begin{pmatrix} u \\ v \end{pmatrix}\right) &= uv - \big(\sin uv + u - \sin uv\big)\big(\sin uv - (\sin uv + u) + (u + v)\big) \\ &= uv - (u)(v) = 0. \end{aligned}$$

b. We need to show that $[\mathbf{D}F(\mathbf{x})] \neq [0, 0, 0]$ for all $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in S$; it is actually true for all $\mathbf{x} \in \mathbb{R}^3$. Indeed, using the product rule, we compute

$$\left[\mathbf{D}F\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = [\underbrace{-(\sin z - x + y) + (x - \sin z)}_{D_1F(\mathbf{x})}\ \underbrace{-(x - \sin z)}_{D_2F(\mathbf{x})}\ \underbrace{1 + \cos z(\sin z - x + y) - \cos z(x - \sin z)}_{D_3F(\mathbf{x})}].$$

When will all three entries be 0? For $D_2F(\mathbf{x})$ to be 0, we need to have $x - \sin z = 0$, but then (looking at the first entry) $y = 0$, and then (looking at the third entry) $1 = 0$, which is a contradiction. So the derivative of F never vanishes on $\mathbb{R}^3$.

c. Given a point $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in S$, we need to show that $\mathbf{g}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ can be solved for u, v. Clearly we can just solve the equations

$$\begin{aligned} x &= \sin uv + u \\ y &= u + v \\ z &= uv. \end{aligned}$$

Part c illustrates that showing that a map is onto is equivalent to showing that some equation has a solution.

The first and third equations give $u = x - \sin z$, and the second then gives $v = y - x + \sin z$. We still need to check that this actually is a solution. Recall that the point $\mathbf{x}$ is a point of S, in particular

$$F(\mathbf{x}) = z - (x - \sin z)(\sin z - x + y) = z - uv = 0,$$

so the third equation is satisfied, and then from the definition of u the first is satisfied, and from the definition of v the second is satisfied.

d. To see that $\mathbf{g}$ is injective, suppose that $\mathbf{g}\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \mathbf{g}\begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$, i.e.,

$$\begin{aligned} \sin u_1 v_1 + u_1 &= \sin u_2 v_2 + u_2 \\ u_1 + v_1 &= u_2 + v_2 \\ u_1 v_1 &= u_2 v_2. \end{aligned}$$

The first and third equations imply $u_1 = u_2$, and then the second equation implies $v_1 = v_2$. Thus $\mathbf{g}$ is injective.

Saying that the derivative

$$\left[\mathbf{Dg}\begin{pmatrix} u \\ v \end{pmatrix}\right] : \mathbb{R}^2 \to \mathbb{R}^3$$

is injective means that its image is a plane in $\mathbb{R}^3$, not a line or a point. It does not mean that the map $\begin{pmatrix} u \\ v \end{pmatrix} \mapsto \left[\mathbf{Dg}\begin{pmatrix} u \\ v \end{pmatrix}\right]$ is injective.

To show that the derivative $\left[\mathbf{Dg}\begin{pmatrix} u \\ v \end{pmatrix}\right]$ is 1–1 at every $\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2$, it is enough to show that $\left[\mathbf{Dg}\begin{pmatrix} u \\ v \end{pmatrix}\right]\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ implies $a = b = 0$. Write

$$\left[\mathbf{Dg}\begin{pmatrix} u \\ v \end{pmatrix}\right]\begin{bmatrix} a \\ b \end{bmatrix} = a\begin{bmatrix} v\cos uv + 1 \\ 1 \\ v \end{bmatrix} + b\begin{bmatrix} u\cos uv \\ 1 \\ u \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The second line tells us that $a = -b$, and the third line says we must have $u = v$. The first line then becomes $u\cos u^2 + 1 = u\cos u^2$, a contradiction, so we must have $a = b = 0$. Thus the derivative of $\mathbf{g}$ is everywhere injective.

3.2.1 a. At $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ the derivative is $[3\,,\,0]$, so the equation for the tangent line is $3(x-1) = 0$, i.e., $x = 1$.

b. At the same point, the tangent space is the kernel of $[3\,,\,0]$, which is the line of equation $\dot{x} = 0$, i.e., the y-axis.

3.2.3 A point $\begin{pmatrix} x \\ y \end{pmatrix}$ is in the tangent line to X_c at $\begin{pmatrix} u \\ v \end{pmatrix}$ if and only if $\begin{pmatrix} x-u \\ y-v \end{pmatrix} \in \ker\,[2u,\ 3v^2]$, i.e., when $2u(x-u) + 3v^2(y-v) = 0$.

The tangent space $T_{\begin{pmatrix} u \\ v \end{pmatrix}} X_c$ is the kernel of $[2u,\ 3v^2]$, i.e., the 1-dimensional subspace of $\mathbb{R}^2$ of equation $2u\dot{x} + 3v^2\dot{y} = 0$.

3.2.5 a. At the point $\begin{pmatrix} u \\ v \\ Au^2+Bv^2 \end{pmatrix}$, the tangent plane to the surface of equation $Ax^2+By^2-z=0$ has equation

$$[\,2Au \quad 2Bv \quad -1\,]\begin{bmatrix} x-u \\ y-v \\ z-(Au^2+Bv^2) \end{bmatrix}=0.$$

Applied to our three points, this means that P_1, P_2, P_3 have equations

$$[\,0 \quad 0 \quad -1\,]\begin{bmatrix} x \\ y \\ z \end{bmatrix}=0, \quad [\,2Aa \quad 0 \quad -1\,]\begin{bmatrix} x-a \\ y \\ z-Aa^2 \end{bmatrix}=0, \quad [\,0 \quad 2Bb \quad -1\,]\begin{bmatrix} x \\ y-b \\ z-Bb^2 \end{bmatrix}=0,$$

which expand to

$$\begin{aligned} z &= 0 \\ 2aAx - z &= a^2A \\ 2bBy - z &= b^2B. \end{aligned}$$

This is easy to solve; we find $\mathbf{q}=\begin{pmatrix} a/2 \\ b/2 \\ 0 \end{pmatrix}$.

b. The volume of the tetrahedron is

$$\frac{1}{6}\left|\det\begin{bmatrix} a & 0 & a/2 \\ 0 & b & b/2 \\ a^2A & b^2B & 0 \end{bmatrix}\right| = \frac{ab}{12}(a^2A+b^2B).$$

3.2.7 For any $\begin{pmatrix} x \\ y \end{pmatrix}$ near $\begin{pmatrix} a \\ b \end{pmatrix}$, the point $f\begin{pmatrix} x \\ y \end{pmatrix}$ is the unique z such that $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is on the surface S. Similarly, $\begin{pmatrix} x \\ g(x,z) \\ z \end{pmatrix}$ and $\begin{pmatrix} h(y,z) \\ y \\ z \end{pmatrix}$ are on S. Note that $f\begin{pmatrix} h(b,z) \\ b \end{pmatrix}=z$ and $g\begin{pmatrix} h(y,c) \\ c \end{pmatrix}=y$. After differentiating these equations with respect to the appropriate variables (z for the first, y for the second) and using the chain rule, we discover that $(D_1f)(D_2h)=1$ and $(D_1g)(D_1h)=1$, where all derivatives are evaluated at $\mathbf{a}=\begin{pmatrix} a \\ b \\ c \end{pmatrix}$. These equations imply that

$$D_1f\neq 0, \quad D_1g\neq 0, \quad D_1h=\frac{1}{D_1g}, \quad D_2h=\frac{1}{D_1f}.$$

We can also combine the functions like this:

$$f\begin{pmatrix} h(y,c) \\ y \end{pmatrix}=c \quad \text{and} \quad g\begin{pmatrix} h(b,z) \\ z \end{pmatrix}=b.$$

After differentiating the first equation with respect to y and the second with respect to z (and using the chain rule) we get

$$(D_1f)(D_1h)+(D_2f)=0 \quad \text{and} \quad (D_1g)(D_2h)+(D_2g)=0.$$

Since $D_1 f$ and $D_1 g$ do not equal zero,

$$D_1 h = -\frac{D_2 f}{D_1 f} \quad \text{and} \quad D_2 h = -\frac{D_2 g}{D_1 g}.$$

Since we know that $D_1 f$ and $D_1 g$ do not equal 0, the equations for the tangent planes at **a** may be written as follows by linearizing the equations $x = h(y,z)$, $y = g(x,z)$, and $z = f(x,y)$ and rearranging terms:

$$\begin{aligned} \dot{x} &= -\frac{D_2 f}{D_1 f}\dot{y} + \frac{1}{D_1 f}\dot{z} \\ \dot{x} &= \ \frac{1}{D_1 g}\dot{y} - \frac{D_2 g}{D_1 g}\dot{z} \\ \dot{x} &= \ (D_1 h)\dot{y} + (D_2 h)\dot{z}. \end{aligned}$$

From the equations we derived above, we see that the first and third tangent planes are the same, as are the second and third. So all tangent planes are the same.

3.2.9 The tangent plane to CX through a point $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ has equation

$$D_1 f \begin{pmatrix} x_0/z_0 \\ y_0/z_0 \end{pmatrix} \frac{x - x_0}{z_0} + D_2 f \begin{pmatrix} x_0/z_0 \\ y_0/z_0 \end{pmatrix} \frac{y - y_0}{z_0} = \frac{z - z_0}{z_0^2} \left(x_0 D_1 f \begin{pmatrix} x_0/z_0 \\ y_0/z_0 \end{pmatrix} + y_0 D_2 f \begin{pmatrix} x_0/z_0 \\ y_0/z_0 \end{pmatrix} \right).$$

Note that this is the plane through the origin that contains the tangent line to the curve X at the point $\begin{pmatrix} x_0/z_0 \\ y_0/z_0 \\ 1 \end{pmatrix}$.

3.2.11 a. If $A \in O(n)$, then $A^\top A = I$, so $A^\top = A^{-1}$. Moreover, if $A, B \in O(n)$, then $(AB)^\top (AB) = B^\top A^\top AB = B^\top B = I$; this shows that a product of orthogonal matrices is orthogonal. Since

$$(A^{-1})^\top A^{-1} = (A^\top)^\top A^{-1} = AA^{-1} = I,$$

we see that A^{-1} is also orthogonal.

b. If $n \mapsto A_n$ is a convergent sequence of orthogonal matrices, converging to A_∞, then

$$A_\infty^\top A_\infty = (\lim_{n\to\infty} A_n^\top)(\lim_{n\to\infty} A_n) \underbrace{=}_{\text{Thm. 1.5.26, part 4}} \lim_{n\to\infty} (A_n^\top A_n) = \lim_{n\to\infty} I = I,$$

so the $O(n)$ is closed in Mat (n,n). It is also bounded; all the columns are of unit length, so $\sum_{i,j=1}^n a_{i,j}^2 = n$.

c. Just compute:

$$(A^\top A - I)^\top = (A^\top A)^\top - I = A^\top (A^\top)^\top - I = A^\top A - I.$$

d. We have $[\mathbf{D}F(A)]H = A^\top H + H^\top A$; we need to show that given any $M \in S(n,n)$, there exists $H \in \text{Mat}\,(n,n)$ such that $A^\top H + H^\top A = M$.

Solution 3.2.11, part d: To compute the derivative $[\mathbf{D}F(A)]$, compute

$$F(A+H) - F(A)$$

and discard all terms in the increment H that are quadratic or higher; see the remark after Example 1.7.17.

Try $H = \frac{1}{2}AM$: since $M = M^\top$, we have

$$A^\top \left(\frac{1}{2}AM\right) + \left(\frac{1}{2}AM\right)^\top A = \frac{1}{2}A^\top AM + \frac{1}{2}M^\top A^\top A = \frac{1}{2}M + \frac{1}{2}M^\top = M.$$

e. That $O(n)$ is a manifold follows immediately from Theorem 3.1.10, and the characterization of the tangent space follows from Theorem 3.2.4: it is $\ker\,[\mathbf{D}F(I)]$, which is the set of $H \in \text{Mat}\,(n,n)$ such that $H + H^\top = [0]$, which is precisely the set of antisymmetric matrices.

3.3.1

$$\begin{aligned} D_1(D_2 f) &= D_1(x^2 + 2xy + z^2) = 2x + 2y \\ D_2(D_3 f) &= D_2(2zy) = 2z \\ D_3(D_1 f) &= D_3(2xy + y^2) = 0 \\ D_1(D_2(D_3 f)) &= D_1(D_2((2yz)) = D_1(2z) = 0. \end{aligned}$$

3.3.3 We have

$$P^2_{f,\mathbf{a}}(\mathbf{a} + \vec{\mathbf{h}}) = \sum_{m=0}^{2} \sum_{I \in \mathcal{I}_3^m} \frac{1}{I!} D_I f(\mathbf{a}) \vec{\mathbf{h}}^I = \underbrace{\frac{1}{0!0!} D_{(0,0,0)} f(\mathbf{a}) h_1^0 h_2^0 h_3^0}_{f(\mathbf{a})} \; +$$

$$\begin{aligned} &+ \frac{1}{1!0!0!} D_{(1,0,0)} f(\mathbf{a}) h_1^1 h_2^0 h_3^0 + \frac{1}{0!1!0!} D_{(0,1,0)} f(\mathbf{a}) h_1^0 h_2^1 h_3^0 + \frac{1}{0!0!1!} D_{(0,0,1)} f(\mathbf{a}) h_1^0 h_2^0 h_3^1 \\ &+ \frac{1}{2!0!0!} D_{(2,0,0)} f(\mathbf{a}) h_1^2 h_2^0 h_3^0 + \frac{1}{1!1!0!} D_{(1,1,0)} f(\mathbf{a}) h_1 h_2 h_3^0 + \frac{1}{1!0!1!} D_{(1,0,1)} f(\mathbf{a}) h_1 h_2^0 h_3 \\ &+ \frac{1}{0!1!1!} D_{(0,1,1)} f(\mathbf{a}) h_1^0 h_2 h_3 + \frac{1}{0!2!0!} D_{(0,2,0)} f(\mathbf{a}) h_1^0 h_2^2 h_3^0 + \frac{1}{0!0!2!} D_{(0,0,2)} f(\mathbf{a}) h_1^0 h_2^0 h_3^2; \end{aligned}$$

i.e.,

$$\begin{aligned} P^2_{f,\mathbf{a}}(\mathbf{a} + \vec{\mathbf{h}}) &= f(\mathbf{a}) + D_{(1,0,0)} f(\mathbf{a}) h_1 + D_{(0,1,0)} f(\mathbf{a}) h_2 + D_{(0,0,1)} f(\mathbf{a}) h_3 \\ &+ \frac{1}{2} D_{(2,0,0)} f(\mathbf{a}) h_1^2 + D_{(1,1,0)} f(\mathbf{a}) h_1 h_2 + D_{(1,0,1)} f(\mathbf{a}) h_1 h_3 \\ &+ D_{(0,1,1)} f(\mathbf{a}) h_2 h_3 + \frac{1}{2} D_{(0,2,0)} f(\mathbf{a}) h_2^2 + \frac{1}{2} D_{(0,0,2)} f(\mathbf{a}) h_3^2\,. \end{aligned}$$

3.3.5

$$\begin{aligned} \text{cardinality } \mathcal{I}_1^m &= \text{ cardinality } \{(m)\} = 1; \\ \text{cardinality } \mathcal{I}_2^m &= \text{ cardinality } \{(0,m), (1, m-1), \dots, (m,0)\}, = m+1; \\ \text{cardinality } \mathcal{I}_3^m &= \text{ cardinality } \{(0, \mathcal{I}_2^m), (1, \mathcal{I}_2^{m-1}), \dots, (m, \mathcal{I}_2^0)\} \\ &= (m+1) + m + \dots + 2 + 1 = \frac{(m+1)(m+2)}{2} \end{aligned}$$

In general,

$$\begin{aligned} \text{cardinality } \mathcal{I}_n^m = \text{cardinality } \mathcal{I}_{n-1}^0 + \text{cardinality } \mathcal{I}_{n-1}^1 + \cdots \\ \cdots + \text{cardinality } \mathcal{I}_{n-1}^{m-1} + \text{cardinality } \mathcal{I}_{n-1}^m, \end{aligned}$$

i.e.,

$$\text{cardinality } \mathcal{I}_n^m = \text{cardinality } \mathcal{I}_n^{m-1} + \text{cardinality } \mathcal{I}_{n-1}^m. \tag{1}$$

Remark. This is reminiscent of the *Pascal's triangle recursion for binomials*,

$$\binom{m}{r} = \binom{m-1}{r-1} + \binom{m-1}{r}. \tag{2}$$

Indeed, setting $m = n+k-1, r = n-1$ gives

$$\binom{n+k-1}{n-1} = \binom{n+k-2}{n-1} + \binom{n+k-2}{n-2}.$$

Thus the formula

$$\text{cardinality } \mathcal{I}_n^k = \binom{n-1+k}{n-1}, \tag{3}$$

satisfies equation 1. Moreover, cardinality $\mathcal{I}_n^k$ satisfies the boundary conditions of Pascal's triangle, namely,

$$\text{cardinality } \mathcal{I}_1^k = 1 = \binom{k}{0} \quad \text{and} \quad \text{cardinality } \mathcal{I}_n^0 = 1 = \binom{n-1}{n-1},$$

and these together with equation 2 completely specify the binomial coefficients. Thus equation 3 is true. $\triangle$

Here is a very nice different solution to Exercise 3.3.5, proposed by Nathaniel Schenker:

Using the "stars and bars" method, consider the problem as counting the ways of allocating m indistinguishable balls to n distinguishable bins. Each configuration of the balls in the bins corresponds to an $I \in \mathcal{I}_n^m$, with each ball in bin j, for $j \in \{1, \dots, n\}$, adding 1 to i_j. Each $I \in \mathcal{I}_n^m$ is represented by a sequence of length $m+n-1$ containing m stars (representing the m balls) and $n-1$ bars (which divide the stars into bins). For example, in $\mathcal{I}_4^3$, we represent $(0,2,0,1)$ by $|**||*$.

Any choice of positions of the m stars in the sequence automatically implies the positions of the $n-1$ bars, and vice versa. Thus the cardinality of $\mathcal{I}_n^m$ is the number of choices of positions for the stars, that is, $\binom{m+n-1}{m}$, or, equivalently, the number of choices of positions for the bars, that is, $\binom{m+n-1}{n-1}$.

3.3.7 If $I \in \mathcal{I}_n^m$, then $\vec{\mathbf{h}}^I = h_1^{i_1} h_2^{i_2} \cdots h_n^{i_n}$, where $i_1 + i_2 + \cdots + i_n = m$, the total degree. Therefore

$$(x\vec{\mathbf{h}})^I = x^{i_1} h_1^{i_1} x^{i_2} h_2^{i_2} \cdots x^{i_n} h_n^{i_n}, \quad \text{i.e.,} \quad x^{i_1 + \cdots + i_n} \vec{\mathbf{h}}^I = x^m \vec{\mathbf{h}}^I.$$

3.3.9 a. Clearly this function has partial derivatives of all orders everywhere, except perhaps at the origin. It is also clear that both first partials exist at the origin, and are 0 there, since f vanishes identically on the axes.

Elsewhere, the partials are given by

$$D_1 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^4y + 3x^2y^3 - 2xy^4}{(x^2+y^2)^2} \quad \text{and} \quad D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{x^5 - x^3y^2 - 2x^4y}{(x^2+y^2)^2}.$$

These partials are continuous on all of $\mathbb{R}^2$, since the limits of the expressions above are 0 as $x, y \to 0$. In particular, f is of class C^1.

b. and c. Since $D_1 f\begin{pmatrix} x \\ 0 \end{pmatrix} = 0$ for all x, and $D_2 f\begin{pmatrix} 0 \\ y \end{pmatrix} = 0$ for all y,

$$D_1(D_1 f)\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \quad \text{and} \quad D_2(D_2 f)\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

Since $D_1 f\begin{pmatrix} 0 \\ y \end{pmatrix} = 0$ for all y and $D_2 f\begin{pmatrix} x \\ 0 \end{pmatrix} = x$ for all x,

$$D_2\left(D_1 f\begin{pmatrix} 0 \\ y \end{pmatrix}\right) = 0 \quad \text{and} \quad D_1\left(D_2 f\begin{pmatrix} x \\ 0 \end{pmatrix}\right) = 1,$$

so the crossed partials are not equal.

d. This does not contradict Theorem 3.3.8 because the first partial derivative $D_2 f$ is not differentiable. The definition of $D_2 f$ being differentiable at $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ is that there should exist a line matrix $[a, b]$ such that

$$\lim_{|\vec{\mathbf{h}}| \to 0} \frac{\left| D_2 f(\mathbf{0} + \vec{\mathbf{h}}) - D_2 f(\mathbf{0}) - [a, b]\vec{\mathbf{h}} \right|}{|\vec{\mathbf{h}}|} = 0. \tag{1}$$

We determine what $[a, b]$ must be by considering $\vec{\mathbf{h}} = \begin{bmatrix} h_1 \\ 0 \end{bmatrix}$ and $\vec{\mathbf{h}} = \begin{bmatrix} 0 \\ h_2 \end{bmatrix}$ (and using the value for $D_2 f$ given in part a):

$$\lim_{h_1 \to 0} \frac{\left| D_2 f\begin{pmatrix} h_1 \\ 0 \end{pmatrix} - D_2 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} - [a, b]\begin{bmatrix} h_1 \\ 0 \end{bmatrix} \right|}{|h_1|} = \frac{|h_1 - ah_1|}{|h_1|}$$

$$\lim_{h_2 \to 0} \frac{\left| D_2 f\begin{pmatrix} 0 \\ h_2 \end{pmatrix} - D_2 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} - [a, b]\begin{bmatrix} 0 \\ h_2 \end{bmatrix} \right|}{|h_2|} = \frac{|b||h_2|}{|h_2|} = 0.$$

To satisfy equation (1), these imply that $[a, b] = [1, 0]$.

But then

$$\lim_{\begin{pmatrix} h \\ h \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}} \frac{\left| D_2 f\begin{pmatrix} h \\ h \end{pmatrix} - D_2 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} - [1, 0]\begin{bmatrix} h \\ h \end{bmatrix} \right|}{\left| \begin{bmatrix} h \\ h \end{bmatrix} \right|} = \frac{\left| \frac{-2h^5}{(2h^2)^2} - h \right|}{\sqrt{2}|h|} = \frac{\frac{3}{2}|h|}{\sqrt{2}|h|} = \frac{3}{2\sqrt{2}} \neq 0.$$

Solution 3.3.11: $f^{(i)}$ denotes the ith derivative of f.

3.3.11 The jth derivative of the numerator of the fraction given in the hint is

$$f^{(j)}(a+h) - \sum_{i=j}^{k} \frac{f^{(i)}(a)}{(i-j)!} h^{i-j},$$

and the jth derivative of the denominator of the fraction given in the hint is

$$\frac{k!}{(k-j)!}h^{k-j}.$$

If we evaluate these two expressions at $h = 0$, they yield 0 for $0 \le j < k$. Thus the hypotheses of l'Hôpital's rule are satisfied until we take the kth derivative, at which point the numerator is 0 and the denominator is $k!$ when evaluated at $h = 0$. Thus the limit in question is 0.

For uniqueness, suppose p_1 and p_2 are two polynomials of degree $\le k$ such that

$$\lim_{h\to 0}\frac{f(a+h)-p_1(a+h)}{h^k} = 0, \quad \lim_{h\to 0}\frac{f(a+h)-p_2(a+h)}{h^k} = 0.$$

By subtraction, we see that

$$\lim_{h\to 0}\frac{p_1(a+h)-p_2(a+h)}{h^k} = 0.$$

If $p_1 \ne p_2$, then $p_1(a+h) - p_2(a+h) = b_0 + \cdots + b_k h^k$ has a first nonvanishing term, i.e., there exists $l \le k$ such that $b_0 = \cdots = b_{l-1} = 0$ and $b_l \ne 0$. Then the limit

$$\begin{aligned}\lim_{h\to 0}\frac{p_1(a+h)-p_2(a+h)}{h^k} &= \lim_{h\to 0}\frac{b_l h^l + \cdots + b_k h^k}{h^k}\\ &= \lim_{h\to 0}\frac{1}{h^{k-l}}(b_l + \cdots + b_k h^{k-l})\end{aligned}$$

is zero only if $\lim_{h\to 0}\frac{1}{h^{k-l}} = 0$, and that is not the case when $l \le k$.

3.3.13 First, we compute the partial derivatives and second partials:

$$D_{(1,0)}f = \frac{1+y}{2\sqrt{x+y+xy}} \qquad D_{(0,1)}f = \frac{1+x}{2\sqrt{x+y+xy}}$$

$$D_{(2,0)}f = \frac{\sqrt{x+y+xy}\cdot 0 - (1+y)\frac{1+y}{2\sqrt{x+y+xy}}}{2(x+y+xy)} = \frac{-(1+y)^2}{4(x+y+xy)^{3/2}}$$

$$D_{(1,1)}f = \frac{\sqrt{x+y+xy}\cdot 1 - \frac{(1+y)(1+x)}{2\sqrt{x+y+xy}}}{2(x+y+xy)} = \frac{2(x+y+xy)-(1+y)(1+x)}{4(x+y+xy)^{3/2}}$$

$$D_{(0,2)}f = \frac{\sqrt{x+y+xy}\cdot 0 - (1+x)\frac{1+x}{2\sqrt{x+y+xy}}}{2(x+y+xy)} = -\frac{(1+x)^2}{4(x+y+xy)^{3/2}}.$$

At the point $\begin{pmatrix}-2\\-3\end{pmatrix}$ these are

$$D_{(1,0)}f\begin{pmatrix}-2\\-3\end{pmatrix} = \frac{1-3}{2\sqrt{-2-3+6}} = -1, \qquad D_{(0,1)}f\begin{pmatrix}-2\\-3\end{pmatrix} = \frac{1-2}{2\cdot 1} = -\frac{1}{2},$$

$$D_{(2,0)}f\begin{pmatrix}-2\\-3\end{pmatrix} = \frac{-(1-3)^2}{4\cdot 1} = -1, \qquad D_{(1,1)}f\begin{pmatrix}-2\\-3\end{pmatrix} = \frac{2\cdot 1-(-2)(-1)}{4} = 0,$$

$$D_{(0,2)}f\begin{pmatrix}-2\\-3\end{pmatrix} = -\frac{(-1)^2}{4\cdot 1} = -\frac{1}{4}.$$

Write $x = -2+u,\ y = -3+v$; i.e., the increment $\vec{\mathbf{h}}$ is $\begin{pmatrix} u \\ v \end{pmatrix}$. The Taylor polynomial of degree 2 of f at $\begin{pmatrix} -2 \\ -3 \end{pmatrix}$ is then $1 - u - \frac{1}{2}v + \frac{1}{2}(-u^2 - \frac{1}{4}v^2)$:

$$P^2_{f,\begin{pmatrix} -2 \\ -3 \end{pmatrix}}\begin{pmatrix} -2+u \\ -3+v \end{pmatrix} = \underbrace{1}_{f\begin{pmatrix} -2 \\ -3 \end{pmatrix}} + \underbrace{-u}_{D_{1,0}f\begin{pmatrix} -2 \\ -3 \end{pmatrix}u^1v^0} + \underbrace{-\frac{1}{2}v}_{D_{(0,1)}f\begin{pmatrix} -2 \\ -3 \end{pmatrix}u^0v^1} + \underbrace{-\frac{1}{2}u^2}_{\frac{1}{2}D_{(2,0)}f\begin{pmatrix} -2 \\ -3 \end{pmatrix}u^2v^0} + \underbrace{\frac{1}{2}\left(-\frac{1}{4}v^2\right)}_{\frac{1}{2}D_{(0,2)}f\begin{pmatrix} -2 \\ -3 \end{pmatrix}u^0v^2}.$$

3.3.15 We know (Theorem 3.1.10) that there exist a neighborhood $U \subset \mathbb{R}^n$ of $\mathbf{a}$ and a C^1 map $\mathbf{f}: U \to R^{n-k}$ with $[\mathbf{Df}(\mathbf{a})]$ onto, such that

$$M \cap U = \{\, \mathbf{u} \mid \mathbf{f}(\mathbf{u}) = \mathbf{0} \,\}.$$

Further, by Theorem 3.2.4, $T_{\mathbf{a}}M = \ker[\mathbf{Df}(\mathbf{a})]$. Thus $U_1 \overset{\text{def}}{=} \mathbf{F}^{-1}(U)$ is a neighborhood of $\mathbf{0} \in \mathbb{R}^n$. Define $\mathbf{g}: U_1 \to \mathbb{R}^{n-k}$ by $\mathbf{g}(\mathbf{u}) = \mathbf{f} \circ \mathbf{F}(\mathbf{u})$. Then

$$\mathbf{F}^{-1}(M) \cap U_1 = \{\, \mathbf{u}_1 \mid \mathbf{g}(\mathbf{u}_1) = \mathbf{0} \,\}.$$

Since

$$[\mathbf{Dg}(\mathbf{0})] = [\mathbf{D}(\mathbf{f} \circ F)(\mathbf{0})] = [\mathbf{Df}(F(\mathbf{0}))][\mathbf{D}F(\mathbf{0})] = [\mathbf{Df}(\mathbf{a})]A,$$

and both $[\mathbf{Df}(\mathbf{a})]$ and A are onto, $[\mathbf{Dg}(\mathbf{0})]$ is onto. But, since

$$\vec{\mathbf{v}} \in \ker[\mathbf{Dg}(\mathbf{0})] \implies A\vec{\mathbf{v}} \in \ker[\mathbf{Df}(\mathbf{a})] \implies \vec{\mathbf{v}} \in A^{-1}\big(\ker[\mathbf{Df}(\mathbf{a})]\big),$$

we have that

$$\ker[\mathbf{Dg}(\mathbf{0})] = A^{-1}(\ker[\mathbf{Df}(\mathbf{a})]) = A^{-1}(T_{\mathbf{a}}M)$$

Here, $\mathbf{g}$ plays the role of $\mathbf{F}$ in the implicit function theorem, and $\mathbf{h}$ plays the role of $\mathbf{g}$.

is the span of $\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_k$. Thus we have $D_1\mathbf{g}(\mathbf{0}) = \cdots = D_k\mathbf{g}(\mathbf{0}) = 0$, so $[D_{k+1}\mathbf{g}(\mathbf{0}), \dots, D_n\mathbf{g}(\mathbf{0})]$ is invertible. Thus the implicit function theorem (Theorem 2.10.14) asserts that there exists a function $\mathbf{h}$ such that $\mathbf{g}(\mathbf{u}_1) = \mathbf{0}$ expresses the last $n-k$ variables as a function $\mathbf{h}$ of the first k variables: $\mathbf{g}\begin{pmatrix} \mathbf{x} \\ \mathbf{h}(\mathbf{x}) \end{pmatrix} = \mathbf{0}$, where $\mathbf{x}$ corresponds to the first k variables. So locally $\mathbf{F}^{-1}(M)$ is the graph of $\mathbf{h}$. ("Locally" must be within U_1, but it may be smaller.) Moreover, by equation 2.10.29,

$$[\mathbf{Dh}(\mathbf{0})] = -[D_{k+1}\mathbf{g}(\mathbf{0}), \dots, D_n\mathbf{g}(\mathbf{0})]^{-1}\underbrace{[D_1\mathbf{g}(\mathbf{0}), \dots, D_k\mathbf{g}(\mathbf{0})]}_{[0]} = [0].$$

3.4.1 Using equation 3.4.9, with $m = 1/2$ and $\sin(x+y)$ playing the role of x in that equation, we have

$$\big(1 + \sin(x+y)\big)^{1/2} = 1 + \frac{1}{2}\sin(x+y) + \frac{-\frac{1}{4}}{2!}\big(\sin(x+y)\big)^2 + \cdots.$$

Equation 3.4.6 tells us that $\sin(x+y) = x + y - \frac{(x+y)^3}{3!}$ plus higher degree terms; discarding terms greater than 2 leaves just $x+y$. Therefore:

$$\sqrt{1+\sin(x+y)} = 1 + \frac{1}{2}(x+y) - \frac{1}{8}(x+y)^2 + \cdots.$$

We could also use the chain rule in the other direction, first developing the sine, then the square root:

$$\sqrt{1+\sin(x+y)} = \sqrt{1+(x+y)+\cdots} = 1 + \frac{1}{2}(x+y) - \frac{1}{8}(x+y)^2 + \cdots .$$

3.4.3 Set $\mathbf{a} = \begin{pmatrix}-2\\-3\end{pmatrix}$ and $\mathbf{u} = \begin{pmatrix}u\\v\end{pmatrix}$. Then

Going from the first to the second line uses the binomial formula, equation 3.4.9.

$$\begin{aligned} f(\mathbf{a}+\mathbf{u}) &= \Big(-2+u-3+v+(-2+u)(-3+v)\Big)^{1/2} = (1-2u-v+uv)^{1/2} \\ &= 1+\frac{1}{2}(-2u-v+uv) - \frac{1}{8}(-2u-v+uv)^2+\cdots \\ &= 1-u-\frac{1}{2}v+\frac{1}{2}uv-\frac{1}{8}(4u^2+v^2+4uv+\cdots)+\cdots \end{aligned}$$

Therefore,

$$P^2_{f,a}(\mathbf{a}+\mathbf{u}) = 1-u-\frac{1}{2}v+\frac{1}{2}uv-\frac{1}{2}u^2-\frac{1}{8}v^2-\frac{1}{2}uv = 1-u-\frac{1}{2}v-\frac{1}{2}u^2-\frac{1}{8}v^2.$$

3.4.5 Write $f(x) = A+Bx+Cx^2+R(x)$ with $R(x)\in o(h^2)$. Then

$$\begin{aligned} h\big(af(0)+bf(h)+cf(2h)\big) &= h\Big(aA+b(A+Bh+Ch^2)+c(A+2Bh+4Ch^2)\Big)+h\Big(bR(h)+cR(2h)\Big) \\ &= hA(a+b+c)+h^2B(b+2c)+h^3C(b+4c)+h\big(bR(h)+cR(2h)\big); \end{aligned}$$

note that $h\big(bR(h)+cR(2h)\big)\in o(h^3)$.

On the other hand,

$$\int_0^{2h} f(t)\,dt = 2Ah+4B\frac{h^2}{2}+8C\frac{h^3}{3}+\int_0^{2h} R(t)\,dt.$$

Note that $\int_0^{2h} R(t)\,dt \in o(h^3)$. Thus we find

$$\begin{aligned} a+b+c &= 2 \\ b+2c &= 2 \\ b+4c &= \frac{8}{3} \end{aligned} \qquad \text{with solution} \qquad \begin{aligned} a &= \frac{1}{3} \\ b &= \frac{4}{3} \\ c &= \frac{1}{3} \end{aligned}.$$

Our error terms above show that with these numbers, the equation

$$h\Big(af(0)+bf(h)+cf(2h)\Big) - \int_0^h f(t)\,dt \in o(h^3)$$

is satisfied by all functions f of class C^3.

3.4.7 Let $f\begin{pmatrix}x\\y\\z\end{pmatrix} = \sin(xyz)-z = 0$ and $\mathbf{a} = \begin{pmatrix}\pi/2\\1\\1\end{pmatrix}$. Then

$$D_3f\begin{pmatrix}x\\y\\z\end{pmatrix} = xy\cos(xyz)-1, \qquad f(\mathbf{a})=0, \qquad D_3f(\mathbf{a}) = -1\neq 0,$$

so $[\mathbf{D}f(\mathbf{a})]$ is onto, and by the implicit function theorem (short version), the equation $f = 0$ expresses one variable as a function of the other two. Since $D_3 f(\mathbf{a}) = -1$ is invertible, the long version of the implicit function theorem says that we can choose z as the "pivotal" or "passive" variable that is a function of the others: $g\begin{pmatrix} x \\ y \end{pmatrix} = z$ in a neighborhood of $\mathbf{p} = \begin{pmatrix} \pi/2 \\ 1 \end{pmatrix}$, where $g(\mathbf{p}) = 1$.

By definition,

$$P^2_{g,\mathbf{p}}\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \end{pmatrix} = \underbrace{1}_{g(\mathbf{p})} + D_1 g(\mathbf{p})\dot{x} + D_2 g(\mathbf{p})\dot{y} + \frac{1}{2}D_1^2 g(\mathbf{p})\dot{x}^2 + D_1 D_2 g(\mathbf{p})\dot{x}\dot{y} + \frac{1}{2}D_2^2 g(\mathbf{p})\dot{y}^2,$$

Of course you could denote the various components of the increment by h_1, h_2, h_3, or by u, v, w. But we recommend against writing x, y, z for those increments, which leads to confusion.

If you didn't use this argument, you would discover that

$$\alpha_1 = \alpha_2 = 0$$

anyway, on coming to equation 3.

where $\dot{x}$ denotes the increment in the x direction and $\dot{y}$ the increment in the y direction. To simplify notation, denote the various coefficients by α_1, α_2, etc.:

$$P^2_{g,\mathbf{p}}\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \end{pmatrix} = 1 + \alpha_1\dot{x} + \alpha_2\dot{y} + \alpha_3\dot{x}^2 + \alpha_4\dot{x}\dot{y} + \alpha_5\dot{y}^2, \qquad (1)$$

and note further that equation 2.10.29 tells us that $D_1 g = D_2 g = 0$:

$$[\mathbf{D}g(\mathbf{p})] = -[D_3 f(\mathbf{a})]^{-1}[D_1 f(\mathbf{a}), D_2 f(\mathbf{a})] = -[-1]^{-1}[0\ 0] = [0\ 0].$$

Thus equation 1 becomes

$$1 + \dot{z} = P^2_{g,\mathbf{p}}\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \end{pmatrix} = \underbrace{1}_{g(\mathbf{p})} + \underbrace{\alpha_3\dot{x}^2 + \alpha_4\dot{x}\dot{y} + \alpha_5\dot{y}^2}_{\dot{z}}. \qquad (2)$$

Compute

$$\begin{aligned} f\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \\ 1 + \dot{z} \end{pmatrix} &= -\underbrace{(1+\dot{z})}_{z\text{-coordinate}} + \sin\underbrace{\left(\left(\frac{\pi}{2} + \dot{x}\right)(1+\dot{y})(1+\dot{z})\right)}_{\text{product of } x,y,z \text{ coordinates}} \\ &= -1 - \dot{z} + \sin\left(\frac{\pi}{2} + \left(\dot{x} + \frac{\pi}{2}\dot{y} + \frac{\pi}{2}\dot{z} + \dot{x}\dot{y} + \dot{x}\dot{z} + \frac{\pi}{2}\dot{y}\dot{z} + \dot{x}\dot{y}\dot{z}\right)\right) \\ &= -1 - \dot{z} + \cos\left(\dot{x} + \frac{\pi}{2}\dot{y} + \frac{\pi}{2}\dot{z} + \dot{x}\dot{y} + \dot{x}\dot{z} + \frac{\pi}{2}\dot{y}\dot{z} + \dot{x}\dot{y}\dot{z}\right) \\ &= -1 - \dot{z} + \underbrace{1 - \frac{1}{2}\left(\dot{x} + \frac{\pi}{2}\dot{y} + \frac{\pi}{2}\dot{z} + \dot{x}\dot{y} + \dot{x}\dot{z} + \frac{\pi}{2}\dot{y}\dot{z} + \dot{x}\dot{y}\dot{z}\right)^2 + \cdots}_{\text{Taylor polynomial of cosine}}, \end{aligned} \qquad (3)$$

which gives

$$P^2_{f,\mathbf{a}}\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \\ 1 + \dot{z} \end{pmatrix} = -\dot{z} - \frac{1}{2}\dot{x}^2 - \frac{\pi^2}{8}\dot{y}^2 - \frac{\pi^2}{8}\dot{z}^2 - \frac{\pi}{2}\dot{x}\dot{y} - \frac{\pi}{2}\dot{x}\dot{z} - \frac{\pi^2}{4}\dot{y}\dot{z}. \qquad (4)$$

Substituting $\dot{z} = \alpha_3\dot{x}^2 + \alpha_4\dot{x}\dot{y} + \alpha_5\dot{y}^2$ from equation 2 into equation 4 gives

$$P^2_{f,\mathbf{a}}\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \\ P^2_{g,\mathbf{p}} \end{pmatrix} = -(\alpha_3\dot{x}^2 + \alpha_4\dot{x}\dot{y} + \alpha_5\dot{y}^2) - \frac{1}{2}\dot{x}^2 - \frac{\pi^2}{8}\dot{y}^2 - \frac{\pi}{2}\dot{x}\dot{y}.$$

(Note the handy fact that you don't need to compute $\frac{\pi^2}{8}\dot{z}^2$, $\frac{\pi}{2}\dot{x}\dot{z}$, or $\frac{\pi^2}{4}\dot{y}\dot{z}$, since all terms are higher than quadratic.)

Since $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$ for any values of x, y, z, then $P^2_{f,\mathbf{a}}\begin{pmatrix} \pi/2 + \dot{x} \\ 1 + \dot{y} \\ P^2_{g,\mathbf{p}} \end{pmatrix} = 0$; setting all coefficients equal to zero gives

$$\alpha_1 = \alpha_2 = 0, \quad \alpha_3 = -\frac{1}{2}, \quad \alpha_4 = -\frac{\pi}{2} \quad \alpha_5 = -\frac{\pi^2}{8},$$

so

$$P^2_{g,\mathbf{p}}(\mathbf{p} + \dot{\mathbf{x}}) = 1 - \frac{1}{2}\dot{x}^2 - \frac{\pi}{2}\dot{x}\dot{y} - \frac{\pi^2}{8}\dot{y}^2,$$

which we could also write as

$$P^2_{g,\mathbf{p}}(\mathbf{x}) = 1 - \frac{1}{2}\left(x - \frac{\pi}{2}\right)^2 - \frac{\pi}{2}\left(x - \frac{\pi}{2}\right)(y-1) - \frac{\pi^2}{8}(y-1)^2.$$

Solution 3.4.7: We said above that we recommend using some other notation than x, y, z to denote increments to x, y, z.

The real danger is in confusing z and the increment to z. If you do that you may be tempted to substitute $z = 1 + \dot{z}$ for $\dot{z}$ in equation 4, counting the 1 twice. In Example 3.4.9 we do insert the entire Taylor polynomial (degree 2) of g into $x^2 + y^3 + xyz^3 - 3 = 0$, but note that we did not count the 1 twice; x^2 became $(1+u)^2$ and y^3 became $(1+v)^3$, but z^3 became $(1 + a_{1,0}u + a_{0,1}v + \frac{a_{2,0}}{2}u^2 + \cdots)^3$, not $(1 + 1 + a_{1,0}u + \cdots)^3$. As long as you can keep variables and increments straight, it does not matter which you do.

3.4.9 Compute

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \overbrace{\sum_0^\infty \frac{(-t^2)^k}{k!}}^{\text{Taylor series of } e^{-t^2}} dt = \frac{2}{\sqrt{\pi}}\sum_0^\infty \frac{(-1)^k}{k!}\int_0^x t^{2k}\, dt \tag{1}$$
$$= \frac{2}{\sqrt{\pi}}\sum_{k=0}^\infty \frac{(-1)^k (x)^{2k+1}}{(2k+1)k!}.$$

Set $P_n(x) = \frac{2}{\sqrt{\pi}}\sum_{k=0}^{n-1} \frac{(-1)^k x^{2k+1}}{(2k+1)k!}$ and

$$E_n = \operatorname{erf}(1/2) - P_n(1/2) = \overbrace{\frac{2}{\sqrt{\pi}}\sum_{k=n}^\infty \frac{(-1)^k (1/2)^{2k+1}}{(2k+1)k!}}^{\text{alternating series because of } (-1)^k}.$$

Since E_n is an alternating series with decreasing terms tending to 0, it converges, and for each n,

$$\left|E_n\right| \le \left|\text{first term of the series } E_n\right| = \frac{2}{\sqrt{\pi}}\frac{(1/2)^{2n+1}}{(2n+1)n!}.$$

Trying different values of n, we find that E_2 may be too big but E_3 is definitely small enough:

$$|E_2| \le \frac{1}{160\sqrt{\pi}} > 10^{-3} \quad \text{and} \quad |E_3| = \frac{1}{2688\sqrt{\pi}} < 10^{-3}.$$

So take $n = 3$; then

$$P_3(x) = \frac{2}{\sqrt{\pi}}\left(x - \frac{1}{3}x^3 + \frac{1}{10}x^5\right).$$

Solution 3.4.9: The function "erf" computes the area under the bell or "Gaussian" curve and is thus important in probability. The exchange of sum and integral in equation 1 is justified because the series $\frac{(-t^2)^k}{k!}$ is absolutely uniformly convergent on $[0, x]$.

3.4.11 a. If we systematically ignore all terms of degree > 3, we can write

$$\left(1+\frac{x+y}{1+xz}\right)^{1/2} \approx (1+(x+y)(1-xz))^{1/2} = \left(1+x+y-x^2z-xyz\right)^{1/2}$$
$$\approx 1+\frac{1}{2}\left(x+y-x^2z-xyz\right)-\frac{1}{8}\left(x+y-x^2z-xyz\right)^2$$
$$+\frac{1}{6}\cdot\frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)(x+y)^3$$
$$\approx 1+\frac{1}{2}(x+y)-\frac{1}{8}(x+y)^2-\frac{1}{2}x^2z-\frac{1}{2}xyz$$
$$+\frac{1}{16}(x^3+y^3+3x^2y+3xy^2).$$

Solution 3.4.11: First we use equation 3.4.9 with $m=-1$ to write

$$\frac{1}{1+zy}\approx 1-xz.$$

Then we use equation 3.4.9 with $m=1/2$ to compute the Taylor polynomial of

$$(1+(1-xz)(x+y))^{1/2}.$$

b. The number $D_{(1,1,1)}f\begin{pmatrix}0\\0\\0\end{pmatrix}$ is exactly the coefficient of xyz, so it is $-1/2$.

3.5.1 We find signature $(2,1)$, since

$$-4z^2+2yz-4xz+2xy+x^2=-4\left(z-\frac{y}{4}+\frac{x}{2}\right)^2+\frac{1}{4}(y+2x)^2+x^2.$$

3.5.3 a. $x^2+xy-3y^2+\frac{y^2}{4}-\frac{y^2}{4}=\left(x+\frac{y}{2}\right)^2-\left(\frac{y\sqrt{13}}{2}\right)^2$; signature $(1,1)$.

b. $x^2+2xy-y^2+y^2-y^2=(x+y)^2-(\sqrt{2}y)^2$; signature $(1,1)$.

c.

$$x^2+xy+zy = x^2+xy+\frac{y^2}{4}-\frac{y^2}{4}+zy$$
$$=\left(x+\frac{y}{2}\right)^2-\left(\frac{y^2}{4}-yz+z^2\right)+z^2$$
$$=\left(x+\frac{y}{2}\right)^2-\left(\frac{y}{2}-z\right)^2+z^2;\quad \text{signature } (2,1).$$

3.5.5 a. The signature is $(1,1)$, since

$$x^2+xy=\left(x^2+xy+\frac{y^2}{4}\right)-\frac{y^2}{4}=\left(x+\frac{y}{2}\right)^2-\left(\frac{y}{2}\right)^2.$$

b. Let us introduce $u=x+y$, so that $x=u-y$. In these variables, our quadratic form is

$$(u-y)y+yz=-\left(y^2-uy-yz+\frac{uz}{2}+\frac{u^2}{4}+\frac{z^2}{4}\right)+\left(\frac{uz}{2}+\frac{u^2}{4}+\frac{z^2}{4}\right)$$
$$=-\left(y-\frac{u}{2}-\frac{z}{2}\right)^2+\left(\frac{u}{2}+\frac{z}{2}\right)^2.$$

Again the quadratic form has signature $(1,1)$, but this time it is degenerate.

Remark. Nathaniel Schenker points out that in part b, the algebra is simpler using a different change of variables: since $xy + yz = (x+z)y$, we can let $u = (x + z - y)$, so that $x + z = u + y$. Then the quadratic form is

$$(u+y)y = \left(y^2 + uy + \frac{u^2}{4}\right) - \frac{u^2}{4} = \left(y + \frac{u}{2}\right)^2 - \left(\frac{u}{2}\right)^2.$$

Often a clever choice of variables can simplify computations. The advantage of the first solution is that it is a systematic approach that always works: when a quadratic form contains no squares, choose the first term and set the new variable u to be the sum of the variables involved (or, as we did in Example 3.5.7, the difference). △

3.5.7 First, assume Q is a positive definite quadratic form on $\mathbb{R}^n$ with signature (k, l):

$$Q(\mathbf{x}) = \big(\alpha_1(\mathbf{x})\big)^2 + \cdots + \big(\alpha_k(\mathbf{x})\big)^2 - \big(\alpha_{k+1}(\mathbf{x})\big)^2 - \cdots - \big(\alpha_{k+l}(\mathbf{x})\big)^2$$

with the α_i linearly independent.

Each linear function

$$\alpha_i : \mathbb{R}^n \to \mathbb{R}$$

is a $1 \times n$ row matrix, the ith row of the matrix T.

We want to show that $k = n$ and $l = 0$. Assume $k < n$. Then, by the dimension formula, the $k \times n$ matrix $T = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{bmatrix}$ satisfies $\ker T \neq \{\mathbf{0}\}$. Let $\vec{\mathbf{y}} \neq \mathbf{0}$ be in $\ker T$. Then $Q(\vec{\mathbf{y}}) \leq 0$, so Q is not positive definite. This shows that Q positive definite implies $k \geq n$. Since $k + l \leq n$ with $k \geq 0$, $l \geq 0$, this shows that $k = n$ and $l = 0$.

In the other direction, if Q has signature $(n, 0)$, then the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ given by the $n \times n$ matrix $T(\vec{\mathbf{x}}) = \begin{bmatrix} \alpha_1(\vec{\mathbf{x}}) \\ \vdots \\ \alpha_n(\vec{\mathbf{x}}) \end{bmatrix}$ is onto $\mathbb{R}^n$, so by the dimension formula, $\ker T = 0$. So for all $\vec{\mathbf{x}} \neq \mathbf{0}$ in $\mathbb{R}^n$, at least one of the $\alpha_i(\vec{\mathbf{x}}) \neq 0$, so squaring the functions ensures that

$$Q(\vec{\mathbf{x}}) = \big(\alpha_1(\vec{\mathbf{x}})\big)^2 + \cdots + \big(\alpha_n(\vec{\mathbf{x}})\big)^2 > 0.$$

3.5.9 As in Example 3.5.7, we use the substitution $u = a - d$ in the following computation:

This result has important ramifications for physics: according to the theory of relativity, spacetime naturally carries a quadratic form of signature $(1, 3)$. Thus this space of Hermitian matrices is a natural model for spacetime.

$$\begin{aligned}\det H &= ad - b^2 - c^2 \\ &= (u+d)d - b^2 - c^2 = d^2 + ud + \left(\frac{u}{2}\right)^2 - \left(\frac{u}{2}\right)^2 - b^2 - c^2 \\ &= \left(d + \frac{u}{2}\right)^2 - \left(\frac{u}{2}\right)^2 - b^2 - c^2 \\ &= \left(\frac{a+d}{2}\right)^2 - \left(\frac{a-d}{2}\right)^2 - b^2 - c^2.\end{aligned}$$

The signature of $\det H$ is thus $(1, 3)$.

3.5.11 a. Set $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$\operatorname{tr}\left(A^2\right) = a^2 + d^2 + 2bc \quad \text{and} \quad \operatorname{tr}\left(A^\top A\right) = a^2 + b^2 + c^2 + d^2.$$

b. The signature of $\operatorname{tr}(A^\top A)$ is $(4,0)$. The trace of A^2 can be written $a^2 + d^2 + \frac{1}{2}(b+c)^2 - \frac{1}{2}(b-c)^2$, so $\operatorname{tr}(A^2)$ has signature $(3,1)$.

3.5.13 a. First, suppose A is a symmetric matrix. Then the function $\vec{\mathbf{v}}, \vec{\mathbf{w}} \mapsto \vec{\mathbf{v}}^\top A\vec{\mathbf{w}}$ satisfies rule 1:

$$(a_1\vec{\mathbf{v}}_1 + a_2\vec{\mathbf{v}}_2)^\top A\vec{\mathbf{w}} = a_1\vec{\mathbf{v}}_1^\top A\vec{\mathbf{w}} + a_2\vec{\mathbf{v}}_2^\top A\vec{\mathbf{w}}.$$

Since A is symmetric, and the transpose of a number equals the number, it satisfies rule 2:

$$\underbrace{\vec{\mathbf{v}}^\top A\vec{\mathbf{w}}}_{\text{a number}} = (\vec{\mathbf{v}}^\top A\vec{\mathbf{w}})^\top = \vec{\mathbf{w}}^\top(\vec{\mathbf{v}}^\top A)^\top = \vec{\mathbf{w}}^\top A^\top\vec{\mathbf{v}} = \vec{\mathbf{w}}^\top A\vec{\mathbf{v}}.$$

In the other direction, given a symmetric bilinear function B, define the matrix A by

$$a_{i,j} = B(\vec{\mathbf{e}}_i, \vec{\mathbf{e}}_j).$$

and define B_A by $B_A(\vec{\mathbf{v}}, \vec{\mathbf{w}}) = \vec{\mathbf{v}}^\top A\vec{\mathbf{w}}$. The matrix A is symmetric since B is symmetric, and

$$B_A(\vec{\mathbf{e}}_i, \vec{\mathbf{e}}_j) = \vec{\mathbf{e}}_i^\top A\vec{\mathbf{e}}_j = a_{i,j} = B(\vec{\mathbf{e}}_i, \vec{\mathbf{e}}_j),$$

so the desired relation $B = B_A$ holds if applied to the standard basis vectors. But then it holds for the linear combinations of standard basis vectors, i.e., for all pairs of vectors.

b. Let Q be a quadratic form:

$$Q(\vec{\mathbf{x}}) = \sum_{1\le i\le j\le n} b_{i,j}x_ix_j.$$

Make a C with entries $c_{i,i} = b_{i,i}$ and $c_{i,j} = \frac{1}{2}b_{i,j}$ if $i \ne j$. Then we can take B to be the symmetric bilinear function $(\vec{\mathbf{x}}, \vec{\mathbf{w}}) \mapsto \vec{\mathbf{x}}^\top C\vec{\mathbf{w}}$:

$$\vec{\mathbf{x}}^\top C\vec{\mathbf{x}} = \sum_i x_i \left(\sum_j c_{i,j}x_j\right) = \sum_{i=1}^n\sum_{j=1}^n c_{i,j}x_ix_j = \sum_{1\le i\le j\le n} b_{i,j}x_ix_j = Q(\vec{\mathbf{x}}).$$

c. There isn't very much to show:

$$\int_0^1 (a_1p_1(t) + a_2p_2(t))q(t)\,dt = a_1\int_0^1 p_1(t)q(t)\,dt + a_2\int_0^1 p_2(t)q(t)\,dt,$$

and

$$\int_0^1 p(t)q(t)\,dt = \int_0^1 q(t)p(t)\,dt.$$

d. If $T : V \to W$ is any linear transformation, and B is a symmetric bilinear function on W, it is easy to check that the function $B(T(\vec{\mathbf{v}}_1), T(\vec{\mathbf{v}}_2))$ is always a symmetric bilinear function on V. To find the matrix, compute

$$B\Big(\Phi_p(\vec{\mathbf{e}}_i), \Phi_p(\vec{\mathbf{e}}_j)\Big) = \int_0^1 t^{i+j-2}dt = \frac{1}{i+j-1}.$$

Thus the matrix is

$$\begin{bmatrix} 1 & \frac{1}{2} & \dots & \frac{1}{k+1} \\ \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{k+2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{k+1} & \frac{1}{k+2} & \dots & \frac{1}{2k+1} \end{bmatrix}.$$

Solution 3.5.13, part d: This matrix is called the $(k+1)\times(k+1)$ *Hilbert matrix*; it is a remarkable fact that its inverse has only integer entries. You might confirm this in a couple of cases, but proving it is quite tricky.

3.5.15 a. Both $(p(t))^2$ and $(p'(t))^2$ are polynomials, whose coefficients are homogeneous quadratic polynomials in the coefficients of p. After integrating, the coefficient of t^k in $(p(t))^2 - (p'(t))^2$ simply gets multiplied by $1/(k+1)$.

Explicitly: setting $p = \sum_{i=0}^{k} a_i t^i$ we get

$$\begin{aligned} Q(p) &= \sum_{i=0}^{k}\sum_{j=0}^{k} a_i a_j \int_0^1 \left(t^{i+j} - ijt^{i+j-2}\right) dt \\ &= \sum_{i=0}^{k}\sum_{j=0}^{k} a_i a_j \left(\frac{1}{i+j+1} - \frac{ij}{i+j-1}\right). \end{aligned}$$

b. We find

$$\int_0^1 \Big((at^2+bt+c)^2 - (2at+b)^2\Big)dt = -\frac{17}{15}a^2 - \frac{2}{3}b^2 + c^2 - \frac{3}{2}ab + \frac{2}{3}ac + bc.$$

It seems easiest to remove c, to find

$$Q(p) = \left(c + \frac{b}{2} + \frac{a}{3}\right)^2 - \frac{17}{15}a^2 - \frac{2}{3}b^2 - \frac{3}{2}ab - \frac{1}{9}a^2 - \frac{1}{4}b^2 - \frac{1}{3}ab.$$

The terms other than the first square are

$$-\frac{56}{45}a^2 - \frac{11}{12}b^2 - \frac{11}{6}ab.$$

It seems easier to remove b next, writing

$$-\frac{11}{12}(a+b)^2 + \frac{11}{12}a^2 - \frac{56}{45}a^2.$$

We are left with $\frac{-59}{180}a^2$. So the quadratic form is

$$\left(c+\frac{a}{3}+\frac{b}{2}\right)^2-\frac{11}{12}(a+b)^2-\frac{59}{180}a^2, \quad \text{with signature } (1,2).$$

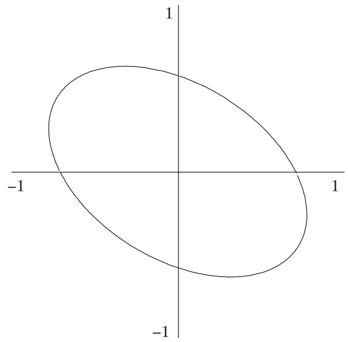


FIGURE 1 FOR SOLUTION 3.5.17, part a. The ellipse of equation
$$2x^2+2xy+3y^2=1.$$

3.5.17 a. The quadratic form corresponding to this matrix is

$$[x\ y]\begin{bmatrix}2&1\\1&3\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=2x^2+2xy+3y^2.$$

Completing squares gives $2x^2+2xy+3y^2=2(x+y/2)^2+5y^2/2$, so the quadratic form has signature $(2,0)$, and the curve of equation

$$2x^2+2xy+3y^2=1$$

is an ellipse, drawn in the margin (Figure 1).

b. The quadratic form corresponding to the matrix $A=\begin{bmatrix}2&1&0\\1&2&1\\0&1&2\end{bmatrix}$ is

$$2(x^2+y^2+z^2+xy+yz)=2\left(\left(x+\frac{y}{2}\right)^2+\left(z+\frac{y}{2}\right)^2+\left(\frac{y}{\sqrt{2}}\right)^2\right),$$

which has signature $(3,0)$. This formula describes an ellipsoid. It is shown in Figure 2.

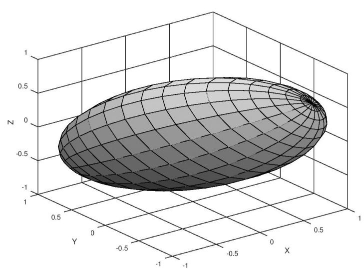


FIGURE 2 for Solution 3.5.17, part b

c. The quadratic form corresponding to the matrix $\begin{bmatrix}2&0&3\\0&0&0\\3&0&-1\end{bmatrix}$ is $2x^2+6xz-z^2=(x+3z)^2-(\sqrt{8}z)^2$. It is degenerate, of signature $(1,1)$, and the surface of equation $2x^2+6xz-z^2=1$ is a hyperbolic cylinder.

As above this can be parametrized by

$$x=\cosh s-\frac{3\sinh s}{\sqrt{8}}$$
$$y=t$$
$$z=\frac{\sinh s}{\sqrt{8}}.$$

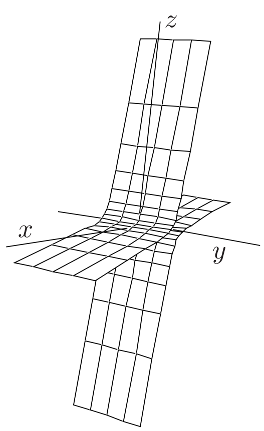


FIGURE 3 FOR SOLUTION 3.5.17, part c. The hyperbolic cylinder.

The resulting surface is shown in Figure 3; its intersection with the (x,z)-plane is shown in Figure 4.

d. The quadratic form corresponding to the matrix $\begin{bmatrix}2&4&-3\\4&1&3\\-3&3&-1\end{bmatrix}$ is $2x^2+y^2-z^2+8xy-6xz+6yz$. Writing this as a sum or difference of squares is a bit more challenging; we find

$$\left(y+4x+\frac{3}{2}y\right)^2-14\left(x-\frac{9}{14}z\right)^2+\frac{71}{28}z^2,$$

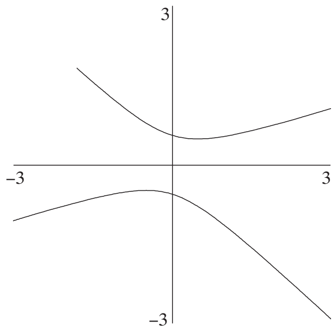


FIGURE 4. FOR SOLUTION 3.5.17, part c. The hyperbola in the (x,z)-plane over which we are constructing the cylinder.

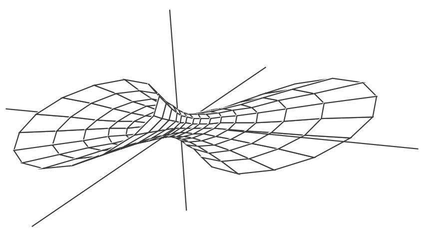


FIGURE 5 FOR SOLUTION 3.5.17, part d. The hyperboloid of one sheet of part d.

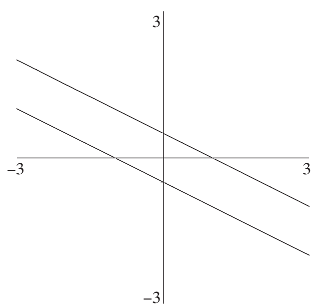


FIGURE 6 FOR SOLUTION 3.5.17, (e) The two lines $x+2y=1$ and $x+2y=-1$ of part e.

so the quadratic form is of signature $(2,1)$. Again the surface of equation $2x^2+y^2-z^2+8xy-6xz+6yz=1$ is a hyperboloid of one sheet, shown in Figure 5. (Parametrizing this hyperboloid was done as in part b, but the computations are quite a bit more unpleasant.)

e. The quadratic form corresponding to the matrix $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ is

$$x^2+4xy+4y^2=(x+2y)^2.$$

It is degenerate, of signature $(1,0)$, and the curve of equation $(x+2y)^2=1$ is the union of the two lines $x+2y=1$ and $x+2y=-1$; see Figure 6.

3.6.1 a. Since all the partials of f are continuous, f is differentiable on all of $\mathbb{R}^3$. In fact,

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right]=[\,2x+y \quad x+\sin y \quad 2z\,], \quad \text{which vanishes at the origin,}$$

and is thus a critical point by Definition 3.6.2.

b. We have

$$P^2_{f,\mathbf{0}}(\mathbf{h})=-1+h_1^2+h_1h_2+\frac{h_2^2}{2}+h_3^2$$

and

$$h_1^2+h_1h_2+\frac{h_2^2}{2}+h_3^2=\left(h_1+\frac{h_2}{2}\right)^2+\frac{h_2^2}{4}+h_3^2,$$

so the critical point has signature $(3,0)$ and is a local minimum by Theorem 3.6.8.

3.6.3 By Proposition 3.5.11 there exists a subspace W of $\mathbb{R}^n$ such that Q is negative definite on W. If a quadratic form on W is negative definite, then by Proposition 3.5.15, there exists a constant $C>0$ such that

$$Q(\vec{\mathbf{x}})\le -C|\vec{\mathbf{x}}|^2 \quad \text{for all } \vec{\mathbf{x}}\in\mathbb{R}^n.$$

Thus if $\vec{\mathbf{h}}\in W$, and $t>0$, there exists $C>0$ such that

$$\frac{f(\mathbf{a}+t\vec{\mathbf{h}})-f(\mathbf{a})}{t^2}=\frac{t^2Q(\vec{\mathbf{h}})+r(t\vec{\mathbf{h}})}{t^2}\le -C|\vec{\mathbf{h}}|^2+\frac{r(t\vec{\mathbf{h}})}{t^2},$$

where $t^2Q(\vec{\mathbf{h}})$ is the quadratic term of the Taylor polynomial, and $r(\vec{\mathbf{h}})$ is the remainder. Since

$$\lim_{t\to 0}\frac{r(t\vec{\mathbf{h}})}{t^2}=0,$$

it follows that $f(\mathbf{a}+t\vec{\mathbf{h}})<f(\mathbf{a})$ for $|t|>0$ sufficiently small.

Solution 3.6.3: For $|t|>0$ sufficiently small,

$$\left|\frac{r(t\vec{\mathbf{h}})}{t^2}\right|<C|\vec{\mathbf{h}}|^2.$$

so we have

$$\frac{f(\mathbf{a}+t\vec{\mathbf{h}})-f(\mathbf{a})}{t^2}<0.$$

3.6.5 a. The critical points are the points where $D_1 f$, $D_2 f$, and $D_3 f$ vanish, i.e., the points satisfying the three equations

$$\begin{aligned} y - z + yz &= 0 \\ x + z + xz &= 0 \\ y - x + xy &= 0. \end{aligned}$$

If we use the first two to express x and y in terms of z and substitute into the third equation, we get

$$\frac{z}{z+1}\left(2 - \frac{z}{z+1}\right) = 0.$$

This leads to the two roots $z = 0$ and $z = -2$, and finally to the critical points $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}$.

A polynomial function like f is its own Taylor polynomial at the origin (but not elsewhere).

b. The quadratic terms at the first point (the origin) are $uv + vw - uw$; see the margin note. Set $u = s - v$, and get rid of u, to find

$$sv - v^2 + vw - sw + vw = -\left(v - w - \frac{s}{2}\right)^2 + w^2 + \frac{1}{4}s^2.$$

The origin is a saddle, with signature $(2, 1)$.

Here we turn the polynomial into a new polynomial, whose variables are the increments from the point $\begin{pmatrix} -2 \\ 2 \\ -2 \end{pmatrix}$.

At the other critical point, set

$$x = -2 + u, \quad y = 2 + v, \quad z = -2 + w$$

and substitute into the function. We find

$$-4 - uv - vw + uw + uvw.$$

The quadratic terms are $-uv - vw + uw$, exactly the opposites of the quadratic terms at the first point, so again we have a saddle, this time with signature $(1, 2)$.

3.6.7 a. We compute

Solution 3.6.7: Remember that

$$(fg)' = gf' + fg'$$
$$(e^f)' = e^f f'.$$

In part b, we use equation 3.4.5 for the Taylor polynomial of e^x.

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \end{pmatrix}\right] = \left[2x(1 + x^2 + y^2)e^{x^2 - y^2}\ ,\ 2y\big(1 - (x^2 + y^2)\big)e^{x^2 - y^2}\right].$$

Thus $D_1 f = 0$ when $x = 0$, and $D_2 f = 0$ when $y = 0$ or $x^2 + y^2 = 1$. Thus the critical points occur when $x = 0$ and either $y = 0$ or $x^2 + y^2 = 1$, that is, at the points $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$.

b. The second-degree terms of the Taylor approximation of f about the origin are x^2 and y^2. (Just multiply $x^2 + y^2$ and $e^u = 1 + u + u^2/2 + \cdots$, where $u = x^2 - y^2$, and take the second-degree terms.) Since $x^2 + y^2$ is a positive definite quadratic form, f has a local minimum at the origin. In order to characterize the other two critical points, we first compute the three second partial derivatives:

$$D_1^2 f\begin{pmatrix} x \\ y \end{pmatrix} = \left(2 + 10x^2 + 2y^2 + 4x^2y^2 + 4x^4\right) e^{(x^2-y^2)}$$

$$D_1 D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = -4xy(x^2+y^2)e^{(x^2-y^2)}$$

$$D_2^2 f\begin{pmatrix} x \\ y \end{pmatrix} = \left(2 - 2x^2 - 10y^2 + 4x^2y^2 + 4y^4\right)e^{(x^2-y^2)}.$$

At both $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$, we have $D_1^2 = 4/e$, $D_1D_2f = 0$, $D_2^2f = -4/e$. Thus the second-degree terms of the Taylor polynomial are

$$\frac{D_1^2 f h_1^2}{2!} = \frac{2h_1^2}{e} \quad \text{and} \quad \frac{D_2^2 f h_2^2}{2!} = \frac{-2h_2^2}{e},$$

which give the quadratic form

$$\frac{2h_1^2}{e} - \frac{2h_2^2}{e}$$

at both points. This form has signature $(1,1)$, so f has a saddle at both of these critical points.

3.7.1 Let us call the three sides of the box x, $2x$, and y. Then the problem is to maximize the volume $2x^2y$, subject to the constraint that the surface area is $2(2x^2 + xy + 2xy) = 10$. This leads to the Lagrange multiplier problem

$$[\,4xy \quad 2x^2\,] = \lambda\,[\,4x + 3y \quad 3x\,].$$

In the resulting two equations

$$4xy = \lambda(4x + 3y) \tag{1}$$

$$2x^2 = \lambda(3x), \tag{2}$$

we see that one solution to equation 2 is $x = 0$, which is certainly not the maximum. The other solution is $2x = 3\lambda$. Substitute this value of λ into equation 1, to find $4x = 3y$, and then the corresponding value of y in terms of x into the constraint equation, to find $x = \sqrt{\frac{5}{6}}$. This leads to the maximum volume

$$V = \frac{20}{9}\sqrt{\frac{5}{6}}.$$

3.7.3 a. Suppose we use Lagrange multipliers, computing the derivative of $\varphi = x^3 + y^3 + z^3$ and the derivatives of the constraint functions $F_1 = x + y + z = 2$ and $F_2 = x + y - z = 3$. Then at a critical point,

$$D\varphi = [3x^2, 3y^2, 3z^2] = \lambda_1[1,1,1] + \lambda_2[1,1,-1];$$

i.e.,

$$x = \pm y = \pm\sqrt{\frac{\lambda_1 + \lambda_2}{3}} \quad \text{and} \quad z = \pm\sqrt{\frac{\lambda_1 - \lambda_2}{3}}.$$

Adding the constraints gives $2(x+y)=5$, hence we must use the positive square root for both x and y. Similarly, subtracting the constraints gives $-2z=1$, so for z we must use the negative square root:

Solution 3.7.3, part a: Here it would be easier to find the critical points using a parametrization. In this case the constraints are linear, so finding a parametrization is easy. But usually finding a parametrization requires the implicit function and is computationally daunting, and less straightforward than using Lagrange multipliers (also computationally daunting).

$$x = y = \sqrt{\frac{\lambda_1+\lambda_2}{3}} \quad \text{and} \quad z = -\sqrt{\frac{\lambda_1-\lambda_2}{3}}.$$

Substituting these values in the constraint functions tells us that

$$\lambda_1 = \frac{87}{32} \quad \text{and} \quad \lambda_2 = \frac{63}{32},$$

which we can use to determine that there is a critical point at $\begin{pmatrix} 5/4 \\ 5/4 \\ -1/2 \end{pmatrix}$.

b. We can't immediately conclude that this critical point is a minimum or a maximum of the constrained function because the domain of that function is not compact. We will analyze the critical point two ways: using a parametrization that incorporates the constraints, and using the augmented Hessian matrix. In this case, a parametrization is computationally much easier.

Solution 3.7.3, part b: If you can find critical points, analyzing them using the Hessian matrix approach is systematic.

First let us use a parametrization. If we write the constraint functions in matrix form and row reduce, we get

$$\begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 1 & -1 & 3 \end{bmatrix}, \quad \text{which row reduces to} \quad \begin{bmatrix} 1 & 1 & 0 & 5/2 \\ 0 & 0 & 1 & -1/2 \end{bmatrix};$$

i.e.,

$$x+y = 5/2 \quad \text{and} \quad z = -1/2.$$

Substituting $x = 5/2 - y$ and $z = -1/2$ in the function $x^3+y^3+z^3$ gives

$$(5/2-y)^3 + y^3 - 1/8 = (5/2)^3 - 3(5/2)^2 y + 3(5/2)y^2 - 1/8,$$

which is the equation for a parabola; the fact that the coefficient for y^2 is positive tells us that the parabola is shaped like a cup, and therefore has a minimum.

Now we will use the augmented Hessian matrix. Since the constraint functions F_1 and F_2 are linear, their second derivatives are 0, so the only nonzero entries of the matrix B of equation 3.7.42 come from the constrained function f, which has second derivatives $6x$, $6y$, and $6z$:

The augmented Hessian matrix H is defined in equation 3.7.43. Since

$$F_1 = x+y+z = 2$$
$$F_2 = x+y-z = 3,$$

$$-[\mathbf{DF}] = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

$$B = \begin{bmatrix} 6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z \end{bmatrix}.$$

Thus the augmented Hessian matrix at the critical point $\begin{pmatrix} 5/4 \\ 5/4 \\ -1/2 \end{pmatrix}$ is

$$H = \begin{bmatrix} 6x\cdot\frac{5}{4} & 0 & 0 & -1 & -1 \\ 0 & 6y\cdot\frac{5}{4} & 0 & -1 & -1 \\ 0 & 0 & -6z\cdot\frac{1}{2} & -1 & 1 \\ -1 & -1 & -1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \end{bmatrix},$$

which (denoting by u and v the two additional variables represented by that matrix) corresponds to the quadratic form

$$\frac{15}{2}x^2 + \frac{15}{2}y^2 - 3z^2 - 2ux - 2xv - 2yu - 2yv - 2zu + 2zv.$$

After some (painful) arithmetic, we express this as a sum of squares:

$$+\left(\sqrt{\frac{15}{2}}\,x - \sqrt{\frac{2}{15}}\,u - \sqrt{\frac{2}{15}}\,v\right)^2 + \left(\sqrt{\frac{15}{2}}\,y - \sqrt{\frac{2}{15}}\,u - \sqrt{\frac{2}{15}}\,v\right)^2$$
$$-\left(\sqrt{3}\,z + \frac{u}{\sqrt{3}} - \frac{v}{\sqrt{3}}\right)^2 + \frac{1}{15}(u-9v)^2 - \frac{80}{15}v^2,$$

which has signature $(3,2)$. Since $m = 2$, Theorem 3.7.13 tells us that $(3,2) = (p+2, q+2)$. So the signature of the constrained critical point is $(1,0)$, which means that the constrained critical point is a minimum.

Solution 3.7.5: We will compute the volume in one octant, then multiply that volume by 8.

3.7.5 The volume of the parallelepiped in the first octant (where $x \geq 0$, $y \geq 0$, $z \geq 0$) is gotten by maximizing xyz subject to the constraint

$$x^2 + 4y^2 + 9z^2 = 9.$$

The Lagrange multiplier theorem says that at such a maximum, there exists a number λ such that

$$[yz,\ xz,\ xy] = \lambda[2x,\ 8y,\ 18z],$$

i.e., we have the three equations $yz = 2\lambda x,\ xz = 8\lambda y,\ xy = 18\lambda z$. Multiply the first by x, the second by y, and the third by z, to find

$$xyz = 2\lambda x^2 = 8\lambda y^2 = 18\lambda z^2.$$

This leads to $x^2 = 4y^2 = 9z^2$, hence $3x^2 = 9$, or

$$x = \sqrt{3},\ y = \frac{\sqrt{3}}{2},\ z = \frac{\sqrt{3}}{3},$$

and finally the maximum of xyz is $\sqrt{3}/2$. There are eight octants, so the total volume is $4\sqrt{3}$.

3.7.7 a. Using the hint, write $z = \frac{1}{c}(1 - ax - by)$. The function now becomes

$$xyz = \frac{1}{c}xy(1 - ax - by) = F\begin{pmatrix} x \\ y \end{pmatrix},$$

with partial derivatives

$$D_1F\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{c}(y - 2axy - by^2) = \frac{y}{c}(1 - 2ax - by),$$
$$D_2F\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{c}(x - ax^2 - 2bxy) = \frac{x}{c}(1 - ax - 2by).$$

There are clearly four critical points:

$$\begin{pmatrix} 0 \\ 0 \\ 1/c \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1/b \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1/a \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1/(3a) \\ 1/(3b) \\ 1/(3c) \end{pmatrix}.$$

b. At these four points the Hessian matrices (matrices whose entries are second derivatives) are

$$\frac{1}{c}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \quad \frac{1}{c}\begin{bmatrix}-2a/b & -1\\ -1 & 0\end{bmatrix}, \quad \frac{1}{c}\begin{bmatrix}0 & -1\\ -1 & -2b/a\end{bmatrix}, \quad \frac{1}{c}\begin{bmatrix}-(2a)/(3b) & -1/3\\ -1/3 & -(2b)/(3a)\end{bmatrix}.$$

By Proposition 3.5.17, each of these symmetric matrices uniquely determines a quadratic form; by Theorem 3.6.8, these quadratic forms can be used to analyze the critical points.

The quadratic form corresponding to the first matrix is

$$\frac{2}{c}xy = \frac{1}{2c}\Big((x+y)^2-(x-y)^2\Big),$$

with signature $(1,1)$, so the first point is a saddle point. Since nothing made the z-coordinate special, the second and third points are also saddles. (This could also be computed directly.)

The fourth corresponds to the quadratic form

$$-\frac{2}{3c}\left(\frac{a}{b}x^2+\frac{b}{a}y^2+xy\right) = -\frac{2}{3c}\left(\left(\sqrt{\frac{a}{b}}x+\frac{1}{2}\sqrt{\frac{b}{a}}y\right)^2+\frac{3b}{4a}y^2\right).$$

Thus the quadratic form is negative definite, and the function is a maximum. (You could also do this conceptually: The region

$$ax+by+cz=1,\ x,y,z\ge 0$$

is compact, and the function is nonnegative there. Thus it must have a positive maximum; since the function is equal to 0 at the other critical points, this maximum must be the fourth critical point.)

3.7.9 If A is any square matrix, then

$$\begin{aligned}Q_A(\vec{\mathbf{a}}+\vec{\mathbf{h}})-Q_A(\vec{\mathbf{a}})-\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}-\vec{\mathbf{a}}^\top A\vec{\mathbf{h}} &= (\vec{\mathbf{a}}+\vec{\mathbf{h}})^\top A(\vec{\mathbf{a}}+\vec{\mathbf{h}})-\vec{\mathbf{a}}^\top A\vec{\mathbf{a}}-\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}-\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}\\ &= \vec{\mathbf{a}}^\top A\vec{\mathbf{a}}+\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}+\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}+\vec{\mathbf{h}}^\top A\vec{\mathbf{h}}-\vec{\mathbf{a}}^\top A\vec{\mathbf{a}}-\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}-\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}\\ &= \vec{\mathbf{h}}^\top A\vec{\mathbf{h}}.\end{aligned}$$

So

$$0\le \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}}\frac{\left|Q_A(\vec{\mathbf{a}}+\vec{\mathbf{h}})-Q_A(\vec{\mathbf{a}})-\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}-\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}\right|}{|\vec{\mathbf{h}}|} = \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}}\frac{|\vec{\mathbf{h}}^\top A\vec{\mathbf{h}}|}{|\vec{\mathbf{h}}|}\le \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}}\frac{|\vec{\mathbf{h}}|^2 A}{|\vec{\mathbf{h}}|}=0$$

Thus the derivative of Q_A at $\vec{\mathbf{a}}$ is the linear function $\vec{\mathbf{h}}\mapsto\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}+\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}$. In our case A is symmetric, so $\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}=\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}$, justifying equation 3.7.52.

We can get the same result using the definition of the directional derivative. The directional derivative of Q_A is

$$\begin{aligned}\lim_{k\to 0}\frac{1}{k}\Big(Q_A(\vec{\mathbf{a}}+k\vec{\mathbf{h}})-Q_A(\vec{\mathbf{a}})\Big) &= \lim_{k\to 0}\frac{1}{k}\Big((\vec{\mathbf{a}}+k\vec{\mathbf{h}})\cdot\big(A(\vec{\mathbf{a}}+k\vec{\mathbf{h}})\big)-\vec{\mathbf{a}}\cdot\big(A\vec{\mathbf{a}}\big)\Big)\\ &= \lim_{k\to 0}\Big(\vec{\mathbf{a}}\cdot\big(A\vec{\mathbf{h}}\big)+\vec{\mathbf{h}}\cdot\big(A\vec{\mathbf{a}}\big)+k\vec{\mathbf{h}}\cdot\big(A\vec{\mathbf{h}}\big)\Big) \qquad (1)\\ &= \vec{\mathbf{a}}\cdot(A\vec{\mathbf{h}})\ +\ \vec{\mathbf{h}}\cdot\big(A\vec{\mathbf{a}}\big)\\ &= \vec{\mathbf{a}}^\top A\vec{\mathbf{h}}+\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}.\end{aligned}$$

This linear function of $\vec{\mathbf{h}}$ depends continuously on $\mathbf{a}$, so the function Q_A is differentiable, and its derivative is

$$[\mathbf{D}Q_A(\vec{\mathbf{a}})]\vec{\mathbf{h}} = \vec{\mathbf{a}}^\top A\vec{\mathbf{h}} + \vec{\mathbf{h}}^\top A\vec{\mathbf{a}}.$$

Using the properties of transposes on products and noting that A is symmetric, we see that

$$(\vec{\mathbf{h}}^\top A\vec{\mathbf{a}})^\top = \vec{\mathbf{a}}^\top A^\top \vec{\mathbf{h}} = \vec{\mathbf{a}}^\top A\vec{\mathbf{h}}.$$

Since $\vec{\mathbf{h}}^\top A\vec{\mathbf{a}}$ is a number, it is symmetric, so we also have

$$\vec{\mathbf{h}}^\top A\vec{\mathbf{a}} = \vec{\mathbf{a}}^\top A\vec{\mathbf{h}}.$$

Substituting this into the last line of equation 1 yields

$$\lim_{k\to 0} \frac{1}{k}\Big(Q_A(\vec{\mathbf{a}} + k\vec{\mathbf{h}}) - Q_A(\vec{\mathbf{a}})\Big) = 2\vec{\mathbf{a}}^\top A\vec{\mathbf{h}}.$$

3.7.11 The curve C is a hyperbola. Indeed, take a point of the plane curve C' of equation $2xy + 2x + 2y + 1 = 0$, and set $z = 1 + x + y$. Then $z^2 = x^2 + y^2 + 1 + 2xy + 2x + 2y = x^2 + y^2$, so this point is both on the plane and on the cone. The equation for C' can be written $(x+1)(y+1) = 1/2$, so C' is the hyperbola of equation $xy = 1/2$, translated by $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$.

In particular, C stretches out to infinity, and there is no point furthest from the origin. But the hyperbola is a closed subset of $\mathbb{R}^3$, and each branch contains a point closest to the origin, i.e., a local minimum.

We are interested in the extrema of the function $x^2 + y^2 + z^2$ (measuring squared distance from the origin), constrained to the two equations defining C. At a critical point of the constrained function, we must have

$$[2x,\ 2y, 2z] = \lambda_1[2x,\ 2y,\ -2z] + \lambda_2[1,\ 1,\ -1].$$

The first two of these equations, $2x = \lambda_1 2x + \lambda_2$ and $2y = \lambda_1 2y + \lambda_2$, imply that either $\lambda_1 = 1$ and $\lambda_2 = 0$ or that $x = y$.

The hypothesis $\lambda_2 = 0$ gives $z = 0$, hence the first constraint (equation of the cone) becomes $x^2 + y^2 = 0$, i.e., $x = y = 0$. This is incompatible with the second constraint.

So we must have $x = y$. Then the two equations $2x^2 - z^2 = 0$ and $2x - z = -1$ give $2x^2 + 4x + 1 = 0$. Solving this quadratic equation, we find the two roots

$$x_1 = -\frac{2+\sqrt{2}}{2} \quad \text{and} \quad x_2 = \frac{-2+\sqrt{2}}{2}.$$

These give the points

$$\begin{pmatrix} (-2-\sqrt{2})/2 \\ (-2-\sqrt{2})/2 \\ -1-\sqrt{2} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} (-2+\sqrt{2})/2 \\ (-2+\sqrt{2})/2 \\ -1+\sqrt{2} \end{pmatrix}.$$

The distances squared from the origin to these points are $6 + 4\sqrt{2}$ and $6 - 4\sqrt{2}$ respectively. Since both of the local minima of the distance function are critical points, these two points are both local minima.

3.7.13 a. The partial derivatives of f are

$$D_1 f\begin{pmatrix} x\\ y\\ z\end{pmatrix} = y-1, \qquad D_2 f\begin{pmatrix} x\\ y\\ z\end{pmatrix} = x+1, \qquad D_3 f\begin{pmatrix} x\\ y\\ z\end{pmatrix} = 2z.$$

The only point at which all three vanish is $\mathbf{x}_0$, so this is the only critical point of f. Since f is a quadratic polynomial, its Taylor polynomial of degree two at $\mathbf{x}_0$ is itself:

$$\begin{aligned} P^2_{f,\mathbf{x}_0}\begin{pmatrix} -1+h_x\\ 1+h_y\\ 0+h_z\end{pmatrix} &= f\begin{pmatrix} -1+h_x\\ 1+h_y\\ 0+h_z\end{pmatrix} \\ &= (-1+h_x)(1+h_y) - (-1+h_x) + (1+h_y) + (h_z)^2 \\ &= -1 + h_x - h_y + h_x h_y + 1 - h_x + 1 + h_y + h_z^2 \\ &= 1 + h_x h_y + h_z^2. \end{aligned}$$

The quadratic terms yield the quadratic form

$$\frac{1}{4}\Big((h_x+h_y)^2 - (h_x-h_y)^2\Big) + h_z^2,$$

which has signature $(2,1)$. Therefore $\mathbf{x}_0$ is a saddle of f (Definition 3.6.9).

b. The function F is our constraint function. Its derivative is

$$\left[\mathbf{D}F\begin{pmatrix} x\\ y\\ z\end{pmatrix}\right] = [\,2(x+1) \quad 2(y-1) \quad 2z\,].$$

By the Lagrange multiplier theorem, a constrained critical point on $S_{\sqrt{2}}(\mathbf{x}_0)$ occurs at a point where

$$\left[\mathbf{D}f\begin{pmatrix} x\\ y\\ z\end{pmatrix}\right] = \lambda\left[\mathbf{D}F\begin{pmatrix} x\\ y\\ z\end{pmatrix}\right], \qquad \text{i.e.,} \qquad \begin{cases} y-1 = 2\lambda(x+1) \\ x+1 = 2\lambda(y-1) \\ 2z = 2\lambda z \end{cases}.$$

The final equation says either $\lambda = 1$ or $z = 0$.

If $z \neq 0$, solving the other equations yields $x = -1$, $y = 1$. Then the constraint function F shows that $z = \pm\sqrt{2}$. Thus $\begin{pmatrix} -1\\ 1\\ \sqrt{2}\end{pmatrix}, \begin{pmatrix} -1\\ 1\\ -\sqrt{2}\end{pmatrix}$ are constrained critical points of f.

If $z = 0$, then the first two equations imply that $\lambda = \pm 1/2$ or $\lambda = 0$. But if $\lambda = 0$, then $y = 1$ and $x = -1$, yielding the point $\mathbf{x}_0$, which is not on $S_{\sqrt{2}}(\mathbf{x}_0)$. Therefore $\lambda = \pm 1/2$. In the case $\lambda = 1/2$, we get $y = x+2$, while if $\lambda = -1/2$, we have $y = -x$. These, plus the equation of the sphere, lead to $x = 0$ or $x = -2$. Thus

$$\begin{pmatrix} 0\\ 2\\ 0\end{pmatrix}, \quad \begin{pmatrix} -2\\ 0\\ 0\end{pmatrix}, \quad \begin{pmatrix} 0\\ 0\\ 0\end{pmatrix}, \quad \begin{pmatrix} -2\\ 2\\ 0\end{pmatrix}$$

are also constrained critical points of f.

c. We showed in part a that the only critical point in the interior of the ball is at $\mathbf{x}_0$, and that this point is a saddle; therefore it cannot be an extremum. We compute f at the constrained critical points (on the surface of the ball) found in part b:

$$f\begin{pmatrix}-1\\1\\\sqrt{2}\end{pmatrix} = f\begin{pmatrix}-1\\1\\-\sqrt{2}\end{pmatrix} = 3, \qquad f\begin{pmatrix}0\\2\\0\end{pmatrix} = f\begin{pmatrix}-2\\0\\0\end{pmatrix} = 2,$$

$$f\begin{pmatrix}0\\0\\0\end{pmatrix} = 0, \qquad f\begin{pmatrix}-2\\2\\0\end{pmatrix} = 0$$

Therefore the maximum value of f on $\overline{B_{\sqrt{2}}(\mathbf{x}_0)}$ is 3 and the minimum value is 0. Note that at the critical point $\mathbf{x}_0$ we have $f(\mathbf{x}_0) = 1$, which does indeed lie between the maximum and minimum values.

3.7.15 a. The set

$$X_1 = \left\{ \begin{pmatrix}x\\y\end{pmatrix} \in X \mid x^2 + y^2 \le 4 \right\} \subset \mathbb{R}^2$$

is closed and bounded, hence compact, and the function f has a minimum on X_1.

The point $\begin{pmatrix}1\\0\end{pmatrix}$ is on X_1, and $f\begin{pmatrix}1\\0\end{pmatrix} = 1$. Since at every $\begin{pmatrix}x\\y\end{pmatrix} \in X - X_1$ we have $f\begin{pmatrix}x\\y\end{pmatrix} > 4$, the minimum of f on X_1 is a global minimum.

b. The Lagrange multiplier equation is

$$[2x,\ \ xy] = \lambda\big[3(x-1)^2,\ \ -2y\big].$$

The second equation $2y = -2\lambda y$ implies either

- $y = 0$, which by $y^2 = (x-1)^3$ gives $x = 1$, so the first equation gives $2 = 0$, a contradiction,

or

- $\lambda = -1$. Then $2x = -3(x-1)^2$, which gives

$$-\frac{2}{3}x = x^2 - 2x + 1, \quad \text{i.e., } x^2 - \frac{4}{3}x + 1 = 0,$$

which has no real solutions, since then $x = -\frac{2}{3} \pm \sqrt{4/9 - 1}$.

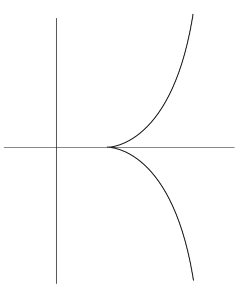


FIGURE FOR SOLUTION 3.7.15 The curve $X \subset \mathbb{R}^2$. It is defined by

$$F\begin{pmatrix}x\\y\end{pmatrix} = (x-1)^3 - y^2 = 0.$$

c. The curve is shown in the figure in the margin. The point $\begin{pmatrix}1\\0\end{pmatrix}$ is the closest point to the origin, hence the minimum of $x^2 + y^2$ on X.

To see this without a drawing, note that $y^2 = (x-1)^3$ implies that $(x-1)^3 \ge 0$, hence $x \ge 1$, so $x^2 + y^2 \ge 1$, with the unique minimum at $\begin{pmatrix}1\\0\end{pmatrix}$.

The Lagrange multiplier theorem (Theorem 3.7.5) requires that $[\mathbf{D}F]$ be surjective. Here,

$$\left[\mathbf{D}F\begin{pmatrix}x\\y\end{pmatrix}\right] = \big[3(x-1)^2,\ \ -2y\big]$$

is not surjective at $\begin{pmatrix}1\\0\end{pmatrix}$.

3.8.1 The outcome

“total 2” has probability 1/100
“total 3” has probability 2/100
“total 4” has probability 3/100
“total 5” has probability 2/20+2/100;
“total 6” has probability 2/20+ 3/100
“total 7” has probability 2/20+4/100
“total 8” has probability 1/4+4/100
“total 9” has probability 2/20+2/100
“total 10” has probability 2/20+1/100
“total 11” has probability 2/100
“total 12” has probability 1/100

FIGURE FOR SOL. 3.8.3, PART C
Daniel Bernoulli (1700–1782)

Consider, Bernoulli wrote, a very poor man waiting for a drawing in which he is equally likely to get nothing or twenty thousand ducats: since the expectation is ten thousand ducats, would he be foolish if he sold his right to the drawing for nine thousand? “I do not think so,” Bernoulli wrote, “although I think that a very rich man would go against his interest if, given the chance of buying this right for that same price, he declined.” Bernoulli's comment justifies insurance: the expected value of insurance is always less than the premiums you pay, but insuring against losses you can't afford still make sense.

3.8.3 a. We have

$$E(f)=\sum_{n=1}^{\infty} f(n)\,\mathbf{P}(\{s\})=\sum_{n=1}^{\infty} n\frac{1}{2^n}=\frac{1}{2}+\frac{2}{2^2}+\frac{3}{2^3}+\cdots=2,$$

since $|x|<1$, we have

$$\sum_{n=1}^{\infty} x^n=\frac{1}{1-x},\quad \text{so differentiating gives}\quad \sum_{n=1}^{\infty} nx^{n-1}=\frac{1}{(1-x)^2},$$

which gives

$$\sum_{n=1}^{\infty} nx^n=\frac{x}{(1-x)^2},\quad \text{hence}\quad \sum_{n=1}^{\infty} n\cdot\frac{1}{2^n}=\frac{1/2}{(1/2)^2}=2.$$

b.

$$E(f)=\sum_{n=1}^{\infty} f(n)\,\mathbf{P}(\{s\})=\sum_{n=1}^{\infty} 2^n\frac{1}{2^n}=\infty$$

c. In theory, you should be willing to ante up any sum less than the expected value, so for f, any sum less than 2. For g, the first thing to find out is how solvent the “bank” is. Obviously no bank can pay out the expected value, so to determine the real expectation you would need to know the maximum the bank will pay. Another thing to consider is how much you are willing to lose; consider the comments of Daniel Bernoulli.

3.8.5 a. The following computation shows that if $\vec{\mathbf{v}}$ is in the kernel of $A^\top A$, it is in the kernel of A:

$$\begin{aligned} 0 &= \vec{\mathbf{v}} \cdot \vec{\mathbf{0}} = \vec{\mathbf{v}} \cdot A^\top A\vec{\mathbf{v}} \\ &= \vec{\mathbf{v}}^\top A^\top A\vec{\mathbf{v}} = (\vec{\mathbf{v}}^\top A^\top)^\top \cdot A\vec{\mathbf{v}} = (A\vec{\mathbf{v}}) \cdot (A\vec{\mathbf{v}}) = |A\vec{\mathbf{v}}|^2. \end{aligned} \tag{1}$$

Similarly, any vector in the kernel of $AA^\top$ is in the kernel of $A^\top$.

Equation 1: It is always true that $A\vec{\mathbf{v}} \cdot \vec{\mathbf{w}} = \vec{\mathbf{v}} \cdot A^\top \vec{\mathbf{w}}$, since

$$\begin{aligned} (A\vec{\mathbf{v}}) \cdot \vec{\mathbf{w}} &= (A\vec{\mathbf{v}})^\top \vec{\mathbf{w}} = \vec{\mathbf{v}}^\top A^\top \vec{\mathbf{w}} \\ &= \vec{\mathbf{v}} \cdot (A^\top \vec{\mathbf{w}}). \end{aligned}$$

b. Since $A^\top A$ and $AA^\top$ are symmetric, by the spectral theorem there are orthonormal eigenbases $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_m$ of $\mathbb{R}^m$ for $A^\top A$ and $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_n$ of $\mathbb{R}^n$ for $AA^\top$. Suppose that

$$A^\top A\vec{\mathbf{v}}_i = \lambda_i \vec{\mathbf{v}}_i \quad \text{and} \quad AA^\top \vec{\mathbf{u}}_j = \mu_j \vec{\mathbf{u}}_j,$$

labeled so that

$$\lambda_1 \ge \dots \ge \lambda_k > \lambda_{k+1} = \dots = \lambda_m = 0 \quad \text{and} \quad \mu_1 \ge \dots \ge \mu_l > \mu_{l+1} = \dots = \mu_n = 0.$$

Let V be the span of $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k$ and let W be the span of $\vec{\mathbf{u}}_1, \dots, \vec{\mathbf{u}}_l$. For each $i \le k$, set $\vec{\mathbf{v}}'_i \stackrel{\text{def}}{=} \frac{1}{\sqrt{\lambda_i}} A\vec{\mathbf{v}}_i$, and for each $j \le l$ set $\vec{\mathbf{u}}'_j \stackrel{\text{def}}{=} \frac{1}{\sqrt{\mu_j}} A^\top \vec{\mathbf{u}}_j$.

Now let us check that the $\vec{\mathbf{v}}'_i$, $i = 1, \dots, k$ are an orthonormal set in W and the $\vec{\mathbf{u}}'_j$ are an orthonormal set in V.

Indeed, the $\vec{\mathbf{v}}'_i$ are orthonormal:

$$\begin{aligned} \vec{\mathbf{v}}'_{i_1} \cdot \vec{\mathbf{v}}'_{i_2} &= \frac{1}{\sqrt{\lambda_{i_1}\lambda_{i_2}}} (A\vec{\mathbf{v}}_{i_1}) \cdot (A\vec{\mathbf{v}}_{i_2}) \\ &= \frac{1}{\sqrt{\lambda_{i_1}\lambda_{i_2}}} \vec{\mathbf{v}}_{i_1} \cdot (A^\top A\vec{\mathbf{v}}_{i_2}) \\ &= \sqrt{\frac{\lambda_{i_2}}{\lambda_{i_1}}} \vec{\mathbf{v}}_{i_1} \cdot \vec{\mathbf{v}}_{i_2} = \begin{cases} 0 & \text{if } i_1 \ne i_2 \\ 1 & \text{if } i_1 = i_2 \end{cases} \end{aligned}$$

and they do belong to W, because they are orthogonal to $\vec{\mathbf{u}}_{l+1}, \dots, \vec{\mathbf{u}}_n$: if $j > l$ then

$$\vec{\mathbf{v}}'_{i_1} \cdot \vec{\mathbf{u}}_j = \frac{1}{\sqrt{\lambda_{i_1}}} (A\vec{\mathbf{v}}_{i_1}) \cdot \vec{\mathbf{u}}_j = \frac{1}{\sqrt{\lambda_{i_1}}} (\mathbf{v}_{i_1}) \cdot A^\top \mathbf{u}_j = 0.$$

The same argument works for the $\vec{\mathbf{u}}'_j$.

Since an orthonormal set is linearly independent, we have $k \le l$ and $l \le k$ so $k = l$, and finally the basis $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_m$ and the basis $\vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_n$, where

$$\vec{\mathbf{w}}_1 = \vec{\mathbf{v}}'_1, \dots, \vec{\mathbf{w}}_k = \vec{\mathbf{v}}'_k, \vec{\mathbf{w}}_{k+1} = \vec{\mathbf{u}}_{k+1}, \dots, \vec{\mathbf{w}}_n = \vec{\mathbf{u}}_n$$

are bases answering the question.

3.9.1 a. If $a = b = 1$, so that the ellipse is a circle with radius 1, by part b we have $\kappa = 1$. If the ellipse is a circle of radius r, we have $\frac{x^2}{r^2} + \frac{y^2}{r^2} = 1$, and again using part b we have

$$\kappa = \frac{r^8}{(r^6)^{3/2}} = \frac{1}{r};$$

the larger the circle, the smaller the curvature.

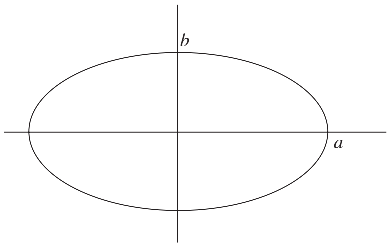


FIGURE FOR SOLUTION 3.9.1
The ellipse given by

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

b. The ellipse of equation

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

(shown in the figure at left) is the union of the graphs of

$$f(x)=\frac{b}{a}\sqrt{a^2-x^2} \quad \text{and} \quad g(x)=-\frac{b}{a}\sqrt{a^2-x^2}.$$

Clearly the curvatures at $\begin{pmatrix} x \\ f(x) \end{pmatrix}$ and $\begin{pmatrix} x \\ g(x) \end{pmatrix}$ are equal, so we will treat only f. Rather than differentiating f explicitly, we will use implicit differentiation to find

$$y'=-\frac{b^2}{a^2}\frac{x}{y} \quad \text{and} \quad y''=-\frac{b^4}{a^2y^3},$$

leading to

$$\kappa(x)=\frac{|y''|}{\big(1+(y')^2\big)^{3/2}}=\frac{b^4}{a^2y^3}\frac{1}{\big(1+\frac{b^4}{a^4}\frac{x^2}{y^2}\big)^{3/2}}=\frac{a^4b^4}{\big(a^4y^2+b^4x^2\big)^{3/2}}. \qquad (1)$$

Since we know that $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ gives the ellipse shown at left, we may wish to see whether this value for the curvature seems reasonable for certain values of a, b, and x. For example, if $a=b=1$, the ellipse is the unit circle, which should have curvature 1. This is indeed what equation 1 gives. At $x=0$ the curvature is b/a^2, while at $x=a$ it is a/b^2, a reciprocal relationship that seems reasonable.

Solution 3.9.3: We discuss the effect of the choice of point on the sign of the mean curvature at the end of the solution.

3.9.3 First, note that any point can be rotated to the point $\mathbf{a}=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, so it is enough to compute the curvatures there. Set $z=f\begin{pmatrix} x \\ y \end{pmatrix}=\sqrt{1-x^2-y^2}$. Then near $\mathbf{a}$ the "best coordinates" are $X=x$, $Y=y$, $Z=z-1$, so that

$$Z=z-1=\sqrt{1-X^2-Y^2}-1.$$

To determine the Taylor polynomial of f at the origin, we use equation 3.4.9:

$$(1+x)^m=1+mx+\frac{m(m-1)}{2!}x^2+\frac{m(m-1)(m-2)}{3!}x^3+\cdots \qquad (1)$$

with $m=1/2$, to get

$$Z\approx 1+\frac{1}{2}(-X^2-Y^2)+\frac{1}{2}\left(1-\frac{1}{2}\right)\frac{1}{2!}(-X^2-Y^2)^2+\cdots-1. \qquad (2)$$

Note that the quadratic terms of equation 2 do not correspond to the quadratic terms of equation 1.

The quadratic terms of the Taylor polynomial are

$$-\frac{X^2}{2}-\frac{Y^2}{2}.$$

Since in an adapted coordinate system the quadratic terms of the Taylor polynomial are

$$\frac{1}{2}(A_{2,0}X^2 + 2A_{1,1}XY + A_{0,2}Y^2),$$

we have $A_{2,0} = A_{0,2} = -1$ and $A_{1,1} = 0$. Definition 3.9.7 of mean curvature then gives

$$\text{mean curvature} = \frac{1}{2}(A_{2,0} + A_{0,2}) = -1$$

and Definition 3.9.8 of Gaussian curvature gives

$$\text{Gaussian curvature} = A_{2,0}A_{0,2} - A_{1,1}^2 = 1.$$

In this simple case, the "best coordinates" are automatic. Even if you did not use the X, Y, Z notation when you computed the Taylor polynomial you would have found no first-degree terms. Thus if you tried to apply Proposition 3.9.10 you would have run into

$$c = \sqrt{a_1^2 + a_2^2} = 0,$$

and all the expressions that involve dividing by c or by c^2 would not make sense. That proposition only applies to cases where at least one of a_1 and a_2 is nonzero.

What if you had chosen to work near $\begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$? In a neighborhood of this point, $z = f\begin{pmatrix} x \\ y \end{pmatrix} = -\sqrt{1 - x^2 - y^2}$, so setting $X = x$, $Y = y$, $Z = z + 1$, we would have $Z = z + 1 = -\sqrt{1 - X^2 - Y^2} + 1$, so that

$$Z \approx -1 - \frac{1}{2}(-X^2 - Y^2) - \frac{1}{2}\left(1 - \frac{1}{2}\right)\frac{1}{2!}(-X^2 - Y^2)^2 - \cdots + 1.$$

This gives a mean curvature of $+1$.

Recall that in the discussion of "best coordinates for surfaces," we said that an adapted coordinate system for a surface S at $\mathbf{a}$ "is a system where X and Y are coordinates with respect to an orthonormal basis of the tangent plane, and the Z-axis is the normal direction." Of course there are two normal directions: pointing up on the Z-axis or pointing down. If we assume that the normal is pointing up, then when we chose $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, we chose a point at which the sphere is curving away from the direction of the normal, giving a negative mean curvature. When we redid the computation at $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$, then the sphere was curving towards the direction of the normal, giving a positive mean curvature.

3.9.5 By equation 3.9.5, the curvature of the curve at $\begin{pmatrix} x \\ f(x) \end{pmatrix}$ is $\kappa = \dfrac{|f''(x)|}{(1 + (f'(x))^2)^{3/2}}$. The surface satisfies

$$a_1 = f'(x),\ a_2 = 0,\ a_{2,0} = f''(x),\ a_{1,1} = a_{0,2} = 0,$$

hence $c^2 = (f'(x))^2$, so equation 3.9.38 says that the absolute value of the mean curvature is

$$|H| = \frac{|f''(x)|}{2(1 + (f'(x))^2)^{3/2}},$$

which is indeed half the curvature as given by Proposition 3.9.2.

Solution 3.9.7 is intended to illustrate Figure 3.9.7. A goat tethered at the North Pole with a chain 2533 kilometers long would have less grass to eat (assuming grass would grow) than a goat tethered at the North Pole on a "flat" earth.

3.9.7 a. Remember (or look up) that the radius of the earth is 6366 km.[3] Remember also that the inclination of the axis of the earth to the ecliptic is 23.45°. (Again, this isn't exact, and the inclination varies periodically with a period of about 30 000 years and amplitude $\approx .5°$.)

Thus the real radius of the arctic circle (the radius as measured in the plane containing the arctic circle) is $6366 \sin(23.45°) = 2533$ km, and its circumference is $2\pi 2533 \approx 15930$ km. If the earth were flat, the circumference of a circle with the radius of the arctic circle as measured from the pole would be $2\pi 2607.5 \approx 16\,383$ km.

b. By the same computation as above, the equation for the angle (in radians) between the pole and a point of the circumference, with vertex at the center of the earth, is

$$2\pi 6366\theta - 2\pi 6366 \sin\theta = 1, \quad \text{i.e.,} \quad \theta - \sin\theta = \frac{1}{2\pi \cdot 6366}.$$

A computer will tell you that the solution to this equation is approximately $\theta \approx .05341$ radians. If you don't have a computer handy, a very good approximation is obtained by using two terms of the Taylor expansion for sine, which gives $(6/(2\pi \cdot 6366))^{1/3} \approx .05313$. The corresponding circle has radius on earth about 340 km.

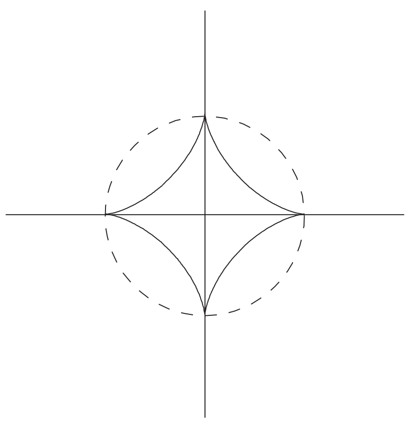


FIGURE FOR SOLUTION 3.9.9 The hypocycloid of part a is the curve inscribed in the dotted circle.

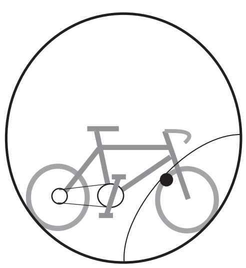


FIGURE FOR SOLUTION 3.9.9 Part b: A dot on the rim of this 2-dimensional bicycle describes a hypocycloid.

3.9.9 a. The hypocycloid looks like the solid curve in the margin (top).

b. If you ride a 2-dimensional bicycle on the inside of the dotted circle, and the radius of your wheels is 1/4 of the radius of the circle, a dot on the rim will describe a hypocycloid, as shown in the bottom figure at left.

c. Using Definition 3.9.5, the length is given by

$$\begin{aligned} &4a \int_0^{\pi/2} \sqrt{(3\sin^2 t \cos t)^2 + (3 \sin t \cos^2 t)^2}\, dt \\ &= 12a \int_0^{\pi/2} \sqrt{\sin^2 t \cos^2 t(\sin^2 t + \cos^2 t)}\, dt \\ &= 6a \int_0^{\pi/2} 2 \sin t \cos t\, dt = 6a. \end{aligned}$$

The hypocycloid has four arcs, so the length of one arc is $3a/2$.

Note that the circle has circumference $2\pi a$, so that the lengths of the circle and of the hypocycloid are close.

3.9.11 The map

$$\varphi : s \mapsto \left[\vec{\mathbf{t}}(s), \vec{\mathbf{n}}(s), \vec{\mathbf{b}}(s)\right], \quad s \in I$$

[3]Actually, the earth is flattened at the poles, and the radius ranges from 6356 kilometers at the poles to 6378 km at the equator, but the easy number to remember is that the circumference of the earth is 40 000 km, so the radius is approximately $40\,000/(2\pi) \approx 6366$ km.

(φ for "phrenet") is a mapping $I \to O(3)$. Exercise 3.2.11 says that the tangent space to $O(3)$ at the identity is the space of antisymmetric matrices. If we adapt our Frenet map φ to consider the map

$$\psi : s \mapsto \left[\vec{\mathbf{t}}(s_0), \vec{\mathbf{n}}(s_0), \vec{\mathbf{b}}(s_0)\right]^{-1} \left[\vec{\mathbf{t}}(s), \vec{\mathbf{n}}(s), \vec{\mathbf{b}}(s)\right], \quad s \in I,$$

part b of Exercise 3.2.11 tells us that this represents a parametrized curve in $O(3)$, such that $\psi(s_0) = I$. Thus its derivative $\psi'(s_0)$ is antisymmetric; temporarily let us call it

$$A = \left[\vec{\mathbf{t}}(s_0), \vec{\mathbf{n}}(s_0), \vec{\mathbf{b}}(s_0)\right]^{-1} \left[\vec{\mathbf{t}}'(s), \vec{\mathbf{n}}'(s), \vec{\mathbf{b}}'(s)\right].$$

This can be rewritten

$$\left[\vec{\mathbf{t}}'(s), \vec{\mathbf{n}}'(s), \vec{\mathbf{b}}'(s)\right] = \left[\vec{\mathbf{t}}(s_0), \vec{\mathbf{n}}(s_0), \vec{\mathbf{b}}(s_0)\right] A,$$

which, if A is the antisymmetric matrix

$$A = \begin{bmatrix} 0 & -\kappa & 0 \\ \kappa & 0 & -\tau \\ 0 & \tau & 0 \end{bmatrix},$$

corresponds to

$$\begin{aligned}
\vec{\mathbf{t}}'(s_0) &= && \kappa(s_0)\vec{\mathbf{n}}(s_0) && \\
\vec{\mathbf{n}}'(s_0) &= -\kappa(s_0)\vec{\mathbf{t}}(s_0) && && + \tau(s_0)\vec{\mathbf{b}}(s_0) \\
\vec{\mathbf{b}}'(s_0) &= && - \tau(s_0)\vec{\mathbf{n}}(s_0). &&
\end{aligned}$$

Solutions for review exercises, chapter 3

3.1 a. The derivative of the mapping $x^3 + xy^2 + yz^2 + z^3 - 4$ is

$$[\,3x^2 + y^2 \quad 2xy + z^2 \quad 2yz + 3z^2\,],$$

which only vanishes if $x = y = z = 0$. (Look at the first entry, then the second.) The origin isn't on X, so X is a smooth surface.

b. At the point $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, the derivative of F is $[\,4 \quad 3 \quad 5\,]$. The tangent plane to the surface is the plane of equation $4x + 3y + 5z = 12$.

The tangent space to the surface is the kernel of the derivative, i.e., it is the plane of equation $4\dot{x} + 3\dot{y} + 5\dot{z} = 0$.

3.3 a. The space X is given by the equations

$$\begin{aligned} (x_1 - 1)^2 + y_1^2 + z_1^2 - 1 &= 0 \\ (x_2 + 1)^2 + y_2^2 + z_2^2 - 1 &= 0 \\ (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 - 4 &= 0. \end{aligned}$$

b. The derivative of the mapping $\mathbb{R}^6 \to \mathbb{R}^3$ defining X is

$$\begin{bmatrix} 2(x_1 - 1) & 2y_1 & 2z_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2(x_2 + 1) & 2y_2 & 2z_2 \\ 2(x_1 - x_2) & 2(y_1 - y_2) & 2(z_1 - z_2) & 2(x_2 - x_1) & 2(y_2 - y_1) & 2(z_2 - z_1) \end{bmatrix}.$$

When

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \tag{1}$$

this matrix is

$$\begin{bmatrix} 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 4 & 0 & 0 & -4 & 0 & 0 \end{bmatrix}, \tag{2}$$

which has rank 3 since the first, second, and fifth columns are linearly independent. So X is a manifold of dimension 3 near the points in equation 1.

c. The tangent space is the kernel of the matrix (2) computed in part b, i.e., the set of

$$\begin{pmatrix} u_1 \\ v_1 \\ w_1 \end{pmatrix}, \quad \begin{pmatrix} u_2 \\ v_2 \\ w_2 \end{pmatrix}$$

such that $2v_1 = 0$, $2v_2 = 0$, $4(u_1 - u_2) = 0$, which is a 3-dimensional subspace of $\mathbb{R}^6$.

3.5 The best way to deal with this is to use the Taylor series for sine and cosine given in Proposition 3.4.2. Set

$$x = \frac{\pi}{6} + u, \quad y = \frac{\pi}{4} + v, \quad z = \frac{\pi}{3} + w,$$

and compute

$$\begin{aligned}\sin\left(\frac{3\pi}{4} + u + v + w\right) &= \sin\left(\frac{3\pi}{4}\right)\cos(u+v+w) + \cos\left(\frac{3\pi}{4}\right)\sin(u+v+w)\\ &= \frac{\sqrt{2}}{2}\left(1 - \frac{1}{2}(u+v+w)^2\right) - \frac{\sqrt{2}}{2}\left(u+v+w - \frac{1}{6}(u+v+w)^3\right) + \cdots\\ &= \frac{\sqrt{2}}{2}\left(1 - (u+v+w) - \frac{1}{2}(u+v+w)^2 + \frac{1}{6}(u+v+w)^3\right) + \cdots.\end{aligned}$$

Moral: If you can, compute Taylor polynomials from known expansions rather than by computing partial derivatives.

Thus the required Taylor polynomial is

$$\frac{\sqrt{2}}{2}\left(1 - (u+v+w) - \frac{1}{2}(u+v+w)^2 + \frac{1}{6}(u+v+w)^3\right).$$

Solution 3.7: The wrong way to go about this problem is to say that $\sin(x^2+y) = x^2 + y - \frac{1}{6}y^3$ plus higher-degree terms, and then apply equation 3.4.7, saying that

$$\begin{aligned}&\cos\Big(1 + \sin(x^2+y)\Big)\\ &\approx \cos(1 + x^2 + y - \frac{1}{6}y^3)\\ &= 1 - \frac{1}{2}\left(1 + x^2 + y - \frac{1}{6}y^3\right)^2\\ &\qquad + \cdots.\end{aligned}$$

This is WRONG (not just a harder way of going about things), because equation 3.4.7 is true only near the origin. At the origin,

$$1 + \sin(x^2+y) = 1,$$

so we cannot treat it as the x of equation 3.4.7, which must be 0 at the origin.

3.7 The margin note discusses the incorrect way to go about this. What we can do is use the formula

$$\cos(a+b) = \cos a \cos b - \sin a \sin b,$$

which gives

$$\cos\big(1 + \sin(x^2+y)\big) = \cos 1\big(\cos(\sin(x^2+y))\big) - \sin 1\big(\sin(\sin(x^2+y))\big).$$

At the origin, $\sin(x^2+y) = 0$, so now we can apply equations 3.4.6 and 3.4.7 (discarding terms higher than degree 3), to get

$$\begin{aligned}\cos\big(1 + \sin(x^2+y)\big) &= \cos 1\left(1 - \frac{(x^2+y)^2}{2}\right)\\ &\quad - \sin 1\left((x^2+y) - \frac{1}{6}(x^2+y)^3 - \frac{1}{6}(x^2+y)^3\right)\\ &= \cos 1 - (\sin 1)y - (\sin 1)x^2 - \frac{\cos 1}{2}y^2 - (\cos 1)x^2y - \frac{\sin 1}{3}y^3.\end{aligned}$$

3.9 Both first partials exist and are continuous for all homogeneous polynomials of degree 4, and

$$D_2\left(D_1 f\begin{pmatrix}0\\y\end{pmatrix}\right) = d \quad \text{and} \quad D_1\left(D_2 f\begin{pmatrix}x\\0\end{pmatrix}\right) = b.$$

It follows that

$$D_2\left(D_1 f\begin{pmatrix}0\\0\end{pmatrix}\right) - D_1\left(D_2 f\begin{pmatrix}0\\0\end{pmatrix}\right) = d - b,$$

and that the condition for the crossed partials to be equal is that $d = b$.

Solution 3.11: We are evaluating

$$D_z(y\cos z - x\sin z)$$

at $\begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix}$, not multiplying.

3.11 a. The implicit function theorem says that this will happen if

$$\begin{aligned} D_z(y\cos z - x\sin z)\begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix} &= -y\sin z - x\cos z\begin{pmatrix} r \\ 0 \\ 0 \end{pmatrix} \\ &= -0\cdot\sin 0 - r\cos 0 = -r \neq 0. \end{aligned}$$

b. The hint gives it away: since the x-axis is contained in the surface, we have $g_r\begin{pmatrix} x \\ 0 \end{pmatrix} = 0$ for all x, so

$$D_x^1 g_r\begin{pmatrix} r \\ 0 \end{pmatrix} = D_x^2 g_r\begin{pmatrix} r \\ 0 \end{pmatrix} = 0.$$

For a picture of this surface, see Figure 3.9.9.

3.13 a. We have

$$\det\begin{bmatrix} 1 & x & y \\ 1 & y & z \\ 1 & z & x \end{bmatrix} = xy + xz + yz - x^2 - y^2 - z^2,$$

which certainly is a quadratic form. But $\det\begin{bmatrix} 0 & x & y \\ x & 0 & z \\ y & z & 0 \end{bmatrix} = 2xyz$ certainly is not.

b. We have $xy + xz + yz - x^2 - y^2 - z^2 = -\left(x - \frac{y}{2} - \frac{z}{2}\right)^2 - \frac{3}{4}(y-z)^2$, so the signature is $(0, 2)$, and the quadratic form is degenerate.

3.15 The quadratic form represented is $ax^2 + 2bxy + dy^2$.

(if) If $a+d > 0$, at least one of a and d is positive, and then if $ad-b^2 > 0$, we must have both $a > 0$, $d > 0$. So we can write

$$\begin{aligned} ax^2 + 2bxy + dy^2 &= \left(ax^2 + 2bxy + \frac{b^2}{a}y^2\right) + \frac{ad-b^2}{a}y^2 \\ &= a\left(x + \frac{b}{a}y\right)^2 + \frac{ad-b^2}{a}y^2, \end{aligned}$$

which is strictly positive if $\begin{bmatrix} x \\ y \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(only if) If you apply the quadratic form to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, you get

$$a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\cdot G\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad d = \begin{bmatrix} 0 \\ 1 \end{bmatrix}\cdot G\begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

so if the quadratic form is positive definite, you find $a > 0$, $d > 0$, hence $a + d > 0$.

Now apply the quadratic form to the vector $\begin{bmatrix} -b \\ a \end{bmatrix}$, to find

$$\begin{bmatrix} -b \\ a \end{bmatrix}\cdot\left(\begin{bmatrix} a & b \\ b & d \end{bmatrix}\begin{bmatrix} -b \\ a \end{bmatrix}\right) = a(ad-b^2).$$

Since this must also be positive, we find $ad - b^2 > 0$.

3.17 a. The critical points are $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, since at those points the partial derivatives $D_1 f = 6x - 6y$ and $D_2 f = -6x + 6y^2$ vanish.

b. We have $D_1^2 f = 6$, $D_1 D_2 f = -6$, and $D_2^2 f = 12y$, so the second-degree term of the Taylor polynomial is $3h_1^2 - 6h_1h_2 + 6yh_2^2$. At the critical point $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, the term $6yh_2^2$ vanishes, so

$$Q(\vec{\mathbf{h}}) = 3(h_1^2 - 2h_1h_2) = 3\big((h_1 - h_2)^2 - (h_2)^2\big);$$

the signature of this quadratic form is $(1,1)$, and the critical point is a saddle.

At the critical point $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, we have $6yh_2^2 = 6h_2^2$, so

$$Q(\vec{\mathbf{h}}) = 3\big(h_1^2 - 2h_1h_2 + 2h_2^2\big) = 3\Big((h_1 - h_2)^2 + h_2^2\Big);$$

the signature of Q is $(2,0)$, and the critical point is a minimum.

3.19 a. The function is $1 + 2xyz - x^2 - y^2 - z^2$. The critical points occur where all three partials vanish, i.e., where

$$\begin{aligned} 2yz - 2x &= 0 \\ 2xz - 2y &= 0 \\ 2xy - 2z &= 0. \end{aligned}$$

Since $x = yz$, we have

$$x^2 = x(yz),$$

and so on.

It follows from these equations that $x^2 = y^2 = z^2 = xyz$. One solution is $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. The other solutions are of the form

$$\begin{pmatrix} a \\ a \\ a \end{pmatrix}, \quad \begin{pmatrix} a \\ a \\ -a \end{pmatrix}, \quad \begin{pmatrix} a \\ -a \\ -a \end{pmatrix}, \quad \begin{pmatrix} a \\ -a \\ a \end{pmatrix},$$

for an appropriate value of a. These are solutions of $a = a^2$ in the first and third, and $a = -a^2$ in the second and fourth, so the critical points are

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}.$$

b. At the origin, the quadratic terms are evidently $-x^2 - y^2 - z^2$, so that point is a maximum.

At the point $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, if we set $x = 1 + u$, $y = 1 + v$, $z = 1 + w$, and multiply out, keeping only the quadratic terms, we find

$$2(1+u)(1+v)(1+w) - (1+u)^2 - (1+v)^2 + (1+w)^2 = -(u-v-w)^2 + (v+w)^2 - (v-w)^2.$$

This point is a $(1,2)$-saddle.

At the point $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$, if we set $x = 1+u,\ y = -1+v,\ z = -1+w$, and multiply out, keeping only the quadratic terms, we find

$$2(1+u)(-1+v)(-1+w)-(1+u)^2-(-1+v)^2+(-1+w)^2 = -(u+v+w)^2+(v+w)^2-(v-w)^2.$$

This point is also a $(1,2)$-saddle. Since the expression for the function is symmetric with respect to the variables, evidently the other two points behave the same way.

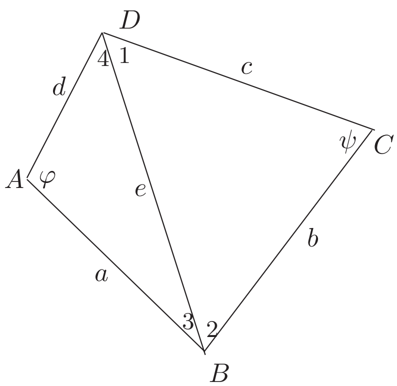


Figure for solution 3.21.

We compute the area of the quadrilateral $ABCD$ by computing the area of the two triangles, using the formula area $= \frac{1}{2}ab\sin\theta$, where θ is the angle between sides a and b.

3.21 a. Consider the figure at left. The cosine law says

$$e^2 = a^2 + d^2 - 2ad\cos\varphi = b^2 + c^2 - 2bc\cos\psi.$$

b. The area of the quadrilateral is given by $\frac{1}{2}(ad\sin\varphi + bc\sin\psi)$.

c. Our quadrilateral satisfies the constraint

$$a^2 + d^2 - 2ad\cos\varphi - b^2 - c^2 + 2bc\cos\psi = 0.$$

So the Lagrange multiplier theorem asserts that at the maximum of the area function, there is a number λ such that

$$[ad\cos\varphi,\ bc\cos\psi] = \lambda[2ad\sin\varphi,\ -2bc\sin\psi].$$

(To find the critical point of the area, we can ignore the 1/2 in the formula for the area.) This immediately gives $\cot\varphi + \cot\psi = 0$, i.e., opposite angles are supplementary (i.e., sum to 180°). It follows from high school geometry that the quadrilateral can be inscribed in a circle; see the figure below.

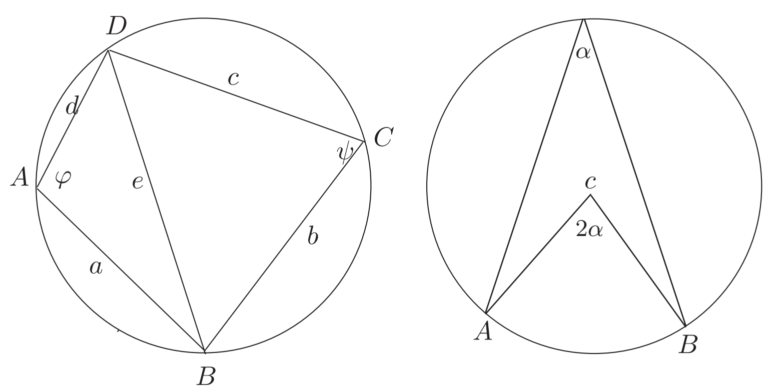


Figure for part c, Solution 3.21. Left: Our quadrilateral, inscribed in a circle. Right: An angle inscribed in a circle is half the corresponding angle at the center of the circle. This is key to proving that a quadrilateral can be inscribed in a circle if and only if opposite angles are supplementary, a result you may remember from high school.

3.23 Computing the derivatives of $f\begin{pmatrix} x \\ y \end{pmatrix} = \sqrt{x^2+y^2}$ gives the following quantities (see equation 3.9.33), which enter in Proposition 3.9.11:

$$D_1 f\begin{pmatrix} a \\ b \end{pmatrix} = a_1 = \frac{a}{\sqrt{a^2+b^2}}, \quad D_2 f\begin{pmatrix} a \\ b \end{pmatrix} = a_2 = \frac{b}{\sqrt{a^2+b^2}}, \quad c = 1,$$

$$a_{2,0} = \frac{b^2}{(a^2+b^2)^{3/2}}, \quad a_{1,1} = -\frac{ab}{(a^2+b^2)^{3/2}}, \quad a_{0,2} = \frac{a^2}{(a^2+b^2)^{3/2}}.$$

Inserting these values into equation 3.9.37, we see that the Gaussian curvature is

$$K = \frac{a^2b^2 - a^2b^2}{2^2(a^2+b^2)^3} = 0.$$

By equation 3.9.38, the mean curvature is

$$H = \frac{1}{2\sqrt{2}\sqrt{a^2+b^2}}.$$

A goat grazing on the flank of a conical volcano, constrained by a chain of a given length, too short for him to reach the top, has access to the same amount of grass as a goat grazing on a flat surface, and constrained by the same length chain.

Thus the mean curvature decreases as $|a|$ and $|b|$ increase. This is what one would expect: our surface is a cone, so it becomes flatter as we move away from the vertex of the cone. If we draw a circle on the cone near its vertex, and determine the minimal surface bounded by that circle, then repeat the procedure further and further away from the vertex, the difference between the area of the disc on the cone and the area of the minimal surface will decrease with distance from the vertex.

One way to explain why the Gaussian curvature is 0 is to look at equation 3.9.32. Our surface is a cone, which can be made out of a flat piece of paper, with no distortion. Thus the area of a disc around a point on the surface (a point not the vertex of the cone) is the same as the area of a flat disc with the same radius. Therefore, K must be 0.

The quadratic form (1) is called the *second fundamental form* of a surface at a point.

It follows from $A_{1,1}^2 = A_{0,2}A_{2,0}$ that $A_{0,2}$ and $A_{2,0}$ are either both positive or both negative.

We could also see that the second fundamental form is degenerate by considering the first equation in Solution 3.28, substituting $A_{2,0}$ for a, $A_{1,1}$ for $2b$, and $A_{0,2}$ for c: the only way the quadratic form

$$a\left(x + \frac{b}{a}y\right)^2 + \frac{ac-b^2}{a}y^2$$

in the third line can be degenerate is if $ac = b^2$, i.e., $A_{1,1}^2 = A_{0,2}A_{2,0}$

For another explanation, note that by Definition 3.9.8, Gaussian curvature 0 implies $A_{1,1}^2 = A_{0,2}A_{2,0}$. If both $A_{0,2}$ and $A_{2,0}$ are positive, this means that the quadratic form

$$\frac{1}{2}(A_{2,0}X^2 + 2A_{1,1}XY + A_{0,2}Y^2) \tag{1}$$

(see equation 3.9.28) can be written

$$(\sqrt{A_{2,0}}\,X + \sqrt{A_{0,2}}\,Y)^2 = A_{2,0}X^2 + 2A_{1,1}XY + A_{0,2}Y^2.$$

If both $A_{0,2}$ and $A_{2,0}$ are negative, it can be written

$$-\big(\sqrt{-A_{2,0}}\,X - \sqrt{-A_{0,2}}\,Y\big)^2$$

Thus the quadratic form (second fundamental form) is degenerate.

Why should the second fundamental form of a cone be degenerate? In adapted coordinates, a surface is locally the graph of a function whose domain is the tangent space and whose values are in the normal line; see equation 3.9.28. Thus the quadratic terms making up the second fundamental form are also a map from the tangent plane to the normal line. In the case of a cone, the tangent plane to a cone contains a line of the cone, so the quadratic terms vanish on that line.

3.25 This is similar to Exercise 3.7.14 but more complicated, since the second partials of the constraint function $\mathbf{F}$ do not all vanish. Therefore, when constructing the augmented Hessian matrix, one must take into account these second derivatives and the Lagrange multipliers corresponding to each critical point. We have

$$D_1 f = 2x, \; D_2 f = 2y, \; D_3 f = dz,$$

so

$$\begin{aligned} D_1 D_1 f &= D_2 D_2 f = D_3 D_3 f = 2 \\ D_1 D_2 f &= D_2 D_1 f = 0 \\ D_2 D_3 f &= D_3 D_2 f = 0. \end{aligned}$$

Moreover,

$$\begin{aligned} D_1 F_1 &= 2x \qquad & D_1 F_2 &= 1 \\ D_2 F_1 &= 2y \qquad & D_2 F_2 &= 0 \\ D_3 F_1 &= 0 \qquad & D_3 F_2 &= -1, \end{aligned}$$

So we have (where B is the matrix in equation 3.7.42)

$$\begin{aligned} B_{1,1} &= \underbrace{2}_{D_1 D_1 f} - \underbrace{2\lambda_1}_{\lambda_1 D_1 D_1 F_1} - \underbrace{0\lambda_2}_{\lambda_2 D_1 D_1 F_2} = 2 - 2\lambda_1 \\ B_{2,1} &= \underbrace{0}_{D_2 D_1 f} - \underbrace{0\lambda_1}_{\lambda_1 D_2 D_1 F_1} - \underbrace{0\lambda_2}_{\lambda_2 D_2 D_1 F_2} = 0 \\ B_{3,1} &= 0 \\ B_{2,2} &= \underbrace{2}_{D_2 D_2 f} - \underbrace{2\lambda_1}_{\lambda_1 D_2 D_2 F_1} - \underbrace{0}_{\lambda_2 D_2 D_2 F_2} = 2 - 2\lambda_1 \\ B_{2,3} &= 0 \\ B_{3,3} &= \underbrace{2}_{D_3 D_3 f} - \underbrace{0}_{\lambda_1 D_3 D_3 F_1} - \underbrace{0}_{\lambda_2 D_3 D_3 F_2} = 2 \end{aligned}$$

This gives the augmented Hessian matrix

$$H = \begin{bmatrix} 2 - 2\lambda_1 & 0 & 0 & -2x & -1 \\ 0 & 2 - 2\lambda_1 & 0 & -2y & 0 \\ 0 & 0 & 2 & 0 & 1 \\ -2x & -2y & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

We now analyze the four critical points

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}.$$

First critical point

At the critical point $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, we have $\lambda_1 = 1, \lambda_2 = 0$, so

$$H = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 2 & 0 & 1 \\ 0 & -2 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

If we label our variables $A - E$, this corresponds to the quadratic form

$$-2AE - 4BD + 2CE + 2C^2.$$

Completing squares gives

$$\left(\sqrt{2}C + \frac{E}{\sqrt{2}}\right)^2 - \left(\frac{E}{\sqrt{2}} + \sqrt{2}A\right)^2 + 2A^2 - (D+B)^2 + (B-D)^2,$$

with signature $(3, 2)$. In this case, $m = 2$ (there are two constraints); therefore, by Theorem 3.7.13, the constrained critical point has signature $(1, 0)$ and the point is a minimum.

Second critical point

At the critical point $\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$, the quadratic form is the same except that we have $+4BD$ rather than $-4BD$:

$$-2AE + 4BD + 2CE + 2C^2,$$

which can be written

$$\left(\sqrt{2}C + \frac{E}{\sqrt{2}}\right)^2 - \left(\frac{E}{\sqrt{2}} + \sqrt{2}A\right)^2 + 2A^2 + (D+B)^2 - (B-D)^2.$$

Again, the constrained critical point is a minimum, with signature $(1, 0)$.

Third critical point

At the critical point $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, we have $\lambda_1 = 2$ and $\lambda_2 = -2$, and the augmented Hessian matrix is

$$H = \begin{bmatrix} -2 & 0 & 0 & -2 & -1 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 \\ -2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

This corresponds to the quadratic form

$$-2A^2 - 2B^2 + 2C^2 - 4AD - 2AE + 2CE,$$

which can be written

$$\left(\sqrt{2}C + \frac{E}{\sqrt{2}}\right)^2 - \left(\frac{E}{\sqrt{2}} + \sqrt{2}A\right)^2 - 2B^2 - (D+A)^2 + (A-D)^2$$

This has signature $(2,3)$, so the constrained critical point has signature $(0,1)$; it is a maximum.

Fourth critical point

At the critical point $\begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$, we have $\lambda_1 = \lambda_2 = 2$, giving

$$H = \begin{bmatrix} -2 & 0 & 0 & 2 & -1 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

This corresponds to

$$-2A^2 - 2B^2 + 2C^2 + 4AD - 2AE + 2CE,$$

which can be written

$$\left(\sqrt{2}C + \frac{E}{\sqrt{2}}\right)^2 - \left(\frac{E}{\sqrt{2}} + \sqrt{2}A\right)^2 - 2B^2 + (A+D)^2 - (A-D)^2.$$

This has signature $(2,3)$, so the constrained critical point has signature $(0,1)$; it is a maximum.

3.27 Proposition 3.9.11 gives a formula for the mean curvature of a surface in terms of the Taylor polynomial, up to degree 2, of the function of which the surface is a graph. Consider Scherk's surface as the graph of

$$z = f\begin{pmatrix} x \\ y \end{pmatrix} = \ln\left(\frac{\cos x}{\cos y}\right) = \ln(\cos x) - \ln(\cos y). \qquad (1)$$

By Proposition 3.9.10, the Taylor polynomial is computed at the origin, so we make the change of variables $x = x_0 + u$, $y = y_0 + v$. Then the Taylor polynomial, to degree 2, of the first term of equation (1) is

$$\begin{aligned} \ln(\cos(x_0+u)) &= \ln(\cos x_0 \cos u - \sin x_0 \sin u) \\ &= \ln\left(\cos x_0\left(\cos u - \frac{\sin x_0}{\cos x_0}\sin u\right)\right) \\ &= \ln(\cos x_0) + \ln\left(\cos u - \frac{\sin x_0}{\cos x_0}\sin u\right) \\ &\overset{(1)}{=} \ln(\cos x_0) + \ln\left(1 - u\frac{\sin x_0}{\cos x_0} - \frac{u^2}{2} + o(u^2)\right) \\ &\overset{(2)}{=} \ln(\cos x_0) - \left(u\frac{\sin x_0}{\cos x_0} + \frac{u^2}{2}\right) - \frac{1}{2}\left(u\frac{\sin x_0}{\cos x_0} + \frac{u^2}{2}\right)^2 + o(u^2) \\ &= \ln(\cos x_0) - u\frac{\sin x_0}{\cos x_0} - \frac{u^2}{2}\left(1 + \left(\frac{\sin x_0}{\cos x_0}\right)^2\right) + o(u^2). \end{aligned}$$

To get the equality marked (1) we use

$$\cos u = 1 - \frac{u^2}{2} + o(u^2)$$

and

$$\sin u = u + o(u^2).$$

To get the equality marked (2) we use

$$\ln(1-x) = -x - \frac{x^2}{2} + o(x^2).$$

A similar formula holds for $\ln\cos(y_0+v)$. Thus

$$\begin{aligned}\ln(\cos x) - \ln(\cos y) &= \ln\big(\cos(x_0+u)\big) - \ln\big(\cos(y_0+v)\big)\\ &= \ln\left(\frac{\cos x_0}{\cos y_0}\right) - u\,\overbrace{\frac{\sin x_0}{\cos x_0}}^{a_1} + v\,\overbrace{\frac{\sin y_0}{\cos y_0}}^{a_2} + \frac{1}{2}\left(\overbrace{-\left(1+\left(\frac{\sin x_0}{\cos x_0}\right)^2\right)}^{a_{2,0}} u^2\right)\\ &\quad + \frac{1}{2}\left(\underbrace{\left(1+\left(\frac{\sin x_0}{\cos x_0}\right)^2\right)}_{a_{0,2}} v^2\right) + o(u^2+v^2).\end{aligned}$$

So by equation 3.9.38, the mean curvature is

$$\left(-\frac{1}{2(1+c^2)^{3/2}}\right)\Big(a_{2,0}(1+a_2^2) - 2a_1a_2a_{1,1} + a_{0,2}(1+a_1^2)\Big),$$

where the second factor is

$$\begin{aligned}&\Big(a_{2,0}(1+a_2^2) - 2a_1a_2a_{1,1} + a_{0,2}(1+a_1^2)\Big)\\ &= -\left(1+\left(\frac{\sin x_0}{\cos x_0}\right)^2\right)\left(1+\left(\frac{\sin y_0}{\cos y_0}\right)^2\right) + \left(1+\left(\frac{\sin x_0}{\cos x_0}\right)^2\right)\left(1+\left(\frac{\sin y_0}{\cos y_0}\right)^2\right)\\ &= 0.\end{aligned}$$

3.29 a. For any angle, B is symmetric:

$$B^\top = (Q_{i,j}^\top A Q_{i,j})^\top = Q_{i,j}^\top A^\top Q_{i,j} = Q_{i,j}^\top A Q_{i,j},$$

since A is symmetric.

This method for diagonalizing symmetric matrices is known as *Jacobi's method*.

We have

$$\begin{bmatrix} b_{i,i} & b_{i,j}\\ b_{j,i} & b_{j,j}\end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix} a_{i,i} & a_{i,j}\\ a_{j,i} & a_{j,j}\end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}.$$

If you carry out the multiplication, you will find

$$b_{i,j} = b_{j,i} = (a_{j,j} - a_{i,i})\sin\theta\cos\theta + a_{i,j}(\cos^2\theta - \sin^2\theta).$$

Thus if

$$\tan 2\theta = \frac{2a_{i,j}}{a_{i,i} - a_{j,j}}$$

we will have $b_{i,j} = b_{j,i} = 0$.

b. This breaks into three parts:

(1) If $k \neq i, k \neq j, l \neq i, l \neq j$, then $a_{k,l} = b_{k,l}$.

(2) Suppose $k \neq i$ and $k \neq j$. Then $b_{k,i}^2 + b_{k,j}^2 = a_{k,i}^2 + a_{k,j}^2$. Indeed, both are the length squared of the projection of $A\vec{\mathbf{e}}_k = B\vec{\mathbf{e}}_k$ onto the plane spanned by $\vec{\mathbf{e}}_i$ and $\vec{\mathbf{e}}_j$; which is also the plane spanned by

$$(\cos\theta)\vec{\mathbf{e}}_i + (\sin\theta)\vec{\mathbf{e}}_j \quad\text{and}\quad (-\sin\theta)\vec{\mathbf{e}}_i + (\cos\theta)\vec{\mathbf{e}}_j.$$

(3) Suppose $k \neq i$ and $k \neq j$. Then $b_{i,k}^2 + b_{k,j}^2 = a_{i,k}^2 + a_{j,k}^2$. This follows from case (2), using that A and B are symmetric, i.e., $a_{i,k} = a_{k,i}$ etc.

It now follows that

$$\sum_{k \neq l} b_{k,l}^2 = \sum_{k \neq l} a_{k,l}^2 - a_{i,j}^2 + a_{j,i}^2,$$

and the desired result follws from taking only the strictly subdiagonal terms.

c. Start with $A_0 = A$, let (i_0, j_0) be the position of the largest off-diagonal $|a_{i,j}|$, find θ_0 so as to cancel it, set $A_1 \overset{\text{def}}{=} Q_0^\top A_0 Q_0$.

Now find the largest off-diagonal entry of A_1, then θ_1 and Q_1 as above and set $A_2 = Q_1^\top A_1 Q_1$.

We have never seen a case where the sequence $k \mapsto A_k$ fails to converge, but here is a proof that the sequence always has a convergent sequence.

Set $P_m = Q_0 Q_1 \cdots Q_m$. This is a sequence in the orthogonal group $O(n)$, so it has a convergent subsequence $i \mapsto P_{m_i}$ that converges, say to the orthogonal matrix P. Then $P^\top AP$ is diagonal.

3.31 Using the Lagrange multipliers theorem, we find that the critical points of the given function occur when

$$\begin{aligned} x_2 x_3 \cdots x_n &= 2\lambda x_1 \\ x_1 x_3 x_4 \cdots x_n &= 4\lambda x_2 \\ &\vdots \\ x_1 \cdots x_{n-1} &= 2n\lambda x_n. \end{aligned} \tag{1}$$

From the first two equations, we find that

$$x_1 = \frac{x_2 \cdots x_n}{2\lambda} \quad \text{and} \quad x_2 = \frac{x_1 x_3 \cdots x_n}{4\lambda} = \frac{x_2 (x_3 \cdots x_n)^2}{8\lambda^2}.$$

(Note we have assumed that none of the x_i are zero, which will certainly be the case at the desired maximum.) Thus

$$\lambda = \frac{x_3 \dots x_n}{2\sqrt{2}}.$$

(Actually, λ may be negative as well, but assuming it is positive has no effect whatsoever on the computation; this should be clear as we proceed.) Substituting this expression for λ back into (1), we find $x_2 = \frac{x_1}{\sqrt{2}}$. Substituting the results we have gathered into each equation successively, we find that $x_i = \frac{x_1}{\sqrt{i}}$. We then substitute these results into the constraint equation $x_1^2 + 2x_2^2 + \cdots + nx_n^2 = 1$ to find $nx_1^2 = 1$, or $x_1 = \frac{1}{\sqrt{n}}$. Thus $x_i = \frac{1}{\sqrt{ni}}$. Thus the maximum of our function is $\frac{1}{n^{n/2}\sqrt{n!}}$, which can also be written $\left(\frac{1}{\sqrt{n}}\right)^n \frac{1}{\sqrt{n!}}$.

Solutions for Chapter 4

4.1.1 a. The area of a dyadic cube $C \in \mathcal{D}_3(\mathbb{R}^2)$ is $(\frac{1}{2^3})^2 = \frac{1}{64}$; the area of a cube $C \in \mathcal{D}_4(\mathbb{R}^2)$ is $(\frac{1}{2^4})^2 = (\frac{1}{16})^2$; the area of a cube $C \in \mathcal{D}_5(\mathbb{R}^2)$ is $(\frac{1}{2^5})^2$.

b. The volume of a dyadic cube $C \in \mathcal{D}_3(\mathbb{R}^3)$ is $(\frac{1}{2^3})^3$; the volume of a cube $C \in \mathcal{D}_4(\mathbb{R}^3)$ is $(\frac{1}{2^4})^3$; the volume of $C \in \mathcal{D}_5(\mathbb{R}^3)$ is $(\frac{1}{2^5})^3$.

4.1.3 a. 2-dimensional volume (i.e., area) $(\frac{1}{2^3})^2 = 1/64$.

b. 3-dimensional volume $(\frac{1}{2^2})^3 = 1/64$

c. 4-dimensional volume $(\frac{1}{2^3})^4 = (\frac{1}{8})^4$

d. 3-dimensional volume $(\frac{1}{2^3})^3 = (\frac{1}{8})^3$

For each cube, the first index gives unnecessary information. This index is $\mathbf{k}$, which tells where the cube is. The number n of entries of $\mathbf{k}$ gives necessary information: if it is 2, the cube is in $\mathbb{R}^2$, if it is 3, the cube is in $\mathbb{R}^3$, and so on. But we do not need to know what the entries are to compute the n-dimensional volume.

4.1.5 a. $$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}.$$

b.
$$\int_{\mathbb{R}} x\mathbf{1}_{[0,1)}(x)\,|dx| = \int_{\mathbb{R}} x\mathbf{1}_{[0,1]}(x)\,|dx| = \int_{\mathbb{R}} x\mathbf{1}_{(0,1]}(x)\,|dx| = \int_{\mathbb{R}} x\mathbf{1}_{(0,1)}(x)\,|dx| = \frac{1}{2}.$$

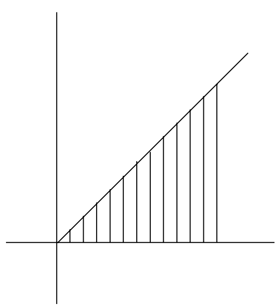


Figure for Solution 4.1.5.

We will write out the first one in detail. Since our theory of integration is based on dyadic decompositions, we divide the interval from 0 to 1 into 2^N pieces, as suggested by the figure at left.

Let us compare upper and lower sums when the interval is $[0,1)$. For the lower sum (which corresponds to the left Riemann sum, since the function is increasing), the width of each piece is $\frac{1}{2^N}$, the height of the first is 0, that of the second is $\frac{1}{2^N}$, and so on, ending with $\frac{2^N-1}{2^N}$. This gives

$$L_N\left(x\mathbf{1}_{[0,1)}(x)\right) = \frac{1}{2^N}\sum_{i=0}^{2^N-1}\frac{i}{2^N} = \frac{1}{2^{2N}}\sum_{i=0}^{2^N-1} i.$$

For the upper sum (corresponding to the right Riemann sum), the height of the first piece is $\frac{1}{2^N}$, that of the second is $\frac{2}{2^N}$, and so on, ending with $\frac{2^N}{2^N} = 1$:

$$U_N\left(x\mathbf{1}_{[0,1)}(x)\right) = \frac{1}{2^N}\sum_{i=1}^{2^N}\frac{i}{2^N} = \frac{1}{2^N}\sum_{i=0}^{2^N-1}\frac{i+1}{2^N}.$$

So

$$U_N - L_N = \frac{1}{2^N},$$

which goes to 0 as $N \to \infty$, so the function is integrable. Using the result from part a, we see that it has integral 1/2:

$$\lim_{N\to\infty} L_N\left(x\mathbf{1}_{[0,1)}(x)\right) = \lim_{N\to\infty}\left(\frac{1}{2^{2N}}\cdot\frac{(2^N-1)(2^N)}{2}\right) = \lim_{N\to\infty}\left(\frac{1}{2}-\frac{1}{2^{N+1}}\right) = \frac{1}{2}.$$

For the interval $[0,1]$, note that

$$U_N\left(x\mathbf{1}_{[0,1]}(x)\right) = U_N\left(x\mathbf{1}_{[0,1)}(x)\right) + \frac{1}{2^N};$$
$$L_N\left(x\mathbf{1}_{[0,1]}(x)\right) = L_N\left(x\mathbf{1}_{[0,1)}(x)\right).$$

Therefore,

$$\int_{\mathbb{R}} x\mathbf{1}_{[0,1)}(x)\,|dx| = \int_{\mathbb{R}} x\mathbf{1}_{[0,1]}(x)\,|dx| = \frac{1}{2}.$$

The upper and lower sums for the interval $(0,1]$ are the same as those for $[0,1]$, and those for the interval $(0,1)$ are the same as those for $[0,1)$, so those integrals are equal also.

c. This is fundamentally easy; the difficulty is all in seeing exactly how many dyadic intervals are contained in $[0,a]$, and fiddling with the little bit at the right that is left over. We will require the notation $\lfloor b\rfloor$, called the *floor* of b, and which stands for the largest integer $\le b$.

Consider first the case of the closed interval, i.e., integrate $x\mathbf{1}_{[0,a]}(x)$. Then we have

$$U_N\left(x\mathbf{1}_{[0,a]}(x)\right) = \sum_{i=0}^{\lfloor 2^N a\rfloor -1}\frac{i+1}{2^{2N}} + \frac{a}{2^N} \quad\text{and}\quad L_N\left(x\mathbf{1}_{[0,a]}(x)\right) = \sum_{i=0}^{\lfloor 2^N a\rfloor -1}\frac{i}{2^{2N}}.$$

The last term in the upper sum is the contribution of the rightmost interval

$$\left[\frac{\lfloor 2^N a\rfloor}{2^N}, \frac{\lfloor 2^N a\rfloor + 1}{2^N}\right),$$

where the maximum value is a and is achieved at a.

So we find

$$U_N\left(x\mathbf{1}_{[0,a]}(x)\right) - L_N\left(x\mathbf{1}_{[0,a]}(x)\right) = \sum_{i=0}^{\lfloor 2^N a\rfloor -1}\frac{1}{2^{2N}} + \frac{a}{2^N} \le \frac{2a}{2^N},$$

which clearly tends to 0 as $N \to \infty$, so the function is integrable. To find the integral, we will use the lower sums (which are slightly less messy). We find

$$\begin{aligned} L_N\left(x\mathbf{1}_{[0,a]}(x)\right) &= \sum_{i=0}^{\lfloor 2^N a\rfloor -1}\frac{i}{2^{2N}} = \frac{1}{2^{2N}}\frac{1}{2}(\lfloor 2^N a\rfloor - 1)\lfloor 2^N a\rfloor \\ &= \frac{1}{2}\frac{\lfloor 2^N a\rfloor - 1}{2^N}\frac{\lfloor 2^N a\rfloor}{2^N} = \frac{a^2}{2}. \end{aligned}$$

For the other integrals, note that whether the interval is open or closed at 0 makes no difference to the upper or lower sums. The interval being open or closed at a makes a difference exactly when $a = k/2^N$, in which case the upper sum contains an extra term when the endpoint is there. But this extra term is $a/2^N$, and does not change the limit as $N \to \infty$.

d. Observe that

$$\begin{aligned}
x\mathbf{1}_{[a,b]}(x) &= x\mathbf{1}_{[0,b]}(x) - x\mathbf{1}_{[0,a)}(x),\\
x\mathbf{1}_{[a,b)}(x) &= x\mathbf{1}_{[0,b)}(x) - x\mathbf{1}_{[0,a)}(x),\\
x\mathbf{1}_{(a,b]}(x) &= x\mathbf{1}_{[0,b]}(x) - x\mathbf{1}_{[0,a]}(x),\\
x\mathbf{1}_{(a,b)}(x) &= x\mathbf{1}_{[0,b)}(x) - x\mathbf{1}_{[0,a]}(x).
\end{aligned}$$

By Proposition 4.1.14, all four right sides are integrable, and the integrals are all $(b^2 - a^2)/2$.

4.1.7 As suggested in the text, write $f_1 = f_1^+ - f_1^-$, $f_2 = f_2^+ - f_2^-$. Then we have

$$\begin{aligned}
\int_{\mathbb{R}^{n+m}} g\,|d^n\mathbf{x}||d^m\mathbf{y}| &= \int_{\mathbb{R}^{n+m}} \left(f_1^+(\mathbf{x}) - f_1^-(\mathbf{x})\right)\left(f_2^+(\mathbf{y}) - f_2^-(\mathbf{y})\right)|d^n\mathbf{x}||d^m\mathbf{y}|\\
&= \int_{\mathbb{R}^{n+m}} \Big(f_1^+(\mathbf{x})f_2^+(\mathbf{y}) - f_1^-(\mathbf{x})f_2^+(\mathbf{y}) - f_1^+(\mathbf{x})f_2^-(\mathbf{y}) + f_1^-(\mathbf{x})f_2^-(\mathbf{y})\Big)|d^n\mathbf{x}||d^m\mathbf{y}|\\
&= \int_{\mathbb{R}^{n+m}} f_1^+(\mathbf{x})f_2^+(\mathbf{y})|d^n\mathbf{x}||d^m\mathbf{y}| - \int_{\mathbb{R}^{n+m}} f_1^-(\mathbf{x})f_2^+(\mathbf{y})|d^n\mathbf{x}||d^m\mathbf{y}|\\
&\quad - \int_{\mathbb{R}^{n+m}} f_1^+(\mathbf{x})f_2^-(\mathbf{y})|d^n\mathbf{x}||d^m\mathbf{y}| + \int_{\mathbb{R}^{n+m}} f_1^-(\mathbf{x})f_2^-(\mathbf{y})|d^n\mathbf{x}||d^m\mathbf{y}|\\
&= \left(\int_{\mathbb{R}^n} f_1^+(\mathbf{x})|d^n\mathbf{x}|\right)\left(\int_{\mathbb{R}^m} f_2^+(\mathbf{y})|d^m\mathbf{y}|\right) - \left(\int_{\mathbb{R}^n} f_1^-(\mathbf{x})|d^n\mathbf{x}|\right)\left(\int_{\mathbb{R}^m} f_2^+(\mathbf{y})|d^m\mathbf{y}|\right)\\
&\quad - \left(\int_{\mathbb{R}^n} f_1^+(\mathbf{x})|d^n\mathbf{x}|\right)\left(\int_{\mathbb{R}^m} f_2^-(\mathbf{y})|d^m\mathbf{y}|\right) + \left(\int_{\mathbb{R}^n} f_1^+(\mathbf{x})|d^n\mathbf{x}|\right)\left(\int_{\mathbb{R}^m} f_2^+(\mathbf{y})|d^m\mathbf{y}|\right)\\
&= \left(\int_{\mathbb{R}^n} (f_1^+(\mathbf{x}) - f_1^-(\mathbf{x}))|d^n\mathbf{x}|\right)\left(\int_{\mathbb{R}^m} (f_2^+(\mathbf{y}) - f_2^-(\mathbf{y}))|d^m\mathbf{y}|\right)\\
&= \left(\int_{\mathbb{R}^n} f_1\,|d^n\mathbf{x}|\right)\left(\int_{\mathbb{R}^m} f_2\,|d^m\mathbf{y}|\right)
\end{aligned}$$

4.1.9 Using the inequality $|\sin a - \sin b| \le |a - b|$ (see the footnote in the text, page 245), we see that within each dyadic square $C \in \mathcal{D}_N(\mathbb{R}^2)$ in the interior of Q, we have

$$\begin{aligned}
|\sin(x_1 - y_1) - \sin(x_2 - y_2)| &\le |(x_1 - y_1) - (x_2 - y_2)|\\
&= |(x_1 - x_2) - (y_2 - y_1)|\\
&\le |x_1 - x_2| + |y_2 - y_1| \le 1/2^N + 1/2^N\\
&= 2^{1-N}.
\end{aligned}$$

Furthermore, at most $8(2^N)$ such squares touch the boundary of the square, where the oscillation is at most 1. Thus we have

$$U_N(f) - L_N(f) \le \frac{1}{2^{N-1}}\frac{2^{2N}}{2^{2N}} + \frac{8 \cdot 2^N}{2^{2N}} = \frac{2+8}{2^N},$$

which evidently goes to 0 as $N \to \infty$.

4.1.11 If $X \subset \mathbb{R}^n$ is any set, and a is a number, we will denote

$$aX = \{\, a\mathbf{x} \mid \mathbf{x} \in X \,\}.$$

If $C \in \mathcal{D}_M(\mathbb{R}^n)$, then $2^N C \in \mathcal{D}_{M-N}(\mathbb{R}^n)$. In particular,

$$\begin{aligned} U_M(D_{2^N} f) &= \sum_{C \in \mathcal{D}_M(\mathbb{R}^n)} M_C(D_{2^N} f)\, \mathrm{vol}_n(C) \\ &= \sum_{C \in \mathcal{D}_M(\mathbb{R}^n)} M_{2^{-N}C}(f)\, \mathrm{vol}_n(C) \\ &= 2^{nN} \sum_{C \in \mathcal{D}_{M+N}(\mathbb{R}^n)} M_C(f)\, \mathrm{vol}_n(C) = 2^{nN} U_{M+N}(f). \end{aligned}$$

An exactly similar computation gives

$$L_M(D_{2^N} f) = 2^{nN} L_{M+N}(f).$$

Putting these inequalities together, we find

$$2^{nN} L_{M+N}(f) = L_M(D_{2^N} f) \le U_M(D_{2^N} f) = 2^{nN} U_{M+N}(f).$$

The outer terms above can be made arbitrarily close, so the inner terms can also. It follows that $D_{2^N} f$ is integrable, and

$$\int_{\mathbb{R}^n} D_{2^N} f(\mathbf{x}) |d^n\mathbf{x}| = 2^{nN} \int_{\mathbb{R}^n} f(\mathbf{x}) |d^n\mathbf{x}|.$$

4.1.13 We have

$$\begin{aligned} \frac{2^N|b-a| - 2}{2^N} &= \frac{2^N|b-a|}{2^N} - \frac{2}{2^N} \le L_N(\mathbf{1}_I) \\ &\le \mathrm{vol}(I) \le U_N(\mathbf{1}_I) \le \frac{2^N|b-a| + 2}{2^N} = \frac{2^N|b-a|}{2^N} + \frac{2}{2^N}. \end{aligned}$$

So taking the limit as $N \to \infty$, we find $\mathrm{vol}(I) = |b-a|$.

4.1.15 a. $X \cap Y$: If $C \cap (X \cap Y) \ne \emptyset$ then $C \cap X \ne \emptyset$, so for a given ϵ we can use the same N as we did for X alone in equation 4.1.62.

Observe that the argument for $X \times Y$ proves a stronger statement; it shows that if X has volume 0, and Y has any *finite* volume, then $X \times Y$ has volume 0.

$X \times Y$: Using Proposition 4.1.20 we see that for any ϵ_1, it suffices to choose N to be the larger of

N in Proposition 4.1.23 for X with $\epsilon = \epsilon_1$, and

N for Y with $\epsilon = 1$.

$X \cup Y$: Whether or not X and Y are disjoint, we have

$$\mathbf{1}_{X \cup Y} \le \mathbf{1}_X + \mathbf{1}_Y,$$

so

Case $X \cup Y$: inequality 1 is Proposition 4.1.14, part 3; equality 2 is Proposition 4.1.14, part 1; equality 3: X and Y have volume 0.

$$\begin{aligned} \int_{\mathbb{R}^n} \mathbf{1}_{X \cup Y}(\mathbf{x}) |d^n\mathbf{x}| &\underset{1}{\le} \int_{\mathbb{R}^n} \Big(\mathbf{1}_X(\mathbf{x}) + \mathbf{1}_Y(\mathbf{x}) \Big) |d^n\mathbf{x}| \\ &\underset{2}{=} \int_{\mathbb{R}^n} \mathbf{1}_X(\mathbf{x}) |d^n\mathbf{x}| + \int_{\mathbb{R}^n} \mathbf{1}_Y(\mathbf{x}) |d^n\mathbf{x}| \\ &\underset{3}{=} 0 + 0 = 0. \end{aligned}$$

One could also argue that $X \cup Y = X \cup (Y - X)$ and apply Theorem 4.1.21 (which concerns disjoint sets), but that requires knowing $Y - X \subset Y$ and $\mathrm{vol}_n Y = 0 \implies \mathrm{vol}_n(Y - X) = 0$, which again requires using $\mathbf{1}_{Y-X} \leq \mathbf{1}_Y$ and applying Proposition 4.1.14, part 3.

b. Let X be the set in question. It intersects $2^N + 1$ squares of $\mathcal{D}_N(\mathbb{R}^2)$, so

$$U_N(\mathbf{1}_X) = \frac{2^N + 1}{2^{2N}}, \quad \text{which tends to 0 as } N \to \infty.$$

c. Let X be the set in question. It intersects $2^{2N} + 2^N + 2^N + 1$ cubes of $\mathcal{D}_N(\mathbb{R}^3)$, so

$$U_N(\mathbf{1}_X) = \frac{2^{2N} + 2^{N+1} + 1}{2^{3N}}, \quad \text{which tends to 0 as } N \to \infty.$$

In Exercises 4.1.17 and 4.1.19, we will suppose that we have made a decomposition

$$0 = x_0 < x_1 < \cdots < x_n = 1,$$

and will try to see what sort of Riemann sum the proposed "integrand" gives.

4.1.17 a. This is an acceptable but "degenerate" integrand: it will give 0, since

$$\sum_{i=0}^{n-1}(x_{i+1} - x_i)^2 = \underbrace{\sum_{i=0}^{n-1} x_{i+1}(x_{i+1} - x_i)}_{\substack{\text{right Riemann} \\ \text{sum for } \int_0^1 x\,dx}} - \underbrace{\sum_{i=0}^{n-1} x_i(x_{i+1} - x_i)}_{\substack{\text{left Riemann} \\ \text{sum for } \int_0^1 x\,dx}}.$$

b. This is an acceptable integrand. We have $|\sin(u) - u| < |u|^3/6$ when $|u| < 1$, since the power series for sine is alternating, with decreasing terms tending to 0. So as soon as all $|x_{i+1} - x_i| < 1$, we have

$$\left| \sum_{i=0}^{n-1} \sin(x_{i+1} - x_i) - \sum_{i=0}^{n-1}(x_{i+1} - x_i) \right| \leq \frac{1}{6}\sum_{i=0}^{n-1}(x_{i+1} - x_i)^3.$$

The right side tends to 0 as the decomposition becomes fine (for instance, it is smaller than the sum in part a), so the two terms on the left have the same limit, which is $\int_0^1 1\,dx = 1 - 0$.

c. This is not an acceptable integrand; it gives an infinite limit. Indeed, if we break up $[0, 1]$ into n intervals of equal length, each will have length $1/n$ and we get

$$\sum_{i=1}^{n} \sqrt{\frac{1}{n}} = \frac{n}{\sqrt{n}} = \sqrt{n},$$

which goes to ∞ as $n \to \infty$.

d. This is an acceptable integrand; the limit it gives is 1. Indeed,

$$\sum_{i=0}^{n-1}\left(x_{i+1}^2 - x_i^2\right) = \sum_{i=0}^{n-1} x_{i+1}(x_{i+1} - x_i) + \sum_{i=0}^{n-1} x_i(x_{i+1} - x_i).$$

As in part a, each of these is a Riemann sum for $\int_0^1 x\,dx$, so the sum tends to

$$\int_0^1 2x\,dx = 1^2 - 0^2.$$

An easier approach is to notice that the sum telescopes:

$$\sum_{i=0}^{n-1}\left(x_{i+1}^2 - x_i^2\right) = x_1^2 - x_0^2 + x_2^2 - x_1^2 + \cdots + x_n^2 - x_{n-1}^2 = x_n^2 - x_0^2 = 1 - 0 = 1.$$

e. This is similar:

$$\sum_{i=0}^{n-1} x_{i+1}^3 - x_i^3 = \sum_{i=0}^{n-1} (x_{i+1}^2 + x_{i+1}x_i + x_i^2)(x_{i+1} - x_i).$$

All three terms of the sum are Riemann sums for $\int_0^1 x^2\,dx$, so the limit of their sum is

$$\int_0^1 3x^2\,dx = 1^3 - 0^3.$$

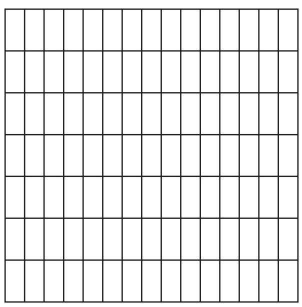


Figure for Solution 4.1.19.

This is the case $N = 15$ and $M = 7$.

Again, this is easier if you notice that the sum telescopes:

$$\sum_{i=0}^{n-1} x_{i+1}^3 - x_i^3 = x_1^3 - x_0^3 + x_2^3 - x_1^3 + \cdots + x_n^3 - x_{n-1}^3 = x_n^3 - x_0^3 = 1 - 0 = 1.$$

4.1.19 Suppose we break $[0,1] \times [0,1]$ into NM congruent rectangles, as shown in the margin. Then the relevant Riemann sum is

$$\underbrace{NM}_{\text{number of pieces}} \left(\frac{1}{N^2}\sqrt{\frac{1}{M}}\right) = \frac{\sqrt{M}}{N}.$$

The limit depends on how fast N goes to infinity relative to M. If $N = \sqrt{M}$, or if N is a multiple of $\sqrt{M}$ ($N = k\sqrt{M}$), then the Riemann sum tends to a finite limit (1 or $1/k$ respectively). If N is much smaller than $\sqrt{M}$, the limit is infinite, and if N is much bigger than M, the limit is 0. Therefore, $|b-a|^2\sqrt{|c-d|}$ is not a reasonable integrand.

4.1.21 a. It is enough to show that

$$U_N(f+g) \le U_n(f) + U_N(g) \quad \text{for every } N,$$

i.e.,

$$M_C(f+g) \le M_C(f) + M_C(g)$$

The sequence is necessary because $f+g$ might fail to achieve its supremum $M_C(f+g)$.

There exists a sequence $n \mapsto \mathbf{x}_n$ in $\mathbb{C}$ such that

$$M_C(f+g) = \lim_{n\to\infty}(f+g)(\mathbf{x}_n).$$

Then

$$M_C(f+g) = \lim_{n\to\infty}(f+g)(\mathbf{x}_n) = \lim_{n\to\infty}\big(f(\mathbf{x}_n) + g(\mathbf{x}_n)\big) \le M_C(f) + M_C(g).$$

b. Let f be the indicator function of the rationals in $[0,1]$, and g the indicator function of the irrationals in $[0,1]$: i.e., $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ and $g = \mathbf{1}_{(\mathbb{R}-\mathbb{Q})\cap[0,1]}$. Then

$$1 = U(f+g) < U(f) + U(g) = 2.$$

4.2.1 The question asked in the exercise is a bit vague: you can't reasonably ask "how likely is this particular outcome". Every individual outcome is very unlikely (though some are more unlikely than others). The question which makes sense is to ask: how many standard deviations is the observed outcome from average? You can also ask whether the number of repetitions of the experiment (in this case 15) is large enough to use the figures of Figure 4.2.5.

Problem 4.2.1: This kind of question comes up frequently in real life. Some town discovers that 3 times as many children die from leukemia as the national average for towns of that size: 9 deaths rather than 3. This creates an uproar, and an intense hunt for the cause.

But is there really a cause to discover? Among towns of that size, the number of leukemia cases will vary, and some place will have the maximum. The question is: with the number of towns that there are, would you expect to find one that far from the expected value?

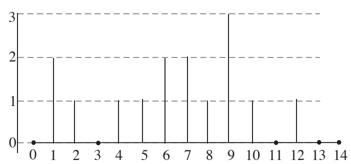


Figure for Solution 4.2.1.

None of the 14 coin tossses resulted in no heads appearing; heads appeared exactly once in two of the 14 toin cosses, and so on.

The quick answer is that these results are not very likely. If we chart results for each integer, the resulting bar graph does not look like a bell curve, as shown in the margin.

To give a more detailed answer, we must compute the standard deviation for the experiment. Figuring out all the possible ways of getting all the various totals would be time consuming, so instead we use the fact that if an experiment with standard deviation σ is repeated n times, the central limit theorem asserts that the average is approximately distributed according to the normal distribution with mean E and standard deviation $\sigma/\sqrt{n}$.

The expectation of getting heads when tossing a coin once is $1/2$. To compute the variance we consider the two cases. If we get heads, then $(f - E(f))^2 = (1 - 1/2)^2 = 1/4$; if we get tails, we also get

$$(f - E(f))^2 = (0 - 1/2)^2 = 1/4. \quad \text{So}$$

$$\mathrm{Var}(f) = E(f - E(f))^2 = E(1/4 + 1/4) = 1/2(1/2) = 1/4$$

and $\sigma(f) = 1/2$.

So the standard deviation of the average results when tossing the coin 14 times is $\frac{1}{2\sqrt{14}}$. But we are interested in the actual number of heads, not the average. Denote this random variable by T; we multiply the standard deviation of the average results by 14 to get

$$\sigma(T) = \frac{\sqrt{14}}{2} \approx 1.9.$$

But is $n = 14$ large? If we were asked in a court of law, we would not be willing to affirm that a person reporting those results was lying or cheating. If 300 people each tossed a coin 14 times, one of them might very likely come up results more than 2.5 standard deviations from the expected value.

The expectation of T is 7, so (see Figure 4.2.5) for n large we should expect 68% of our results to be between 5.1 and 8.9 heads (within one standard deviation). The actual figure is $5/15$ (but if we expanded the range to between 5 and 9, it would be $9/15 \approx 60\%$). For n large we would expect 95% to be within 3.2 and 10.8; the actual figure is $11/15 \approx 73.3\%$.

4.2.3 a.

$$E(f) = \int_{-\infty}^{\infty} f(x)\frac{1}{2a}\mathbf{1}_{[-a,a]}(x)\,dx = \int_{-a}^{a} f(x)\frac{1}{2a}\,dx = \int_{-a}^{a}\frac{x^2}{2a}\,dx$$
$$= \left[\frac{x^3}{6a}\right]_{-a}^{a} = \frac{a^2}{6} - \frac{(-a)^3}{6a} = \frac{a^2}{6} + \frac{a^2}{6} = \frac{a^2}{3}$$

$$\mathrm{Var}(f) = \int_{-\infty}^{\infty}\big(f(x) - E(f)\big)^2\frac{1}{2a}\mathbf{1}_{[-a,a]}\,dx = \int_{-a}^{a}\left(x^2 - \frac{a^2}{3}\right)^2\frac{1}{2a}\,dx$$
$$= \int_{-a}^{a}\frac{1}{2a}\left(x^4 - \frac{2a^2x^2}{3} + \frac{a^4}{9}\right)dx$$
$$= \frac{1}{2a}\left(\left[\frac{x^5}{5}\right]_{-a}^{a} - \frac{2}{3}a^2\left[\frac{x^3}{3}\right]_{-a}^{a} + \frac{a^4}{9}[x]_{-a}^{a}\right) = \frac{2}{2a}\left(\frac{a^5}{5} - \frac{2}{3}\frac{a^5}{3} + \frac{a^5}{9}\right)$$
$$= a^4\left(\frac{1}{5} - \frac{2}{9} + \frac{1}{9}\right) = \frac{4a^4}{45}.$$
$$\sigma(f) = \sqrt{\frac{4a^4}{45}} = \frac{2a^2}{3\sqrt{5}}$$

b.

$$E(f) = \int_{-\infty}^{\infty} f(x)\frac{1}{2a}\mathbf{1}_{[-a,a]}(x)\,dx = \int_{-a}^{a} f(x)\frac{1}{2a}\,dx = \int_{-a}^{a}\frac{x^3}{2a}\,dx$$
$$= \left[\frac{x^4}{8a}\right]_{-a}^{a} = \frac{a^3}{8} - \frac{(-a)^4}{8a} = 0$$
$$\mathrm{Var}(f) = \int_{-\infty}^{\infty}\big(f(x) - E(f)\big)^2\frac{1}{2a}\mathbf{1}_{[-a,a]}\,dx = \int_{-a}^{a}\left(x^3 - 0\right)^2\frac{1}{2a}\,dx$$
$$= \int_{-a}^{a}\frac{x^6}{2a}\,dx = \left[\frac{x^7}{14a}\right]_{-a}^{a} = \frac{a^7}{14a} - \frac{(-a)^7}{14a} = \frac{2a^7}{14a} = \frac{a^6}{7}$$
$$\sigma(f) = \sqrt{\frac{a^6}{7}} = \frac{a^3\sqrt{7}}{7}$$

Solution 4.3.1, part a: The first equality in the second line is justified because by Proposition 4.1.14, part 3,
$$\int f^+ \geq 0 \text{ and } \int f^- \geq 0.$$

4.3.1 a. If f is integrable, then f^+ and f^- are integrable (Corollary 4.1.15) so $|f| = f^+ + f^-$ is integrable (Proposition 4.1.14, part a). Now

$$\left|\int f\right| = \left|\int (f^+ - f^-)\right| \overbrace{=}^{\text{Prop. 4.1.14, part 1}} \left|\left(\int f^+\right) - \left(\int f^-\right)\right|$$
$$\leq \left|\int f^+\right| + \left|\int f^-\right| = \left(\int f^+\right) + \left(\int f^-\right) = \int (f^+ + f^-)$$
$$= \int |f|.$$

b. Converse counterexample: $f = 1/2 - \mathbf{1}_{\mathbb{Q}}$ restricted to $[0,1]$ has $|f| = 1/2$ (integrable) but f is not integrable.

4.3.3 a. Since the circle is symmetric, consider only the upper right quadrant, and keep in mind that we are looking just for an upper bound, not for a sharp upper bound. As suggested in the hint, divide that quadrant into two by drawing the diagonal through the origin. This line intersects the circle at the point $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$. In the part of the circle from $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to the diagonal, the slope of the circle is always less than 1, so if we look at columns of cubes, the circle never intersects more than two cubes in any one column. Starting at the origin, and going to the point $\begin{pmatrix} 1/\sqrt{2} \\ 0 \end{pmatrix}$, there are $\frac{2^N}{\sqrt{2}}$ columns. Since we can't deal in fractions of cubes, take the fractional part of that number (denoted with square brackets, $[\frac{2^N}{\sqrt{2}}]$), and add 1 for good measure. Multiplying by two gives

$$2\left(\left[\frac{2^N}{\sqrt{2}}\right] + 1\right)$$

for that eighth of the circle. This is an over-estimate, since in many columns the circle intersects only one cube, but that doesn't matter.

Solution 4.3.3: A first error to avoid is confusing the unit circle with the unit disc; the problem concerns a curve, not a surface. Second, if you looked carefully at equation 4.1.12 defining dyadic cubes, you may have realized that at level $N = 0$, six cubes are needed, rather than the obvious four, because of the points $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. This is not important, since we are concerned with the number needed as N gets big.

In the second segment of the upper right quadrant of the circle, the slope is much steeper; indeed, at $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the tangent to the circle is vertical. So if we use the same columns, then as $N \to \infty$ there is no limit to the number of cubes the circle will intersect. But if we rotate the circle by 90°, we have the same situation as before, with slope less than 1. (This method counts cubes at the point $\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$ twice, once when we look at the circle the normal way, and again when we have rotated it.) There are eight such segments in all, so for the whole circle we have an (admittedly generous) upper bound of

$$16\left(\left[\frac{2^N}{\sqrt{2}}\right] + 1\right) \approx 2^N \frac{16}{\sqrt{2}}.$$

b. Denote the unit circle by S^1. We have

$$0 \le L_N\left(\mathbf{1}_{S^1}\right) \le U_N\left(\mathbf{1}_{S^1}\right) \le \frac{1}{2^{2N}} 2^N \frac{16}{\sqrt{2}}$$

and

$$\lim_{N\to\infty} \frac{1}{2^{2N}} 2^N \frac{16}{\sqrt{2}} = 0.$$

Another solution to Exercise 4.3.5 is given in the margin on the next page.

4.3.5 Let us apply Theorem 4.3.9, which says that if f is bounded with bounded support, and continuous except on a set of volume 0 (area in this case), then it is integrable. The function $\mathbf{1}_P f$ is bounded (by 1), and its support is in P, which is bounded. Moreover, it is continuous except on the boundary of P. This boundary is the union of the arc of parabola $y = x^2$, $-1 \le x \le 1$, which is the graph of a continuous function, hence has area 0, and the segment $y = 1$, $-1 \le x \le 1$, which is also the graph of a continuous function, hence also has area 0.

4.4.1 Clearly, if a set X has measure 0 using open boxes, then it also has measure 0 using closed boxes; the closures of the open boxes used to cover X also cover X. The volume of an open box equals the volume of its closure.

The other direction is a little more difficult. For any closed box B, let B' be the open box with the same center but twice the sidelength. Note that $\mathrm{vol}_n(B') = 2^n \mathrm{vol}_n B$. If X is a set such that there exists a sequence of closed boxes with $X \subset \cup B_i$ and $\sum \mathrm{vol}_n B_i \le \epsilon$, then $X \subset \cup B'_i$ and $\sum \mathrm{vol}_n B'_i \le 2^n\epsilon$. By taking ϵ sufficiently small, we can make $2^n\epsilon$ arbitrarily small.

Solution 4.3.5: Another way to approach this problem is to claim that over a dyadic square $C \in \mathcal{D}_N(\mathbb{R}^2)$ in P that doesn't intersect the boundary of P, the oscillation of $\sin(y^2)$ is strictly less than $2\sqrt{2}/2^N$. Bounding the oscillation of a function f usually involves bounding the derivative of f and applying the mean value theorem. If $f\begin{pmatrix} x \\ y \end{pmatrix} = \sin(y^2)$, then

$$[\mathbf{D}f(\begin{smallmatrix} x \\ y \end{smallmatrix})] = [0,\ 2y\cos(y^2)]$$

and

$$0 < 2y\ \cos(y^2) < 2 \text{ for } 0 < y < 1.$$

The above result then follows from Corollary 1.9.2 and the fact that for two points $\mathbf{x}_1, \mathbf{x}_2$ in such a dyadic square,

$$|\mathbf{x}_1 - \mathbf{x}_2| \le \sqrt{2}/2^N.$$

Now we estimate how many dyadic squares can intersect the boundary of P. Since when $|x| < 1$ the curve $y = x^2$ has slope at most 2, it will intersect a vertical column of dyadic squares in at most three squares, and the horizontal line $y = 1$ will intersect exactly one square of the column.

Moreover, the right edge will intersect 2^N squares of $\mathcal{D}_N(\mathbb{R}^2)$, so altogether, at most $5 \cdot 2^N$ dyadic squares will intersect the boundary, with total area $5 \cdot 2^{-N}$. Thus for every $\epsilon > 0$, there exists N such that the function has oscillation $< \epsilon$ except on a set of dyadic cubes of total area $< \epsilon$.

4.4.3 A set $X \subset \mathbb{R}^n$ has measure 0 if for any $\epsilon > 0$ it can be covered by a sequence $i \mapsto B_i$ of open boxes such $\sum_i \mathrm{vol}_n(B_i) < \epsilon$. If X is compact, then by the Heine Borel theorem (Theorem A3.3) it is covered by finitely many of these boxes; by Proposition 4.4.6, the union of these boxes has boundary of measure 0 (in fact, it has volume 0). Let B be the union of these boxes, then $\mathbf{1}_B$ is bounded with bounded support, continuous except on a set of measure 0, hence integrable, so for N sufficiently large we have $U_N(\mathbf{1}_B) < \epsilon$. So

$$0 \le L_N(\mathbf{1}_X) \le U_N(\mathbf{1}_X) \le U_N(\mathbf{1}_B) < \epsilon,$$

and $\lim_{N\to\infty}(U_N(\mathbf{1}_X) - L_N(\mathbf{1}_X)) < \epsilon$; since ϵ is arbitrary, the limit is 0.

4.4.5 We will prove it false by showing a counterexample. Let f be the function

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } x = \frac{p}{q} \text{ is rational, written in lowest terms, and } |x| \le 1 \\ 0 & \text{if } x \text{ is irrational, or } |x| > 1, \end{cases}$$

and let $g(x) = 0$ for all $x \in \mathbb{R}$. Then g (obviously) and f (see Example 4.4.7) are integrable, both with integral 0, and $g \le f$, but

$$\{\, x \in \mathbb{R} \mid g(x) < f(x) \,\} = \mathbb{Q} \cap [0,1],$$

and $\mathbb{Q} \cap [0,1]$ does not have volume (see Example 4.4.3).

4.4.7 Let $X \subset [0,1]$ be the set of numbers that can be written using only the digits $1, \dots, 6$.

It takes six intervals of length 1/10th to cover X, namely

$$[.1, .2], \quad [.2, .3], \quad \dots \quad , [.6, .7].$$

Similarly it takes 36 intervals of length 1/100th to cover X, and more generally 6^n intervals of length $1/10^n$ to cover X. Their total length is $(6/10)^n$, with tends to 0 as n tends to infinity, so X has measure 0, in fact this proves that it has 1-dimensional volume 0. In this case we did not need Heine-Borel (as we did in Exercise 4.4.3) to show that measure 0 implies volume 0.

4.5.1 By Proposition 4.3.5, any bounded part of a line in the plane has 2-dimensional volume (area) 0, so the values of the function on a line do not affect the value of the integral.

4.5.3 a. The integral becomes

$$\int_{-1}^{0}\left(\int_{-y}^{2+y} f\begin{pmatrix}x\\y\end{pmatrix}\,dx\right)dy+\int_{0}^{1}\left(\int_{\sqrt{y}}^{2-\sqrt{y}} f\begin{pmatrix}x\\y\end{pmatrix}\,dx\right)dy.$$

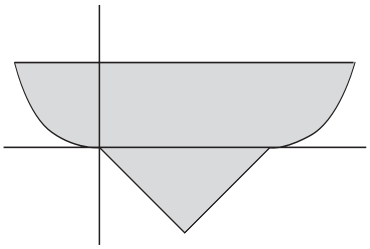


FIGURE FOR SOLUTION 4.5.3.

b. Below the x-axis, the domain of integration would be unchanged, as shown in Figure 4.5.4 in the textbook. Above the x-axis, it would be the region bounded on the left by the "left branch" of the parabola $y = x^2$, bounded on the right by the "right branch" of the parabola $y = (x-2)^2$, and bounded above by the line $y = 1$. This domain of integration is shown in the figure in the margin.

4.5.5 a. We use induction over k. Suppose we know that $\beta_{2k} = \frac{\pi^k}{k!}$. The relation $\beta_i = c_i\beta_{i-1}$ allows us to express β_{2k+2} in terms of β_{2k}, c_{2k+2}, and c_{2k+1}:

$$\beta_{2k+2} = c_{2k+2}\beta_{2k+1} = c_{2k+2}c_{2k+1}\beta_{2k}.$$

The expression for c_i is

$$\underbrace{c_i = c_{2k} = c_0\prod_{j=1}^{k}\frac{2j-1}{2j}}_{\text{for } i \text{ even}} \quad \text{and} \quad \underbrace{c_i = c_{2k+1} = c_1\prod_{j=1}^{k}\frac{2j}{2j+1}}_{\text{for } i \text{ odd}}.$$

Thus

$$c_ic_{i-1} = \frac{(i-1)!}{i!}c_0c_1 = \frac{2\pi}{i}.$$

Substituting this into the equation for β_{2k+2} gives

$$\beta_{2k+2} = c_{2k+2}c_{2k+1}\beta_{2k} = \frac{2\pi}{2k+2}\frac{\pi^k}{k!} = \frac{\pi^{k+1}}{(k+1)!},$$

as desired.

b. Suppose we know that

$$\beta_{2k-1} = \frac{\pi^{k-1}(k-1)!2^{2k-1}}{(2k-1)!}.$$

Then, as before, we have

$$\beta_{2k+1} = c_{2k+1}c_{2k}\beta_{2k-1} = \frac{2\pi}{2k+1}\frac{\pi^{k-1}(k-1)!2^{2k-1}}{(2k-1)!} = \frac{\pi^k(k-1)!2^{2k}}{(2k+1)(2k-1)!}.$$

Multiplying top and bottom by $2k$ gives

$$\beta_{2k+1} = \frac{\pi^k k!2^{2k+1}}{(2k+1)!}, \quad \text{as desired.}$$

4.5.7 We will do z, y, x first:

$$\int_0^1 \left(\int_0^{\frac{1}{2}(1-x)} \left(\int_0^{\frac{1}{3}(1-x-2y)} xyz\, dz \right) dy \right) dx.$$

The integral is *not*

$$\int_0^{1-2y-3z} \left(\int_0^{\frac{1}{2}(1-x-3z)} \left(\int_0^{\frac{1}{3}(1-x-2y)} xyz\, dz \right) dy \right) dx.$$

Solution 4.5.7: Note that since you are not asked to actually compute the integral, whether the function being integrated is xyz or something else is irrelevant.

Why not? The above, incorrect integral does not take into account all the available information. When we integrate first with respect to x, to determine the upper limit of integration we ask what is the most that x can be. Since $x, y, z \geq 0$, it follows that x is biggest when $y, z = 0$; substituting 0 for y and z in the inequality $x+2y+3z \leq 1$ gives $x \leq 1$, so the upper limit of integration is 1. When we next integrate with respect to y, we are doing so for all values of x between 0 and 1. Given that constraint, y is biggest when z is 0, which gives $x + 2y \leq 1$, or $y \leq 1/2(1-x)$. When we get to the third variable, which in this case is z, we are integrating for all permissible values of x and y, so the upper limit of integration is $1/3(1 - x - 2y)$.

The five other ways of setting up the integral are:

1. $\displaystyle\int_0^{1/2}\left(\int_0^{1-2y}\left(\int_0^{\frac{1}{3}(1-x-2y)} xyzdz\right)dx\right)dy$
2. $\displaystyle\int_0^{1/3}\left(\int_0^{\frac{1}{2}(1-3z)}\left(\int_0^{1-2y-3z} xyz\, dx\right)dy\right)dz$
3. $\displaystyle\int_0^{1/2}\left(\int_0^{\frac{1}{3}(1-2y)}\left(\int_0^{1-2y-3z} xyz\, dx\right)dz\right)dy$
4. $\displaystyle\int_0^{1/3}\left(\int_0^{1-3z}\left(\int_0^{\frac{1}{2}(1-x-3z)} xyz\, dy\right)dx\right)dz$
5. $\displaystyle\int_0^{1}\left(\int_0^{\frac{1}{3}(1-x)}\left(\int_0^{\frac{1}{2}(1-x-3z)} xyz\, dy\right)dz\right)dx.$

4.5.9 As indicated in the text, the integral for the half above the diagonal is

$$\begin{aligned}
&\int_0^1 \int_{x_1}^1 \int_0^{\frac{x_1}{y_1}} \overbrace{\left(\int_0^{\frac{y_1 x_2}{x_1}} \underbrace{(x_2 y_1 - x_1 y_2)}_{-\det} dy_2 + \int_{\frac{y_1 x_2}{x_1}}^1 \underbrace{(x_1 y_2 - x_2 y_1)}_{+\det} dy_2 \right)}^{\text{for } x_2 \text{ to the left of the vertical line with } x\text{-coordinate } x_1} dx_2\, dy_1\, dx_1 \\
&+ \quad \int_0^1 \int_{x_1}^1 \int_{\frac{x_1}{y_1}}^1 \underbrace{\left(\int_0^1 (x_2 y_1 - x_1 y_2)\, dy_2 \right)}_{\substack{\text{for } x_2 \text{ to the right of the vertical line} \\ \text{with } x\text{-coordinate } x_1;\ -\det}} dx_2\, dy_1\, dx_1.
\end{aligned} \tag{1}$$

The limit of integration $\frac{y_1 x_2}{x_1}$ is the value of y_2 on the line separating the shaded and unshaded regions, the line on which the first dart lands; since that line is given by $y = \frac{y_1}{x_1} x$, we have $y_2 = \frac{y_1}{x_1} x_2$.

We will compute the integral for the half of the square above the diagonal, and then add it to the integral for the bottom half, which by equation 4.5.30

is $\frac{13}{108}$. (Can you think of a way to take advantage of Fubini to make this computation easier?)

Rather surprisingly, when we compute the total integral this way, we encounter some logarithms. The first inner integral in the first line is

$$\int_0^{\frac{y_1x_2}{x_1}} (x_2y_1 - x_1y_2)\,dy_2 = \left[x_2y_1y_2 - x_1\frac{y_2^2}{2}\right]_0^{\frac{y_1x_2}{x_1}} = x_2y_1\frac{y_1x_2}{x_1} - x_1\frac{y_1^2x_2^2}{2x_1^2}$$
$$= \frac{x_2^2y_1^2}{2x_1}.$$

The second inner integral in the first line is

$$\int_{\frac{y_1x_2}{x_1}}^1 (x_1y_2 - x_2y_1)\,dy_2 = \left[x_1\frac{y_2^2}{2} - x_2y_1y_2\right]_{\frac{y_1x_2}{x_1}}^1$$
$$= \frac{x_1}{2} - x_2y_1 - x_1\frac{x_2^2y_1^2}{2x_1^2} + x_2y_1\frac{x_2y_1}{x_1} = \frac{x_1}{2} - x_2y_1 + \frac{y_1^2x_2^2}{2x_1}.$$

So the iterated integral on the first line of equation 1 is

$$\begin{aligned}
&\int_0^1\int_{x_1}^1\int_0^{\frac{x_1}{y_1}}\left(\frac{x_1}{2} - x_2y_1 + \frac{y_1^2x_2^2}{x_1}\right)dx_2\,dy_1\,dx_1 = \int_0^1\int_{x_1}^1\left[\frac{x_1x_2}{2} - \frac{x_2^2y_1}{2} + \frac{x_2^3y_1^2}{3x_1}\right]_0^{\frac{x_1}{y_1}} dy_1\,dx_1\\
&= \int_0^1\int_{x_1}^1\left(\frac{x_1^2}{2y_1} - \frac{x_1^2y_1}{2y_1^2} + \frac{x_1^3y_1^2}{3y_1^3x_1}\right)dy_1\,dx_1 = \int_0^1\left(\int_{x_1}^1\frac{x_1^2}{3y_1}\,dy_1\right)dx_1 = \int_0^1\frac{x_1^2}{3}\Big[\ln y_1\Big]_{x_1}^1 dx_1 \qquad (2)\\
&= \int_0^1 \underbrace{\frac{x_1^2}{3}(-\ln x_1)}_{f'\,g}\,dx_1 = \left[\underbrace{-\frac{x_1^3}{9}\ln x_1}_{f\,g}\right]_0^1 + \int_0^1\underbrace{\left(\frac{x_1^3}{9}\cdot\frac{1}{x_1}\right)}_{f\,g'}dx_1 = \int_0^1\frac{x_1^2}{9}\,dx_1 = \left[\frac{x_1^3}{27}\right]_0^1 = \frac{1}{27}
\end{aligned}$$

Since $\ln 0 = -\infty$, you might worry that in the third line of formula 2, the first term would lead to $(0^3/9)\cdot(-\infty)$, which is undefined. But as $x_1 \to 0$, x_1^3 goes to 0 much faster than $\ln x_1$ goes to $-\infty$.[4]

The iterated integral on the second line of equation 1 is

$$\begin{aligned}
&\int_0^1\int_{x_1}^1\int_{\frac{x_1}{y_1}}^1\left(\int_0^1(x_2y_1 - x_1y_2)\,dy_2\right)dx_2\,dy_1\,dx_1\\
&= \int_0^1\int_{x_1}^1\int_{\frac{x_1}{y_1}}^1\left[x_2y_1y_2 - \frac{x_1y_2^2}{2}\right]_0^1 dx_2\,dy_1\,dx_1\\
&= \int_0^1\int_{x_1}^1\int_{\frac{x_1}{y_1}}^1\left(x_2y_1 - \frac{x_1}{2}\right)dx_2\,dy_1\,dx_1 = \int_0^1\int_{x_1}^1\left[\frac{x_2^2y_1}{2} - \frac{x_1x_2}{2}\right]_{\frac{x_1}{y_1}}^1 dy_1\,dx_1\\
&= \int_0^1\int_{x_1}^1\left(\frac{y_1}{2} - \frac{x_1}{2} - \frac{x_1^2}{y_1^2}\frac{y_1}{2} + \frac{x_1^2}{2y_1}\right)dy_1\,dx_1 = \int_0^1\left[\frac{y_1^2}{4} - \frac{x_1y_1}{2}\right]_{x_1}^1 dx_1\\
&= \int_0^1\left(\frac{1}{4} - \frac{x_1}{2} - \frac{x_1^2}{4} + \frac{x_1^2}{2}\right)dx_1 = \left[\frac{1}{4}x_1 - \frac{x_1^2}{4} + \frac{x_1^3}{12}\right]_0^1 = \frac{1}{4} - \frac{1}{4} + \frac{1}{12} = \frac{1}{12}.
\end{aligned}$$

[4]Powers dominate logarithms, and exponentials dominate powers. This follows almost immediately from l'Hôpital's rule. For powers and exponentials, see Solution 6.10.7.

So the integral for the half above the diagonal is $\frac{1}{27}+\frac{1}{12}=\frac{13}{108}$, which gives a total integral of 13/54.

This computation should convince you that taking advantage of symmetries is worthwhile. But there is a better way to compute the integral for the upper half, as we realized after working through the above computation. The computation is simpler if we integrate with respect to x_2 first, rather than y_2:

Since the line separating the shaded and unshaded regions is given by $y = \frac{y_1}{x_1}x$, we have

$$y_2 = \frac{y_1}{x_1}x_2, \text{ i.e., } x_2 = \frac{x_1 y_2}{y_1}.$$

$$\int_0^1 \int_{x_1}^0 \int_0^1 \left(\int_0^{\frac{x_1 y_2}{y^1}} \underbrace{(x_1 y_2 - x_2 y_1)}_{+\det} dx_2 + \int_{\frac{x_1 y_2}{y^1}}^1 \underbrace{(x_2 y_1 - x_1 y_2)}_{-\det} dx_2 \right) dy_2\, dy_2\, dx_1.$$

4.5.11 a. We have

$$\int_0^\pi \left(\int_y^\pi \frac{\sin x}{x}\, dx \right) dy = \int_T \frac{\sin x}{x}\, |dx\, dy|,$$

where T is the triangle $0 \le y \le x \le \pi$.

b. Written the other way, this becomes

$$\int_0^\pi \left(\int_0^x \frac{\sin x}{x} dy \right) dx = \int_0^\pi x \frac{\sin x}{x}\, dx = \cos 0 - \cos \pi = 2.$$

4.5.13 a. If $D_2(D_1(f))$ and $D_1(D_2(f))$ both exist and are continuous on U, and if there exists $\epsilon > 0$ such that

$$D_2(D_1(f))(\mathbf{a}) - D_1(D_2(f))(\mathbf{a}) > \epsilon, \tag{1}$$

then there exists $\delta > 0$ such that on the square

$$S = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid a_1 \le x \le a_1 + \delta, a_2 \le y \le a_2 + \delta \right\}$$

$D_2(D_1(f)) - D_1(D_2(f)) > \epsilon$. In particular,

$$\int_S D_2(D_1(f)) - D_1(D_2(f))\, |dx\, dy| > \delta^2 \epsilon > 0.$$

Note that this uses the continuity of the second partials.

b. Using Fubini's theorem, these two integrals can be computed:

$$\begin{aligned}
\int_S D_2\big(D_1(f)\big)\, |dx\, dy| &= \int_{a_1}^{a_1+\delta} \left(\int_{a_2}^{a_2+\delta} D_2\big(D_1(f)\big)\, dy \right) dx \\
&= \int_{a_1}^{a_1+\delta} \left(D_1(f) \begin{pmatrix} x \\ a_2+\delta \end{pmatrix} - D_1(f) \begin{pmatrix} x \\ a_2 \end{pmatrix} \right) dx \\
&= f \begin{pmatrix} a_1+\delta \\ a_2+\delta \end{pmatrix} - f \begin{pmatrix} a_1 \\ a_2+\delta \end{pmatrix} - f \begin{pmatrix} a_1+\delta \\ a_2 \end{pmatrix} + f \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.
\end{aligned}$$

$$\int_S D_1\big(D_2(f)\big)\,|dx\,dy| = \int_{a_2}^{a_2+\delta}\left(\int_{a_1}^{a_1+\delta} D_1\big(D_2(f)\big)\,dx\right)dy$$
$$= \int_{a_2}^{a_2+\delta}\left(D_2(f)\begin{pmatrix}a_1+\delta\\ y\end{pmatrix} - D_1(f)\begin{pmatrix}a_1\\ y\end{pmatrix}\right)dy$$
$$= f\begin{pmatrix}a_1+\delta\\ a_2+\delta\end{pmatrix} - f\begin{pmatrix}a_1\\ a_2+\delta\end{pmatrix} - f\begin{pmatrix}a_1+\delta\\ a_2\end{pmatrix} + f\begin{pmatrix}a_1\\ a_2\end{pmatrix}.$$

So the two integrals are equal. Thus if $D_2(D_1(f))$ and $D_1(D_2(f))$ both exist and are continuous on U, then the crossed partials are equal; the second partials cannot be continuous and satisfy inequality (1).

c. The second partials of the function

$$f\begin{pmatrix}x\\ y\end{pmatrix} = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{if } \begin{pmatrix}x\\ y\end{pmatrix} \neq \begin{pmatrix}0\\ 0\end{pmatrix} \\ 0 & \text{otherwise}\end{cases}$$

are not continuous at the origin, so although

$$D_1(D_2(f))(\mathbf{0}) > D_2(D_1(f))(\mathbf{0})$$

(see Example 3.3.9), there is no square on which one crossed partial is larger than the other.

Solution 4.5.15: This region looks roughly like a plastic Easter egg, standing on one end, the bottom half $z \geq x^2+y^2$, the top half $z \leq 10 - x^2 - y^2$; the two halves join at $z = 5$.

4.5.15 Since the two halves are symmetric, we will just compute the bottom half and multiply by 2. The volume is given by the following integral. In the second line we make the change of variables $y = \sqrt{z}\,u$, so that $dy = \sqrt{z}\,du$:

$$2\int_0^5\left(\int_{-\sqrt{z}}^{\sqrt{z}}\left(\int_{-\sqrt{z-y^2}}^{\sqrt{z-y^2}} dx\right)dy\right)dz = 2\int_0^5\left(\int_{-\sqrt{z}}^{\sqrt{z}}\underbrace{\left(\Big[x\Big]_{-\sqrt{z-y^2}}^{\sqrt{z-y^2}}\right)}_{2\sqrt{z-y^2}}dy\right)dz$$

Of course one does not need to use the order $dx\,dy\,dz$.

This exercise would be much easier to do using cylindrical coordinates, discussed in Section 4.10.

$$= 4\int_0^5\left(\int_{-\sqrt{z}}^{\sqrt{z}}\sqrt{z-y^2}\,dy\right)dz = 4\int_0^5\left(\int_{-1}^{1}\sqrt{z-zu^2}\sqrt{z}\,du\right)dz$$
$$= 4\int_0^5 z\left[\frac{u}{2}\sqrt{1-u^2}+\frac{1}{2}\arcsin u\right]_{-1}^{1}dz = 4\frac{\pi}{2}\left[\frac{z^2}{2}\right]_0^5 = 2\pi\frac{25}{2} = 25\pi.$$

4.5.17 a. Rearranging the order of a set of numbers does not change the value of the rth smallest number. Thus $M_r(\mathbf{x})$ satisfies the hypothesis of Exercise 4.3.4, and

$$\int_{Q_{0,1}^n} M_r(\mathbf{x})\,|d^n\mathbf{x}| = n!\int_{P_{0,1}^n} M_r(\mathbf{x})\,|d^n\mathbf{x}|\,.$$

Since $0 \leq x_1 \leq x_2 \cdots \leq x_n \leq 1$ in $P_{0,1}^n$, the rth smallest coordinate in the second integral is x_r, i.e., $M_r(\mathbf{x}) = x_r$. We can integrate with respect to the various x_i in different orders, but the order $x_1, x_2, \dots, x_n$ is convenient. With this order, the integral in question becomes

$$n! \int_0^1 \left(\cdots \left(\int_0^{x_2} x_r \, dx_1 \right) \cdots \right) dx_n.$$

Let us first consider how to do the integral if $r \le n-2$ (and $n > 2$). After integrating with respect to x_i for $1 \le i < r$, the innermost integral is

$$\int_0^{x_{i+2}} \frac{x_r x_{i+1}^i dx_{i+1}}{i!}.$$

(This formula is valid for any $i < r$ and $i \le n-2$ even if $r > n-2$.) After we integrate with respect to x_{r-1}, the M_r function becomes relevant. The innermost integral in this case will be

$$\int_0^{x_{r+1}} \frac{x_r^r \, dx_r}{(r-1)!} = \int_0^{x_{r+1}} \frac{r x_r^r \, dx_r}{r!}.$$

Thus, after integrating with respect to x_i for $r \le i < n-2$, the innermost integral is

$$\int_0^{x_{i+2}} r \frac{x_{i+1}^{i+1} \, dx_{i+1}}{(i+1)!}.$$

It is now simple to evaluate the whole integral. After using the above formula to get the result of the first $n-2$ integrations, our integral is

$$n! \int_0^1 \left(\int_0^{x_n} r \frac{x_{n-1}^{n-1} \, dx_{n-1}}{(n-1)!} \right) dx_n = n! \int_0^1 r \frac{x_n^n \, dx_n}{(n)!} \, dx_n = \frac{r(n!)}{(n+1)!} = \frac{r}{n+1}.$$

If $r = n-1$, then we can use our formula for the innermost integral after $n-2$ integrations with $i < r$:

$$n! \int_0^1 \left(\int_0^{x_n} \frac{x_{n-1}^{n-2} x_{n-1} \, dx_{n-1}}{(n-2)!} \right) dx_n = n! \int_0^1 \left(\int_0^{x_n} (n-1) \frac{x_{n-1}^{n-1} \, dx_{n-1}}{(n-1)!} \right) dx_n$$

$$= n! \int_0^1 (n-1) \frac{x_n^n \, dx_n}{(n)!} = \frac{(n-1)(n!)}{(n+1)!} = \frac{n-1}{n+1} \left(= \frac{r}{n+1} \right).$$

For $r = n$, we can use the same formula:

$$n! \int_0^1 \left(\int_0^{x_n} \frac{x_{n-1}^{n-2} x_n \, dx_{n-1}}{(n-2)!} \right) dx_n = n! \int_0^1 \frac{x_n^{n-1} x_n \, dx_n}{(n-1)!}$$

$$= n! \int_0^1 n \frac{x_n^n \, dx_n}{(n)!} = \frac{(n)(n!)}{(n+1)!} = \frac{n}{n+1} \left(= \frac{r}{n+1} \right).$$

b. The minimum of a set of numbers is the same, regardless of the order of the numbers, so the min function also satisfies the hypothesis of Exercise 4.3.4, and

$$\int_{Q_{0,1}^n} \min \left(1, \frac{b}{x_1}, \dots, \frac{b}{x_n} \right) |d^n \mathbf{x}| = n! \int_{P_{0,1}^n} \min \left(1, \frac{b}{x_1}, \dots, \frac{b}{x_n} \right) |d^n \mathbf{x}| .$$

In $P^n_{0,1}$, $x_n > x_i$ for $i < n$. Thus $b/x_n < b/x_i$ for $i < n$ (since $b > 0$). With this in mind, it is clear that $\min(1, b/x_1, \dots, b/x_n) = \min(1, b/x_n)$ in $P^n_{0,1}$. To evaluate the integral it is convenient to use the same order of integration that we chose in part a:

$$n! \int_0^1 \left(\cdots \left(\int_0^{x_2} \min\left(1, \frac{b}{x_n}\right) dx_1 \right) \cdots \right) dx_n.$$

The expression for the innermost integral after integrating with respect to x_i $(i \le n-2)$ will be nearly the same as that for part a, since the function we are integrating is only a function of x_n (the expression is the same, except that $\min(1, b/x_n)$ replaces x_r):

$$\int_0^{x_{i+2}} \frac{\min\left(1, \frac{b}{x_n}\right) x_{i+1}^i dx_{i+1}}{i!}.$$

It is now simple to evaluate the integral, using the above expression for the innermost integral after $n-2$ integrations. First perform one additional integration:

$$n! \int_0^1 \left(\int_0^{x_n} \min\left(1, \frac{b}{x_n}\right) \frac{x_{n-1}^{n-2} dx_{n-1}}{(n-2)!} \right) dx_n = n! \int_0^1 \min\left(1, \frac{b}{x_n}\right) \frac{x_n^{n-1} dx_n}{(n-1)!}.$$

To integrate the min function, we have to split the above into two integrals. For $0 \le x_n \le b$, we have $1 < b/x_n$ and for $b \le x_n \le 1$, we have $b/x_n < 1$. So

$$\begin{aligned} n! \int_0^1 \min\left(1, \frac{b}{x_n}\right) \frac{x_n^{n-1}\, dx_n}{(n-1)!} &= n! \int_0^b \frac{x_n^{n-1}\, dx_n}{(n-1)!} + n! \int_b^1 \frac{b x_n^{n-2}\, dx_n}{(n-1)!} \\ &= \frac{(n!) b^n}{n!} + \frac{n!}{(n-1)!} \left(\frac{b}{n-1} - \frac{b^n}{n-1} \right) \\ &= \frac{b^n(n-1) + nb - nb^n}{n-1} = \frac{nb - b^n}{n-1}. \end{aligned}$$

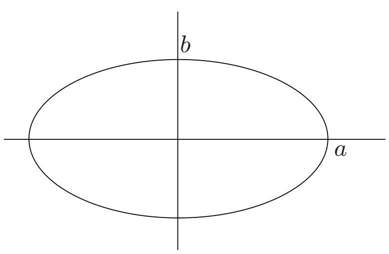


FIGURE FOR SOLUTION 4.5.19
The ellipse
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

In the second line we write $(1 - z^3)$ not $(z^3 - 1)$ because $|a\,b|$ cannot be negative; we can do so, of course, since
$$(1 - z^3)^2 = (z^3 - 1)^2.$$

4.5.19 Use the fact that the area bounded by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ is $\pi|ab|$. This gives

$$\begin{aligned} &\int_{-1}^1 \left(\text{Area bounded by ellipse given by } \frac{x^2}{(z^3-1)^2} + \frac{y^2}{(z^3+1)^2} = 1 \right) dz \\ &= \int_{-1}^1 \pi(1 - z^3)(1 + z^3)\, dz = \pi \int_{-1}^1 (1 - z^6)\, dz \\ &= \pi \left[z - \frac{z^7}{7} \right]_{-1}^1 = \frac{12}{7}\pi. \end{aligned}$$

4.6.1 a. For the unit square, we find

$$\frac{1}{36}\Bigg(f\begin{pmatrix}0\\0\end{pmatrix}+f\begin{pmatrix}0\\1\end{pmatrix}+f\begin{pmatrix}1\\0\end{pmatrix}+f\begin{pmatrix}1\\1\end{pmatrix}$$
$$+4\bigg(f\begin{pmatrix}1/2\\0\end{pmatrix}+f\begin{pmatrix}1/2\\1\end{pmatrix}+f\begin{pmatrix}0\\1/2\end{pmatrix}+f\begin{pmatrix}1\\1/2\end{pmatrix}\bigg)+16f\begin{pmatrix}1/2\\1/2\end{pmatrix}\Bigg).$$

For the unit cube, we find

$$\frac{1}{216}\left(f\begin{pmatrix}0\\0\\0\end{pmatrix}+f\begin{pmatrix}1\\0\\0\end{pmatrix}+f\begin{pmatrix}0\\1\\0\end{pmatrix}+f\begin{pmatrix}1\\1\\0\end{pmatrix}+f\begin{pmatrix}0\\0\\1\end{pmatrix}+f\begin{pmatrix}1\\0\\1\end{pmatrix}+f\begin{pmatrix}0\\1\\1\end{pmatrix}x+f\begin{pmatrix}1\\1\\1\end{pmatrix}\right.$$
$$+4\left(f\begin{pmatrix}1/2\\0\\0\end{pmatrix}+f\begin{pmatrix}1/2\\1\\0\end{pmatrix}+f\begin{pmatrix}1/2\\0\\1\end{pmatrix}+f\begin{pmatrix}1/2\\1\\1\end{pmatrix}+f\begin{pmatrix}0\\1/2\\0\end{pmatrix}+f\begin{pmatrix}1\\1/2\\0\end{pmatrix}\right.$$
$$\left.+f\begin{pmatrix}0\\1/2\\1\end{pmatrix}+f\begin{pmatrix}1\\1/2\\1\end{pmatrix}+f\begin{pmatrix}0\\0\\1/2\end{pmatrix}+f\begin{pmatrix}1\\0\\1/2\end{pmatrix}+f\begin{pmatrix}0\\1\\1/2\end{pmatrix}+f\begin{pmatrix}1\\1\\1/2\end{pmatrix}\right)$$
$$+16\left(f\begin{pmatrix}0\\1/2\\1/2\end{pmatrix}+f\begin{pmatrix}1\\1/2\\1/2\end{pmatrix}+f\begin{pmatrix}1/2\\0\\1/2\end{pmatrix}+f\begin{pmatrix}1/2\\1\\1/2\end{pmatrix}+f\begin{pmatrix}1/2\\1/2\\0\end{pmatrix}+f\begin{pmatrix}1/2\\1/2\\1\end{pmatrix}\right)$$
$$\left.+64f\begin{pmatrix}1/2\\1/2\\1/2\end{pmatrix}\right).$$

b. The formula for the unit square gives

$$\frac{1}{36}\left(\Big(1+\frac{1}{2}+\frac{1}{2}+\frac{1}{3}\Big)+4\Big(\frac{2}{3}+\frac{2}{5}+\frac{2}{3}+\frac{2}{5}\Big)+16\frac{1}{2}\right)=\frac{1}{36}\left(2+\frac{1}{3}+\frac{128}{15}+\frac{16}{2}\right)\approx .524074.$$

The exact value of the integral is

$$\int_{\text{unit square}} f\begin{pmatrix}x\\y\end{pmatrix}\,|dx\,dy| = 3\ln 3-4\ln 2\approx .523848.$$

4.6.3 If we set

$$X=\frac{b-a}{2}x+\frac{a+b}{2},\quad dX=\frac{b-a}{2}\,dx$$

and make the corresponding change of variables, we find

$$\int_a^b f(X)\,dX=\int_{-1}^1 f\left(\frac{b-a}{2}x+\frac{a+b}{2}\right)\frac{b-a}{2}\,dx.$$

This integral is approximated by

$$\sum_{i=1}^p w_i f\left(\frac{b-a}{2}x_i+\frac{a+b}{2}\right)\frac{b-a}{2},$$

so to make the "same approximation" for $\int_a^b f(X)\,dX$, we need to use

$$X_i=\frac{b-a}{2}x_i+\frac{a+b}{2}\quad\text{and}\quad W_i=\frac{b-a}{2}w_i.$$

4.6.5 a. By Fubini's theorem,

$$\begin{aligned}\int_{[a,b]\times[a,b]} f(x)g(y)\,|dx\,dy| &= \int_{[a,b]}\Big(\int_{[a,b]} f(x)g(y)\,|dx|\Big)|dy|\\ &= \left(\int_{[a,b]} g(y)\,|dy|\right)\left(\int_{[a,b]} f(x)\,|dx|\right)\\ &= \left(\sum_{i=1}^{n} c_i f(x_i)\right)\left(\sum_{i=1}^{n} c_i g(x_i)\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j f(x_i)g(x_j).\end{aligned}$$

b. Since both quadratic and cubic polynomials are integrated exactly by Simpson's rule, this function is integrated exactly, giving

$$\left[\frac{x^3}{3}\right]_0^1 \left[\frac{y^4}{4}\right]_0^1 = \frac{1}{12}.$$

4.6.7 a. In the expression $E_{\mathbf{T}}(E_T(f^2))$, we have summed $f^2(s)$ once for every T with $s \in T$. So every $s \in S$ has been sampled the same (huge) number of times, and in the expectation we have divided by this huge number, to find the average of f^2 over S, i.e., $E_S(f^2)$.

We already used equation 1 in equation 3.8.19.

For the next two, note that $\overline{f}_T$ is a constant on T, so $E_T(\overline{f}_T^2) = \overline{f}_T^2$. By the same argument,

$$E_T(f\overline{f}_T) = \overline{f}_T E_T(f) = \overline{f}_T^2. \tag{1}$$

b.

$$\begin{aligned}&\frac{N}{N-1}E_{\mathbf{T}}\Big(E_T\big((f-\overline{f}_T)^2\big)\Big)\\ &\qquad= \frac{N}{N-1}E_{\mathbf{T}}\Big(E_T(f^2) - E_T(2f\overline{f}_T) + E_T(\overline{f}_T^2)\Big)\\ &\qquad= \frac{N}{N-1}E_{\mathbf{T}}\Big(E_T(f^2) - \overline{f}_T^2\Big)\\ &\qquad= \frac{N}{N-1}\Big(E_S(f^2) - E_{\mathbf{T}}(\overline{f}_T^2)\Big)\end{aligned} \tag{2}$$

c. Continuing equation 2, we have

$$\begin{aligned}&\frac{N}{N-1}\Big(E_S(f^2) - E_T(\overline{f}_T^2)\Big)\\ &\qquad= \frac{N}{N-1}\bigg(\Big(\mathrm{Var}_S(f) + \big(E_S(f)\big)^2\Big) - \Big(\mathrm{Var}_{\mathbf{T}}(\overline{f}_T) + \big(E_{\mathbf{T}}(\overline{f}_T)\big)^2\Big)\bigg).\end{aligned}$$

But $\big(E_S(f)\big)^2 = \big(E_{\mathbf{T}}(\overline{f}_T)\big)^2$ so we are left with

$$\frac{N}{N-1}\Big(\mathrm{Var}_S(f) - \mathrm{Var}_{\mathbf{T}}(\overline{f}_T)\Big).$$

d. The crucial issue is that

$$\mathrm{Var}_{\mathbf{T}}\left(\overline{f}_T\right) = \frac{1}{N}\,\mathrm{Var}_S(f). \tag{3}$$

To see this, we need to develop out:

The notation $|\mathbf{T}|$ means the number of elements of $\mathbf{T}$.

$$\begin{aligned}\mathrm{Var}_{\mathbf{T}}(\overline{f}_T) &= \frac{1}{|\mathbf{T}|}\sum_{T\in\mathbf{T}}\Big(\overline{f}_T - E_S(f)\Big)^2 = \frac{1}{|\mathbf{T}|}\sum_{T\in\mathbf{T}}\frac{1}{N}\left(\sum_{t\in T}\Big(f(t)-E_S(f)\Big)\right)^2\\ &= \frac{1}{N}\frac{1}{|S|}\sum_{s\in S}\Big(f(s)-E_S(f)\Big)^2 = \frac{1}{N}\,\mathrm{Var}_S(f)\end{aligned}$$

Using equation (3), we have

$$\frac{N}{N-1}\left(\mathrm{Var}_S(f) - \frac{1}{N}\,\mathrm{Var}_S(f)\right) = \frac{N}{N-1}\,\frac{N-1}{N}\,\mathrm{Var}_S(f) = \mathrm{Var}_S(f).$$

4.7.1 a. Theorem 4.7.4 asserts that you can compute integrals by partitioning the plane into squares with vertices at the points $\begin{pmatrix} n/N \\ m/N\end{pmatrix}$ and sidelength $1/N$. The expression

Note that one N goes with the m to make up the y; the other two give the area of the square.

$$\lim_{N\to\infty}\frac{1}{N^3}\sum_{0\le n,m<N} me^{-nm/N^2} = \frac{1}{N^2}\sum\sum\frac{m}{N}e^{-nm/N^2}$$

is the left Riemann sum for the integral

$$\int_0^1\int_0^1 ye^{-xy}\,dx\,dy.$$

This comes down to computing the integral of a function on $\mathbb{R}^2$ that is continuous except on the boundary of the unit square; that boundary has 2-dimensional volume 0, so the function is integrable, and the limit exists.

Solution 4.7.1, part b: This is one of those cases where if you apply Fubini one way, the integral can be evaluated in elementary terms, and in the other it can't.

b. The direction in which the integral can be evaluated in elementary terms is

$$\begin{aligned}\int_0^1\left(\int_0^1 ye^{-xy}\,dx\right)dy &= \int_0^1\left(\left[-\frac{y}{y}e^{-xy}\right]_0^1\right)dy\\ &= \int_0^1 (1-e^{-y})\,dy = 1 + e^{-1} - 1 = \frac{1}{e}.\end{aligned}$$

4.8.1

Solution 4.8.1: It follows from Theorem 4.8.8 that computing the determinant using development by the first row is equivalent to computing it using the first column. For the matrix A, using the first row is easier, because of the 0 in the first row. For B and C, using the first column is easier.

$$\begin{aligned}\det A &= \det\begin{bmatrix} 0 & 1 & 2\\ -1 & 2 & 1\\ 2 & 1 & 0\end{bmatrix} + 2\det\begin{bmatrix} 4 & 1 & 2\\ 5 & 2 & 1\\ 3 & 1 & 0\end{bmatrix} + 3\det\begin{bmatrix} 4 & 0 & 2\\ 5 & -1 & 1\\ 3 & 2 & 0\end{bmatrix}\\ &= -8 - 6 + 54 = 40.\end{aligned}$$

$\det B = -5; \qquad \det C = 91$

4.8.3 1. Multilinearity: Set $\begin{bmatrix} b\\ d\end{bmatrix} = \alpha\begin{bmatrix} e\\ f\end{bmatrix} + \beta\begin{bmatrix} g\\ h\end{bmatrix}$. Then

$$\det\begin{bmatrix} a & b\\ c & d\end{bmatrix} = \det\begin{bmatrix} a & \alpha e+\beta g\\ c & \alpha f + \beta h\end{bmatrix} = a\alpha f + a\beta h - c\alpha e - c\beta g,$$

which is identical to

$$\alpha \det \begin{bmatrix} a & e \\ c & f \end{bmatrix} + \beta \det \begin{bmatrix} a & g \\ c & h \end{bmatrix} = \alpha af - \alpha ce + \beta ah - \beta cg.$$

2. Antisymmetry: $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ and $\det \begin{bmatrix} b & a \\ d & c \end{bmatrix} = bc - ad$, so

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = -\det \begin{bmatrix} b & a \\ d & c \end{bmatrix}.$$

3. Normalization: $\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1.$

4.8.5 The trace is the sum of the elements on the diagonal. Set $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$. Then

$$AB = \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix} \quad \text{and} \quad \operatorname{tr}(AB) = ae+bg+cf+dh;$$

$$BA = \begin{bmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{bmatrix} \quad \text{and} \quad \operatorname{tr}(BA) = ea+fc+gb+hd.$$

4.8.7 a. Suppose the column of 0's is the ith column. The hint suggested multiplying that column by 2 or -4. Instead, let's multiply by 0. Then by multilinearity,

$$\det[\vec{\mathbf{a}}_1, \dots, \vec{\mathbf{0}}, \dots, \vec{\mathbf{a}}_n] = \det[\vec{\mathbf{a}}_1, \dots, 0 \cdot \vec{\mathbf{0}}, \dots, \vec{\mathbf{a}}_n] = 0 \det[\vec{\mathbf{a}}_1, \dots, \vec{\mathbf{0}}, \dots, \vec{\mathbf{a}}_n] = 0.$$

b. By antisymmetry, exchanging two columns changes the sign of the determinant, but exchanging two identical columns leaves it unchanged, therefore the determinant must be 0.

4.8.9 a. i. $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$; ii. $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$; iii. $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

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

4.8.11 Use development of the first column (equation 4.8.5), and work by induction on n, the dimension of A. It is true when $n = 1$, i.e., when $A = [a]$, since then $\det \begin{bmatrix} a & C \\ 0 & B \end{bmatrix} = a \det B$.

Now suppose it is true for $n-1$, and that A is an $n \times n$ matrix. Then

$$\det \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det \begin{bmatrix} A_{[i,1]} & C \\ 0 & B \end{bmatrix} = \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det A_{[i,1]} \det B$$

$$= \det B \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det A_{[i,1]} = \det A \det B.$$

4.8.13

$$\begin{aligned}
\operatorname{sgn}\sigma_1 &= \det M_{\sigma_1} = \det I = +1\\
\operatorname{sgn}\sigma_2 &= \det M_{\sigma_2} = \det\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix} = +1\\
\operatorname{sgn}\sigma_3 &= \det M_{\sigma_3} = \det\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix} = +1\\
\operatorname{sgn}\sigma_4 &= \det M_{\sigma_4} = \det\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix} = -1\\
\operatorname{sgn}\sigma_5 &= \det M_{\sigma_5} = \det\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix} = -1\\
\operatorname{sgn}\sigma_6 &= \det M_{\sigma_6} = \det\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix} = -1.
\end{aligned}$$

Part a of Solution 4.8.15: We do not need to complete the operations to echelon form because any remaining operations add a multiple of one column to another column, which has $\mu = 1$.

4.8.15 a. We use row operations; column operations would work as well. Each row operation (or column operation) is equivalent to multiplying the determinant by a factor μ; we mark only those μ that are $\neq 1$:

$$\begin{bmatrix}2&1&0&1\\1&1&3&2\\2&0&2&1\\1&0&4&2\end{bmatrix}\underbrace{\to}_{\mu_1=1/2}\begin{bmatrix}1&1/2&0&1/2\\1&1&3&2\\2&0&2&1\\1&0&4&2\end{bmatrix}\to\begin{bmatrix}1&1/2&0&1/2\\0&1/2&3&3/2\\0&-1&2&0\\0&-1/2&4&3/2\end{bmatrix}\underbrace{\to}_{\mu_2=2}\begin{bmatrix}1&1/2&0&1/2\\0&1&6&3\\0&-1&2&0\\0&-1/2&4&3/2\end{bmatrix}$$

$$\to\begin{bmatrix}1&0&-3&-1\\0&1&6&3\\0&0&8&3\\0&0&7&3\end{bmatrix}\underbrace{\to}_{\mu_3=1/8}\begin{bmatrix}1&0&-3&-1\\0&1&6&3\\0&0&1&3/8\\0&0&7&3\end{bmatrix}\to\begin{bmatrix}1&0&0&1/8\\0&1&0&1\\0&0&1&3/8\\0&0&0&3/8\end{bmatrix}\underbrace{\to}_{\mu_4=8/3}\begin{bmatrix}1&0&0&1/8\\0&1&0&1\\0&0&1&3/8\\0&0&0&1\end{bmatrix}.$$

Thus the determinant is $\dfrac{1}{\mu_1\mu_2\mu_3\mu_4} = 3$.

b. The numbers $1, 2, 3, 4$ have 24 permutations, but only 10 contribute nonzero terms to the sum

$$\det A = \sum_{\sigma\in\operatorname{Perm}(1,\dots,n)} \operatorname{sgn}(\sigma)a_{1,\sigma(1)}\dots a_{n,\sigma(n)}. \quad \text{of equation 4.8.43:}$$

$$\begin{aligned}
&\sigma_1 = (1234) \to +a_{1,1}a_{2,2}a_{3,3}a_{4,4} = +2\cdot1\cdot2\cdot2 = +8;\\
&\sigma_2 = (1243) \to -a_{1,1}a_{2,2}a_{3,4}a_{4,3} = -2\cdot1\cdot1\cdot4 = -8;\\
&\sigma_3 = (2134) \to -1\cdot1\cdot2\cdot2 = -4; \quad \sigma_4 = (2143) \to 1\cdot1\cdot1\cdot4 = +4\\
&\sigma_5 = (2341) \to -1\cdot3\cdot1\cdot1 = -3; \quad \sigma_6 = (2314) \to 1\cdot3\cdot2\cdot2 = +12\\
&\sigma_7 = (2413) \to -1\cdot2\cdot2\cdot4 = -16 \quad \sigma_8 = (2431) \to +1\cdot2\cdot2\cdot1 = +4\\
&\sigma_9 = (4231) \to -1\cdot1\cdot2\cdot1 = -2; \qquad \sigma_{10} = (4213) \to +1\cdot1\cdot2\cdot4 = +8.
\end{aligned}$$

So

$$\det\begin{bmatrix}2&1&0&1\\1&1&3&2\\2&0&2&1\\1&0&4&2\end{bmatrix} = +8-8-4+4-3+12-16+4-2+8=3.$$

4.8.17 Set $A=\begin{bmatrix}a&b\\c&d\end{bmatrix}$, so $A^\top A=\begin{bmatrix}a^2+c^2&ab+cd\\ab+cd&b^2+d^2\end{bmatrix}$. We need to find the maximal eigenvalue λ of $A^\top A$. To find the eigenvalues, we use the characteristic polynomial:

$$\begin{aligned}\chi_{A^\top A}(\lambda) &= \det(\lambda I - A^\top A) = \det\begin{bmatrix}\lambda-a^2-c^2 & -ab-cd\\ -ab-cd & \lambda-b^2-d^2\end{bmatrix}\\ &= (\lambda-a^2-c^2)(\lambda-b^2-d^2)-(-ab-cd)(-ab-cd)\\ &= \lambda^2-\lambda(\underbrace{a^2+b^2+c^2+d^2}_{|A|^2})+a^2b^2+a^2d^2+b^2c^2+c^2d^2\\ &\qquad -(a^2b^2+2abcd+c^2d^2)\\ &= \lambda^2-|A|^2-(ad-bc)^2=\lambda^2-|A|^2-(\det A)^2.\end{aligned}$$

So using the quadratic formula to solve

$$\lambda^2-|A|^2-(ad-bc)^2=\lambda^2-|A|^2-(\det A)^2=0,$$

we get

$$\lambda=\frac{|A|^2\pm\sqrt{|A|^4-4(\det A)^2}}{2}$$

To get the larger eigenvalue, take the positive square root.

Solution 4.8.19: Remember that the determinant of a matrix is a number, so multiplication of determinants is commutative.

We have

$$\begin{aligned}&\det(QJ_kP^{-1})\\ &\quad=\det Q\det J_k\det P^{-1}\\ &\quad=0\end{aligned}$$

since $\det J_k=0$; see (for instance) Theorem 4.8.9.

4.8.19 Using Exercise 2.39, the the problem is not too hard.

a. Assume A has rank $k=n-1$. We want to show that the linear transformation $[\mathbf{D}\det(A)]:\mathrm{Mat}\,(n,n)\to\mathbb{R}$ is not the zero linear transformation, i.e., that there is a matrix B such that $[\mathbf{D}\det(A)]B\neq 0$.

Write $[\mathbf{D}\det(A)]B$ as follows, where P, Q, J_k are as in Exercise 2.39:

$$\begin{aligned}[\mathbf{D}\det(A)]B &= [\mathbf{D}\det(QJ_kP^{-1})]B=\lim_{h\to 0}\frac{\det\big(QJ_kP^{-1}+hB\big)-\overbrace{\det\big(QJ_kP^{-1}\big)}^{0}}{h}\\ &=\lim_{h\to0}\frac{\det(QJ_kP^{-1}+hQQ^{-1}BPP^{-1})}{h}=\lim_{h\to0}\frac{\det\Big(Q(J_k+hQ^{-1}BP)P^{-1}\Big)}{h}\\ &=\lim_{h\to0}\frac{\det Q\det P^{-1}\det(J_k+hQ^{-1}BP)}{h}\\ &=\det Q\det P^{-1}\lim_{h\to0}\frac{\det(J_k+hQ^{-1}BP)}{h}.\end{aligned}$$

Since Q and P^{-1} are invertible, we know that $\det Q\neq0$ and $\det P^{-1}\neq0$, so we just need to show that there exists a matrix B satisfying

$$\det(J_k+hQ^{-1}BP)\neq0.$$

Set $K = Q^{-1}BP$, so that now we need to show that $\det(J_k + hK) \neq 0$, and let $k_{i,j}$ denote the i,jth entry of K, so that

$$\det(J_k + hQ^{-1}BP) = \det(J_k + hK) = \det \begin{bmatrix} hk_{1,1} & hk_{1,2} & \dots & hk_{1,n} \\ hk_{2,1} & 1 + hk_{2,2} & \dots & hk_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ hk_{n,1} & hk_{n,2} & \dots & 1 + hk_{n,n} \end{bmatrix},$$

We are assuming $k = n - 1$, so the first column and first row of J_k consist of 0's, so the first column and first row of $J_k + hK$ equal the first column and first row of hK.

Then (as you can easily check in the case $n = 3$ by working out an example), when we compute the determinant, there is a single term, the term $hk_{1,1}$, that is of degree 1 in h; all the other terms have at least h^2, and thus disappear after dividing by h and taking the limit as $h \to 0$. This yields

$$\lim_{h \to 0} \frac{1}{h} \det(J_k + hQ^{-1}BP) = k_{1,1}.$$

Thus if B is any $n \times n$ matrix such that the $(1,1)$ entry of $K = Q^{-1}BP$ is not 0, then

$$[\mathbf{D} \det(A)]B \neq 0.$$

Clearly a matrix K' exists with $k'_{1,1} \neq 0$; setting $B = QK'P^{-1}$ gives $Q^{-1}BP = Q^{-1}QK'P^{-1}P = K'$.

b. In this case, entries of the first two columns and first two rows of $J_k + hK$ have an h. For instance, if $n = 4$ and $k = n - 2 = 2$, we have

Solution 4.8.19, part b: You can compute

$$\det(J_k + hK) = \det(J_k + hQ^{-1}BP)$$

using development by the first column or Theorem 4.8.12, which says that you must choose in each term one element from each row and each column. In the first $n-k$ columns, every entry in the matrix has a factor of h, so each term of the determinant has a factor of h^2.

$$\det(J_k + hK) = \det \begin{bmatrix} hk_{1,1} & hk_{1,2} & hk_{1,3} & hk_{1,4} \\ hk_{2,1} & hk_{2,2} & hk_{2,3} & hk_{2,4} \\ hk_{3,1} & hk_{3,2} & 1 + hk_{3,3} & hk_{3,4} \\ hk_{4,1} & hk_{4,2} & hk_{4,3} & 1 + hk_{4,4} \end{bmatrix}$$

When you compute this, the lowest degree term in the expansion will have a factor h^{n-k} (for the case above, h^2), giving

$$\lim_{h \to 0} \frac{1}{h} \det(J_k + hQ^{-1}BP) = 0$$

Thus if $\operatorname{rank} A \leq n - 2$, we have $[\mathbf{D} \det(A)]B = 0$ for all $B \in \operatorname{Mat}(n, n)$.

4.8.21 By definition, $M_\sigma(\vec{\mathbf{e}}_i) = \vec{\mathbf{e}}_{\sigma(i)}$. Thus

$$M_\tau M_\sigma(\vec{\mathbf{e}}_i) = M_\tau(\vec{\mathbf{e}}_{\sigma(i)}) = \vec{\mathbf{e}}_{\tau(\sigma(i))} = \vec{\mathbf{e}}_{(\tau \circ \sigma)(i)},$$

whereas

$$M_{\tau \circ \sigma}(\vec{\mathbf{e}}_i) = \vec{\mathbf{e}}_{(\tau \circ \sigma)(i)}.$$

Thus the linear transformations $M_\tau M_\sigma$ and $M_{\tau \circ \sigma}$ do the same things to the standard basis vectors (alternatively, the matrices have the same columns), so they coincide.

4.8.23 a. If A is orthogonal,

$$1 = \det I = \det A^\top A = (\det A^\top)(\det A) = (\det A)^2,$$

so $\det A = \pm 1$.

Solution 4.8.23, part b: If we had introduced dot products in $\mathbb{C}^n$, and had shown that orthogonal matrices are special cases of unitary matrices, the proof would be easier.

b. If λ is a real eigenvalue with eigenvector $\vec{\mathbf{v}} \in \mathbb{R}^n$, then, since A is orthogonal and preserves lengths,

$$|\vec{\mathbf{v}}| = |A\vec{\mathbf{v}}| = |\lambda||\vec{\mathbf{v}}|, \quad \text{so } |\lambda| = 1.$$

If $\lambda = a + ib \notin \mathbb{R}$, we saw in Exercise 2.14 that $A\mathbf{u} = a\mathbf{u} - b\mathbf{w}$ and $A\mathbf{w} = b\mathbf{u} + a\mathbf{w}$. Since A preserves all lengths,

$$\begin{aligned} |\mathbf{u}|^2 + |\mathbf{w}|^2 &= |A\mathbf{u}|^2 + |A\mathbf{w}|^2 \\ &= a^2|\mathbf{u}|^2 + b^2|\mathbf{w}|^2 - 2ab(\mathbf{u}\cdot\mathbf{w}) + b^2|\mathbf{u}|^2 + a^2|\mathbf{w}|^2 + 2ab(\mathbf{u}\cdot\mathbf{w}) \\ &= (a^2+b^2)(|\mathbf{u}|^2 + |\mathbf{w}|^2). \end{aligned}$$

So we get $|\lambda|^2 = (a^2+b^2) = 1$, so indeed $|\lambda| = 1$.

c. The eigenvalues of A are $+1$, -1, and nonreal pairs $e^{\pm i\theta}$. Since (Corollary 4.8.25) the determinant is the product of the eigenvalues with multiplicity, and since each pair of complex conjugate eigenvalues contributes $+1$ as do the eigenvalues $+1$, of course, we see that the only way that the determinant can be -1 is if -1 occurs as an eigenvalue with odd multiplicity.

d. We will prove a stronger result: that if an $n \times n$ matrix A is orthogonal, there exists an orthonormal basis $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ of $\mathbb{R}^n$ such that if $T = [\vec{\mathbf{v}}_1 \ \dots \vec{\mathbf{v}}_n]$, and we set $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, then we have:

The matrix $\begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix}$ gives rotation by θ_i in the plane spanned by $\mathbf{v}_{2i}$ and $\mathbf{v}_{2i+1}$; see Example 1.3.9. The entry -1 is reflection with respect to some hyperplane.

$$T^{-1}AT = \begin{bmatrix} \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \begin{bmatrix} c_k & s_k \\ -s_k & c_k \end{bmatrix} & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & -1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & -1 \end{bmatrix}$$

The matrix $T^{-1}AT$: In the subspace spanned by the eigenvectors with eigenvalue 1, A is the identity. In the subspace spanned by the eigenvectors with eigenvalue -1, A is minus the identity, i.e., a reflection with respect to the origin. In the planes spanned by $\vec{\mathbf{v}}_{2i-1}, \vec{\mathbf{v}}_{2i}$ A is rotation by some angle.

(all the blank entries are 0).

Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, and $V \subset \mathbb{R}^n$. We will say that A is *orthogonal on* V if $A(V) \subset V$ and A preserves all dot products of elements of V.

Work by induction on the dimension of a subspace of $\mathbb{R}^n$ on which A is orthogonal. Thus suppose that any subspace $V \subset \mathbb{R}^n$ with $AV \subset V$ of dimension $< m$ has an orthonormal basis in which $A : V \to V$ has a matrix as above, and let us prove it for a subspace V of dimension m.

By the fundamental theorem of algebra, every linear transformation has an eigenvalue (perhaps complex). Let $\vec{\mathbf{v}} \in \mathbb{C}^n$ be an eigenvector of $A : V \to V$, i.e., $A\vec{\mathbf{v}} = \lambda\vec{\mathbf{v}}$. Suppose first that λ is real, i.e., $\lambda = \pm 1$. Then $A(\vec{\mathbf{v}}^\perp) = \vec{\mathbf{v}}^\perp$: if $\vec{\mathbf{w}} \cdot \vec{\mathbf{v}} = 0$, then

$$0 = A\vec{\mathbf{w}} \cdot A\vec{\mathbf{v}} = \pm A\vec{\mathbf{w}} \cdot \vec{\mathbf{v}}.$$

But $\vec{\mathbf{v}}^{\perp}$ has dimension $n-1$.

If $\lambda = a + ib = e^{i\theta}$ with $b \neq 0$, we can write $\vec{\mathbf{v}} = \vec{\mathbf{u}} + i\vec{\mathbf{w}}$ with $\vec{\mathbf{u}}, \vec{\mathbf{w}} \in \mathbb{R}^n$. We need to show three things:

1. that $\vec{\mathbf{u}}$ and $\vec{\mathbf{w}}$ are orthogonal
2. that they have the same length, so they can be rescaled to be unit vectors forming an orthonormal basis of the plane P that they span.
3. if a third vector $\vec{\mathbf{z}}$ is orthogonal to the plane P, then $A\vec{\mathbf{z}}$ is also orthogonal to P.

For the first, since A is orthogonal it preserves all dot products, so

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{w}} = A\vec{\mathbf{u}} \cdot A\vec{\mathbf{w}} = (a\vec{\mathbf{u}} - b\vec{\mathbf{w}}) \cdot (b\vec{\mathbf{u}} + a\vec{\mathbf{w}} = (a^2 - b^2)(\vec{\mathbf{u}} \cdot \vec{\mathbf{w}}),$$

and by subtracting

$$(a^2 - b^2)(\vec{\mathbf{u}} \cdot \vec{\mathbf{w}}) = \vec{\mathbf{u}} \cdot \vec{\mathbf{w}} \quad \text{from} \quad (a^2 + b^2)(\vec{\mathbf{u}} \cdot \vec{\mathbf{w}}) = \vec{\mathbf{u}} \cdot \vec{\mathbf{w}}$$

we get $b^2(\vec{\mathbf{u}} \cdot \vec{\mathbf{w}}) = 0$. Since $b \neq 0$ we see that $\vec{\mathbf{u}} \cdot \vec{\mathbf{w}} = 0$.

For the second,

$$|\vec{\mathbf{u}}|^2 = |A\vec{\mathbf{u}}|^2 = |a\vec{\mathbf{u}} - b\vec{\mathbf{w}}|^2 = a^2|\vec{\mathbf{u}}|^2 + b^2|\vec{\mathbf{w}}|^2,$$

so $(1-a^2)|\vec{\mathbf{u}}|^2 = b^2|\vec{\mathbf{u}}|^2 = b^2|\vec{\mathbf{w}}|^2$, and since $b \neq 0$ we have $|\vec{\mathbf{u}}| = |\vec{\mathbf{w}}|$. Thus the vectors $\vec{\mathbf{u}}$ and $\vec{\mathbf{w}}$ have the same length and are orthogonal to each other, so they can be scaled to be an orthonormal basis of P. In that basis the matrix of $A : P \to P$ is

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad \text{if } \lambda = e^{i\theta}.$$

Again $P^{\perp}$ is mapped to itself by A, and all dot products of $P^{\perp}$ are preserved, and again P has lower dimension.

4.9.1 It doesn't matter what A is, so long as it has positive and finite volume: since T has a triangular matrix,

$$\frac{\mathrm{vol}_n(T(A))}{\mathrm{vol}_n(A)} = |\det T| = n!,$$

by Theorem 4.8.9.

4.9.3 a. By Fubini's theorem, this volume is

$$\begin{aligned} \int_0^1 \int_0^{1-z} \int_0^{1-z-y} dx\,dy\,dz &= \int_0^1 \int_0^{1-z} 1 - z - y\,dy\,dz \\ &= \int_0^1 \left[(1-z)y - \frac{y^2}{2}\right]_0^{1-z} dz \\ &= \int_0^1 \frac{1}{2}(1-z)^2 dz = \left[-\frac{1}{6}(1-z)^3\right]_0^1 = \frac{1}{6} \end{aligned}$$

Solution 4.9.3, part b: By "maps the tetrahedron T_1 onto the tetrahedron T_2" we mean that S is an onto map (in fact, it is bijective) that takes a point in T_1 and gives a point in T_2; for every point $\mathbf{y} \in T_2$ there is a point $\mathbf{x} \in T_1$ such that $S\mathbf{x} = \mathbf{y}$.

b. The matrix $S = \begin{bmatrix} 2 & -1 & -2 \\ 1 & 3 & -5 \\ 1 & 1 & 2 \end{bmatrix}$ maps the tetrahedron T_1 onto the tetrahedron T_2. Thus

$$\operatorname{vol} T_2 = |\det S| \operatorname{vol} T_1 = \frac{33}{6} = \frac{11}{2}.$$

4.9.5 The function A is given by

$$A(x) = \det \begin{bmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{bmatrix} = f(x)g'(x) - f'(x)g(x).$$

So its derivative is given by

$$\begin{aligned} A'(x) &= f'(x)g'(x) + f(x)g''(x) - f'(x)g'(x) - f''(x)g(x) \\ &= q(x)\big(f(x)g(x) - f(x)g(x)\big) = 0. \end{aligned}$$

Here we use $(ab)' = ab' + ba'$.

So the function is constant: $A(x) = A(0)$.

4.9.7 a. First let us see why there exists such a function and why it is unique. Existence is easy (since we have Theorem and Definition 4.8.1): the function $|\det T|$ satisfies properties 1–5.

The proof of uniqueness is essentially the same as the proof of uniqueness for the determinant: to compute $\widetilde{\Delta}(T)$ from properties 1–5, column reduce T. Clearly operations of type 2 and 3 do not change the value of $\widetilde{\Delta}$. Operations of type 3 switch two columns, which by property 2 does not change the value of $\widetilde{\Delta}$. Operations of type 2 add a multiple of one column onto another, which by property 4 does not change the value either. By property 3, an operation of type 1 multiplies $\widetilde{\Delta}(T)$ by $|\mu|$, the factor you multiply a column by. So keep track, for each column operation of type 1, of the numbers $\mu_1, \dots, \mu_k$ that you multiply columns by.

At the end we see that

$$\widetilde{\Delta}(T) = \begin{cases} 0 & \text{if } T \text{ does not column reduce to the identity} \\ \frac{1}{|\mu_1 \cdots \mu_k|} = \frac{1}{|\mu_1 \cdots \mu_k|}\widetilde{\Delta}(I) & \text{if } T \text{ column reduces to the identity.} \end{cases}$$

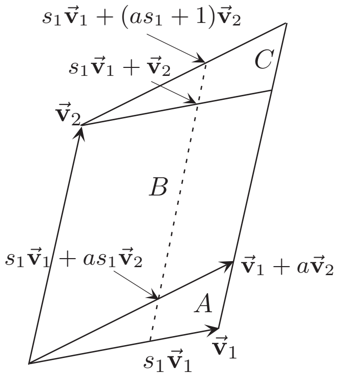
By property 5,

$$\frac{1}{|\mu_1 \cdots \mu_k|}\widetilde{\Delta}(I) = \frac{1}{|\mu_1 \cdots \mu_k|}.$$

b. Now we show that the mapping $T \mapsto \operatorname{vol}_n(T(Q))$ satisfies properties 1–5. Properties 1, 2, and 5 are straightforward, but 3 and 4 are quite delicate.

Property 1: It is not actually necessary to prove that $\operatorname{vol}_n(T(Q)) \geq 0$, but it follows immediately from Definition 4.1.17 of vol_n and from the fact (Definition 4.1.1) that the indicator function is never negative.

Property 2: Since the mapping $Q \mapsto T(Q)$ consists of taking each point in Q, considering it as a vector, and multiplying it by T, changing the order of the vectors does not change the result; it just changes the labeling of the vertices of $T(Q)$. So $\mathrm{vol}_n(T(Q))$ remains unchanged.



Here $T = [\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2]$, and the subset $T(Q)$ is $A \cup B$; the subset $[\vec{\mathbf{v}}_1 + a\vec{\mathbf{v}}_2, \vec{\mathbf{v}}_2](Q)$ is $B \cup C$. We need to show that

$$\mathrm{vol}_2\, A \cup B = \mathrm{vol}_2\, B \cup C,$$

which it does, since

$$B \cup C = B \cup (A + \vec{\mathbf{v}}_2)$$

and (by Proposition 4.1.22),

$$\mathrm{vol}_2(A + \vec{\mathbf{v}}_2) = \mathrm{vol}_2(A).$$

Properties 3 and 4 reflect properties that vol_n obviously ought to have if our definition of n-dimensional volume is to be consistent with our intuition about volume (not to mention the formulas for area and volume you learned in high school). The n vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ span an n-dimensional "parallelogram" (parallelogram in $\mathbb{R}^2$, parallelepiped in $\mathbb{R}^3, \dots$).

Saying that vol_n should have property 3 says exactly that if you multiply one side of such a parallelogram by a number a, keeping the others constant, it should multiply the volume by $|a|$.

Saying that vol_n should have property 4 is a way of saying that translating an object should not change its volume: once we know that vol_n is invariant by translation, property 4 will follow.

It isn't crystal clear that our definition using dyadic decompositions actually has these properties.

The main tool to do this problem is Proposition 4.1.22; we will apply it to prove property 4. Consider the following three subsets of $\mathbb{R}^n$

$$\begin{aligned}
A &= \{\, s_1\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_2 + s_3\vec{\mathbf{v}}_3 + \dots + s_n\vec{\mathbf{v}}_n \mid 0 \le s_i < 1,\ t \in [0, as_1)\,\} \\
B &= \{\, s_1\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_2 + s_3\vec{\mathbf{v}}_3 + \dots + s_n\vec{\mathbf{v}}_n \mid 0 \le s_i < 1,\ t \in [as_1, 1)\,\} \\
C &= \{\, s_1\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_2 + s_3\vec{\mathbf{v}}_3 + \dots + s_n\vec{\mathbf{v}}_n \mid 0 \le s_i < 1,\ t \in [1, 1 + as_1)\,\}.
\end{aligned}$$

These are shown in the margin for $n = 2$. Then

$$[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) = A \cup B, \qquad [\vec{\mathbf{v}}_1 + a\vec{\mathbf{v}}_2, \dots, \vec{\mathbf{v}}_n](Q) = B \cup C, \qquad C = A + \vec{\mathbf{v}}_2.$$

Property 4 follows.

For property 3, we will start by showing that it is true for $a = 1/2,\ 1/3,\ 1/4, \dots$. Choose $m > 0$, and define A_k by

$$A_{k,m} = \left\{ t\vec{\mathbf{v}}_1 + s_2\vec{\mathbf{v}}_2 + s_3\vec{\mathbf{v}}_3 + \dots + s_n\vec{\mathbf{v}}_n \;\middle|\; 0 \le s_i < 1, \frac{k-1}{m} \le t < \frac{k}{m} \right\}.$$

Then for any m, the $A_{k,m}$ are disjoint, $A_k = A_1 + \frac{k-1}{m}\vec{\mathbf{v}}_1$ so they all have the same volume, and

$$\bigcup_{k=1}^{m} A_k = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q), \quad \text{so} \quad \mathrm{vol}_n(A_k) = \frac{1}{m}[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q).$$

Now take $a > 0$ to be any positive real. Choose m, and let p be the number such that $p/m \le a < (p+1)/m$. Then

$$\cup_{k=1}^{p} A_k \subset [a\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) \subset \cup_{k=1}^{p+1} A_k, \quad \text{giving}$$

$$\frac{p}{m}\,\mathrm{vol}_n[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) \le \mathrm{vol}_n[a\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) \le \frac{p+1}{m}\,\mathrm{vol}_n[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q).$$

Clearly, the left and right terms can be made arbitrarily close by taking m sufficiently large, so

$$\begin{aligned}
\mathrm{vol}_n[a\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) &= \lim_{m\to\infty} \frac{p}{m}\,\mathrm{vol}_n[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) \\
&= a\,\mathrm{vol}_n[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q).
\end{aligned}$$

Finally, if $a < 0$, we have

$$[a\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) = \big[|a|\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n\big](Q) - a\vec{\mathbf{v}}_1,$$

so

$$\mathrm{vol}_n[a\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) = \mathrm{vol}_n[-a\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q) = |a| \,\mathrm{vol}_n[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n](Q).$$

Property 5: By Proposition 4.1.20,

$$I \mapsto \mathrm{vol}_n[\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_n](Q) = \mathrm{vol}_n(Q) = 1.$$

4.10.1 Since $x^2 + y^2 \le R^2$, y goes from $-R$ to R, and x goes from $-\sqrt{R^2 - y^2}$ to $\sqrt{R^2 - y^2}$. So

$$\begin{aligned}
\int_{D_R} (x^2+y^2)\,dx\,dy &= \int_{-R}^{R} \int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} (x^2+y^2)\,dx\,dy = \int_{-R}^{R} \left(\left[\frac{x^3}{3} + y^2 x \right]_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} \right) dy \\
&= \int_{-R}^{R} \left(\frac{2(R^2-y^2)^{3/2}}{3} + 2\sqrt{R^2-y^2}\, y^2 \right) dy \\
&= \int_{-R}^{R} \left(\frac{2\Big(R^2\big(1 - (\frac{y}{R})^2\big)\Big)^{3/2}}{3} + 2\sqrt{R^2(1 - \frac{y^2}{R^2})}\, \frac{R^2 y^2}{R^2} \right) dy \\
&= \int_{-R}^{R} \left(\frac{2R^3(1 - (\frac{y}{R})^2)^{3/2}}{3} + 2R^3 \sqrt{1 - \Big(\frac{y}{R}\Big)^2} \Big(\frac{y}{R}\Big)^2 \right) dy.
\end{aligned}$$

Now set $\sin\theta = \frac{y}{R}$ (which we can do because y goes from $-R$ to R) so that $dy = R\cos\theta\, d\theta$, and continue:

$$\begin{aligned}
\int_{D_R} (x^2+y^2)\,dx\,dy &= R^4 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{2}{3}\cos^4\theta + 2\cos^2\theta \sin^2\theta \right) d\theta \\
&= R^4 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{1}{6}(1 + 2 \overbrace{\cos 2\theta}^{\text{integrates to 0}} + \cos^2 2\theta) + \frac{1}{2}(1 - \cos^2 2\theta) \right) d\theta \\
&= R^4 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{1}{6} + \frac{1}{12} + \overbrace{\frac{\cos 4\theta}{12}}^{\text{integrates to 0}} + \frac{1}{2} - \frac{1}{4} - \overbrace{\frac{\cos 4\theta}{4}}^{\text{integrates to 0}} \right) d\theta \\
&= R^4 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2}\, d\theta = \frac{R^4 \pi}{2}.
\end{aligned}$$

In the second and third lines of this equation, $\cos 2\theta$ and $\cos 4\theta$ integrate to 0 because 2θ goes from $-\pi$ to π, where cos is as often negative as positive.

4.10.3 Since $|z^2 - \frac{1}{2}| = -\frac{1}{2}$ has no solutions, squaring both sides of $|z^2 - \frac{1}{2}| = \frac{1}{2}$ does not add more points to the graph. Squaring both sides gives $|z^2 - \frac{1}{2}|^2 = \frac{1}{4}$, which, using de Moivre's formula (equation 0.7.13), can be rewritten as

Remember (Definition 0.7.3) that $|a+ib|^2 = a^2+b^2$; it does not equal $a^2+2iab-b^2$.

$$\frac{1}{4} = \left| r^2\cos^2\theta + ir^2\sin^2\theta - \frac{1}{2}\right|^2 = r^4\cos^2 2\theta - r^2\cos 2\theta + \frac{1}{4} + r^4\sin^2 2\theta,$$

which gives the desired result:

$$r^2(r^2-\cos 2\theta) = 0, \quad \text{i.e.,} \quad r^2 = \cos 2\theta.$$

Solution 4.10.5, part b: This linear transformation is given by $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$. By Theorem 4.8.9, its determinant is abc.

4.10.5 a. Use the linear change of variables $T\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} au \\ bv \end{pmatrix}$, which maps the unit circle to the ellipse, i.e., the transformation given by $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$; see Example 4.9.9. The area of the ellipse is the area of the circle multiplied by $|\det T| = |ab|$, so the area is $\pi|ab|$.

b. This time, use the linear change of variables $T\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} au \\ bv \\ cw \end{pmatrix}$. This maps the unit sphere to the ellipsoid, whose volume is therefore $4\pi/3$ multiplied by $|\det T| = |abc|$.

Solution 4.10.7: For example, if A is the identity, then X_A is the unit ball:

$$Q_I(\vec{\mathbf{x}}) = \vec{\mathbf{x}} \cdot \vec{\mathbf{x}} = |\vec{\mathbf{x}}|^2.$$

More generally, X_A is an n-dimensional ellipsoid; we are computing the n-dimensional volume of such ellipsoids.

The volumes of the unit balls are computed in Example 4.5.7.

4.10.7 Let $X_A \subset \mathbb{R}^n$ denote the subset $Q_A(\vec{\mathbf{x}}) \le 1$. Suppose that there exists a symmetric matrix C such that $C^2 = C^\top C = A$. Note that $|\det C| = \sqrt{\det A}$.

The equation $Q_A(\vec{\mathbf{x}}) \le 1$ can be rewritten $(C\vec{\mathbf{x}})\cdot(C\vec{\mathbf{x}}) \le 1$:

$$1 \ge Q_A(\vec{\mathbf{x}}) = \mathbf{x}\cdot A\vec{\mathbf{x}} = \vec{\mathbf{x}}\cdot C^2\vec{\mathbf{x}} = \vec{\mathbf{x}}^\top C^\top C\vec{\mathbf{x}} = (C\vec{\mathbf{x}})^\top C\vec{\mathbf{x}} = C\vec{\mathbf{x}}\cdot C\vec{\mathbf{x}}.$$

In other words, the linear transformation C turns X_A into the unit ball $B_n \subset \mathbb{R}^n$: it takes a point in $\mathbb{R}^n$ satisfying $\mathbf{x}\cdot A\mathbf{x} \le 1$ and returns a point $C\mathbf{x} \in \mathbb{R}^n$ satisfying $|C\mathbf{x}|^2 \le 1$. This gives

$$\mathrm{vol}_n(X_A) \underbrace{=}_{\text{thm. 4.9.1}} \frac{\mathrm{vol}_n(B_n)}{|\det C|} = \frac{\mathrm{vol}_n(B_n)}{\sqrt{\det A}}. \tag{1}$$

Checking that C exists requires the spectral theorem (Theorem 3.7.15) and also some notions about changes of basis. By Theorem 3.7.15, there exists an orthonormal basis $\vec{\mathbf{v}}_1, \dots \vec{\mathbf{v}}_n$ of $\mathbb{R}^n$ such that $A\vec{\mathbf{v}}_i = \lambda_i\vec{\mathbf{v}}_i$. Since the quadratic form $Q(\vec{\mathbf{v}}) = \vec{\mathbf{v}}\cdot A\vec{\mathbf{v}}$ is positive definite by definition,

$$0 < \vec{\mathbf{v}}_i \cdot A\vec{\mathbf{v}}_i = \vec{\mathbf{v}}_i\cdot\lambda_i\vec{\mathbf{v}}_i = \lambda_i|\vec{\mathbf{v}}_i|^2,$$

which implies that all the eigenvalues λ_i are positive.

Recall (Proposition and Definition 2.4.19) that T is an orthogonal matrix if and only if $T^\top T = I$: if $T = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n]$, then the i,jth entry of $T^\top T$ is $\vec{\mathbf{v}}_i^\top\vec{\mathbf{v}}_j = \vec{\mathbf{v}}_i\cdot\vec{\mathbf{v}}_j$. Since the $\vec{\mathbf{v}}_i$ form an orthonormal basis, this dot product is 1 when $i = j$ and 0 otherwise. Thus $T^\top T = I$.

Let $T = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n]$. We need two properties of T:

1. Since T is an orthogonal matrix, $T^\top T = I$.
2. By Proposition 2.7.5,

$$T^{-1}AT = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}. \tag{2}$$

Let us set

$$C = T\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}T^{-1} = T\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}T^{\top}.$$

Then C is symmetric:

$$C^{\top} = \left(T\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}T^{\top}\right)^{\top} \underbrace{=}_{\text{Theorem 1.2.17}} (T^{\top})^{\top}\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}^{\top} T^{\top}$$

$$= T\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}T^{\top} = C.$$

Moreover,

$$C^2 = T\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}\underbrace{T^{\top}T}_{I}\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}T^{\top}$$

$$= T\begin{bmatrix}\sqrt{\lambda_1} & 0 & \dots & 0\\ 0 & \sqrt{\lambda_2} & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \sqrt{\lambda_n}\end{bmatrix}^2 T^{\top} = T\begin{bmatrix}\lambda_1 & 0 & \dots & 0\\ 0 & \lambda_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & \lambda_n\end{bmatrix}T^{\top} \underbrace{=}_{\text{eq. (2)}} A.$$

This confirms our initial assumption that there exists a symmetric matrix C such that $C^2 = A$, and equation 1 is correct: the volume of X_A is

$$\mathrm{vol}_n(X_A) = \frac{\mathrm{vol}_n(B_n)}{\sqrt{\det A}}, \quad \text{where } B_n \text{ is the unit ball in } \mathbb{R}^n.$$

Solution 4.10.7: Our construction of the square root C of A is a special case of *functional calculus*. Given any real-valued function of a real variable, we can apply the function to real *symmetric* matrices by the same procedure: diagonalize the matrix using the spectral theorem, apply the function to the eigenvalues, and "undiagonalize" back. The same idea applied to linear operators in infinite-dimensional vector spaces is a fundamental part of functional analysis.

4.10.9 Assume that a, b, and c are all positive, and use the "elliptical change of coordinates" E shown in the margin. The region V corresponds to the region

$$U = \left\{\begin{pmatrix} r\\ \theta\\ \varphi\end{pmatrix} \in \mathbb{R}^3 \,\middle|\, 0 < r \le 1,\ 0 \le \theta \le \frac{\pi}{2},\ 0 \le \varphi < \frac{\pi}{2}\right\}$$

in our coordinates. Before we can use these coordinates to integrate the desired function, we must calculate $|\det[\mathbf{D}E]|$. To get $[\mathbf{D}E]$ from $[\mathbf{D}S]$ (calculated in equation 4.10.22), we need only multiply the first, second, and third rows of $[\mathbf{D}S]$ by a, b, and c respectively, to get

$$\left[\mathbf{D}E\begin{pmatrix} r\\ \theta\\ \varphi\end{pmatrix}\right] = \begin{bmatrix} a\cos\theta\cos\varphi & -ar\sin\theta\cos\varphi & -ar\cos\theta\sin\varphi\\ b\sin\theta\cos\varphi & br\cos\theta\cos\varphi & -br\sin\theta\sin\varphi\\ c\sin\varphi & 0 & cr\cos\varphi\end{bmatrix}.$$

Elliptical change of coordinates for Solution 4.10.9:

$$E\colon \begin{pmatrix} r\\ \theta\\ \varphi\end{pmatrix} \mapsto \begin{pmatrix} x = ar\cos\theta\cos\varphi\\ y = br\sin\theta\cos\varphi\\ z = cr\sin\varphi\end{pmatrix}$$

Theorem 4.8.8, together with equation 4.8.16, tells us that multiplying a row of a matrix by a scalar a scales the determinant of the matrix by a, so $\det[\mathbf{D}E] = abc \det[\mathbf{D}S]$, and

By equation 4.10.23,
$$\left|\det\left[\mathbf{D}S\begin{pmatrix} r\\ \theta\\ \varphi\end{pmatrix}\right]\right| = r^2\cos\varphi.$$

$$\left|\det\left[\mathbf{D}E\begin{pmatrix} r\\ \theta\\ \varphi\end{pmatrix}\right]\right| = abc\, r^2\cos\varphi.$$

Integration is now a simple matter. First change coordinates under the multiple integral:

$$\int_V xyz|dx\,dy\,dz| = \int_U (ar\cos\theta\cos\varphi)(br\sin\theta\cos\varphi)(cr\sin\varphi)abcr^2\cos\varphi|dr\,d\theta\,d\varphi|.$$

Rearranging terms and changing to an iterated integral yields

$$(abc)^2\int_0^1\left(\int_0^{\pi/2}\left(\int_0^{\pi/2} r^5\sin\theta\cos\theta\cos^3\varphi\sin\varphi\,d\theta\right)d\varphi\right)dr.$$

This is quite simple to evaluate if one notes that

$$\cos\theta\,d\theta = d\sin\theta \quad \text{and} \quad -\sin\varphi\,d\varphi = d\cos\varphi.$$

The result is $(abc)^2/48$.

4.10.11 Use spherical coordinates, to find

$$\begin{aligned}\int_0^R\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} r^2\cos\varphi\,d\varphi\,d\theta\,dr &= \int_0^R\int_0^{2\pi}\left[r^2\sin\varphi\right]_{-\pi/2}^{\pi/2}d\theta\,dr\\ &= \int_0^R\left[2r^2\theta\right]_0^{2\pi}dr = \left[\frac{4\pi r^3}{3}\right]_0^R = \frac{4R^3\pi}{3}.\end{aligned}$$

4.10.13 Cylindrical coordinates are appropriate here. We find

$$\int_0^{2\pi}\int_0^1\int_{r^2}^r r\,dz\,dr\,d\theta = 2\pi\int_0^1 r(r-r^2)\,dr = 2\pi\left(\frac{1}{3}-\frac{1}{4}\right) = \frac{\pi}{6}.$$

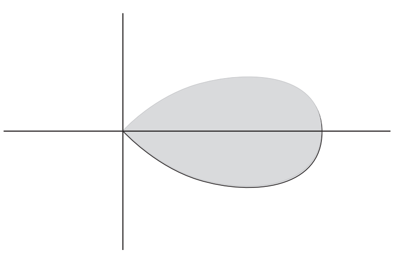


FIGURE FOR SOLUTION 4.10.15
The curve of part a

4.10.15 a. The curve is drawn at left.

b. The x-coordinate of the center of gravity is the point

$$\overline{x} = \frac{\int_{-\pi/4}^{\pi/4}\int_0^{\cos 2\theta}(r\cos\theta)r\,dr\,d\theta}{\int_{-\pi/4}^{\pi/4}\int_0^{\cos 2\theta} r\,dr\,d\theta}.$$

We now compute these two integrals.

The numerator gives

$$\frac{1}{3}\int_{-\pi/4}^{\pi/4} \cos^3 2\theta \cos\theta \, d\theta = \frac{1}{12}\int_{-\pi/4}^{\pi/4} (\cos 6\theta \cos\theta + 3\cos 2\theta \cos\theta)\, d\theta$$
$$= \frac{1}{24}\int_{-\pi/4}^{\pi/4} (\cos 7\theta + \cos 5\theta + 3\cos 3\theta + 3\cos\theta)\, d\theta$$
$$= \frac{\sqrt{2}}{24}\left(-\frac{1}{7} - \frac{1}{5} + 1 + 3\right).$$

The denominator gives

$$\int_{-\pi/4}^{\pi/4}\int_0^{\cos 2\theta} r\, dr\, d\theta = \frac{1}{2}\int_{-\pi/4}^{\pi/4} \cos^2 2\theta \, d\theta = \frac{\pi}{8}.$$

After some arithmetic, this comes out to $\overline{x} = (44\sqrt{2})/(35\pi) \approx .600211\ldots$.

By symmetry, we have $\overline{y} = 0$.

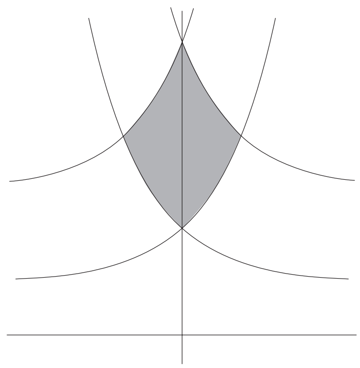


FIGURE FOR SOLUTION 4.10.17 The shaded region is the region A, bounded by the arcs of the curves

$$y = e^x + 1 \text{ and } y = e^a(e^x + 1)$$

where $0 \le x \le a$, and the arcs of the curves

$$y = e^{-x} + 1 \text{ and } y = e^a(e^{-x} + 1)$$

where $-a \le x \le 0$.

4.10.17 a. The region A is shown in the figure in the margin. It is bounded by the arcs of the curves $y = e^x + 1$ and $y = e^a(e^x + 1)$ where $0 \le x \le a$, and the arcs of the curves $y = e^{-x} + 1$ and $y = e^a(e^{-x} + 1)$ where $-a \le x \le 0$.

b. In this case, we can find an explicit inverse of Φ: we need to solve, for u and v in terms of x and y, the system of equations

$$\begin{aligned} u - v &= x \\ e^u + e^v &= y \end{aligned}.$$

Write $u = x + v$, and substitute in the second equation, to find

$$e^{x+v} + e^v = e^v(e^x + 1) = y.$$

Since $y \ge 0$ in A and $e^x + 1 > 0$ always, this leads to $v = \ln \dfrac{y}{e^x + 1}$.

An exactly analogous argument gives

$$u = \ln \frac{y}{e^{-x} + 1}.$$

c. We have

$$\det\left[\mathbf{D}\Phi\begin{pmatrix} x \\ y \end{pmatrix}\right] = \det\begin{bmatrix} 1 & -1 \\ e^u & e^v \end{bmatrix} = e^u + e^v.$$

The change of variables formula gives

$$\int_A y|dx\, dy| = \int_{Q_a} (e^u + e^v)(e^u + e^v)\, |du\, dv|$$
$$= \int_0^a \int_0^a \left(e^{2u} + 2e^u e^v + e^{2v}\right) du\, dv$$
$$= a(e^{2a} - 1) + 2(e^a - 1)^2.$$

4.10.19 a. The image of $0 \le u \le 1$, $v = 0$, is the arc of parabola parametrized by $u \mapsto \begin{pmatrix} u \\ u^2 \end{pmatrix}$, i.e., the arc of the parabola $y = x^2$ where $0 \le x \le 1$.

The image of $0 \le u \le 1$, $v = 1$, is the arc of parabola parametrized by $u \mapsto \begin{pmatrix} u-1 \\ u^2+1 \end{pmatrix}$, i.e., the arc of the parabola $y = (x+1)^2 + 1$ where $-1 \le x \le 0$.

The image of $u = 0$, $0 \le v \le 1$, is the arc of parabola parametrized by $v \mapsto \begin{pmatrix} -v^2 \\ v \end{pmatrix}$, i.e., the arc of the parabola $x = -y^2$ where $0 \le y \le 1$.

The image of $u = 1$, $0 \le v \le 1$, is the arc of parabola parametrized by $v \mapsto \begin{pmatrix} 1-v^2 \\ 1+v \end{pmatrix}$, i.e., the arc of the parabola $x - 1 = -(y-1)^2$ where $1 \le y \le 2$.

These curves are shown in the margin.

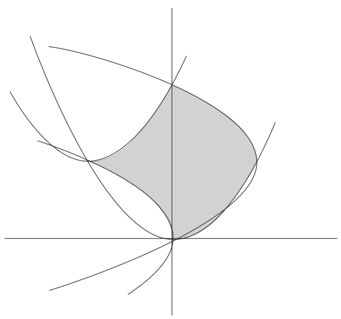


FIGURE FOR SOLUTION 4.10.19
The shaded area is $\Phi(Q)$. We have two vertical parabolas; the shaded region is inside the lower one and outside the upper one.

There are also two horizontal parabolas; the shaded region is inside the rightmost one and outside the leftmost one.

b. We need to show that Φ is 1–1 so that we will be able to apply the change of variables formula in part c. If $\Phi\begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = \Phi\begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$, we get

$$\begin{aligned} u_1 - v_1^2 &= u_2 - v_2^2 \\ u_1^2 + v_1 &= u_2^2 + v_2. \end{aligned}$$

These equations can be rewritten

$$\begin{aligned} u_1 - u_2 = v_1^2 - v_2^2 &= (v_1 - v_2)(v_1 + v_2) \\ (u_1 - u_2)(u_1 + u_2) &= v_2 - v_1. \end{aligned}$$

In the region under consideration, $v_1 + v_2 \ge 0$ and $u_1 + u_2 \ge 0$. It follows from the first equation that if $v_1 \ne v_2$, then $u_1 \ne u_2$, and both $v_1 - v_2$ and $u_1 - u_2$ have the same sign. The same argument for the second equation says that if $u_1 \ne u_2$, then $v_1 \ne v_2$, and $v_1 - v_2$ and $u_1 - u_2$ have opposite sign. So $u_1 = u_2$ and $v_1 = v_2$.

c. We have $\det\left[\mathbf{D}\Phi\begin{pmatrix} u \\ v \end{pmatrix}\right] = \det\begin{bmatrix} 1 & -2v \\ 2u & 1 \end{bmatrix} = 1 + 4uv$, which is positive in Q, so

$$\begin{aligned} \int_A x\,|dx\,dy| = \int_Q (u - v^2)(1+4uv)|du\,dv| &= \int_0^1 \int_0^1 (u + 4u^2v - v^2 - 4uv^3)\,du\,dv \\ &= \int_0^1 \left[\frac{u^2}{2} + \frac{4u^3v}{3} - v^2u - 2u^2v^3\right]_{u=0}^{u=1} dv = \int_0^1 \left(\frac{1}{2} + \frac{4v}{3} - v^2 - 2v^3\right) dv \\ &= \left[\frac{v}{2} + \frac{2v^2}{3} - \frac{v^3}{3} - \frac{v^4}{2}\right]_0^1 = \frac{1}{3}. \end{aligned}$$

4.10.21 a. Let $\Phi\begin{pmatrix} r \\ \theta \\ z \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ z \end{pmatrix}$, and let the region B in (r, θ, z)-space be a subset on which Φ is 1–1. Set $A = \Phi(B)$. If $f : A \to \mathbb{R}$ is integrable, then $rf\left(\Phi\begin{pmatrix} r\\ \theta\\ z\end{pmatrix}\right)$ is integrable over B, and

$$\int_B rf\left(\Phi\begin{pmatrix} r\\ \theta\\ z\end{pmatrix}\right)|dr\,d\theta\,dz| = \int_A f\begin{pmatrix} x\\ y\\ z\end{pmatrix}|dx\,dy\,dz|.$$

In this case, the appropriate cylindrical coordinates consist of keeping x, and writing $\begin{pmatrix} y\\ z\end{pmatrix}$ in polar coordinates. Then r is the distance to the x-axis, so the integral becomes

$$\int_a^b\int_0^{2\pi}\int_0^{f(x)} r\,r^2\,dr\,d\theta\,dx = \frac{\pi}{2}\int_a^b \big(f(x)\big)^4\,dx.$$

b. Applying the formula above gives that the moment of inertia is

$$\begin{aligned}\frac{\pi}{2}\int_{-\pi/2}^{\pi/2}\cos^4 x\ dx &= \frac{\pi}{2}\int_{-\pi/2}^{\pi/2}\cos^2 x(1-\sin^2 x)\ dx\\ &= \frac{\pi}{2}\int_{-\pi/2}^{\pi/2}\left(\cos^2 x - \frac{\sin^2 2x}{4}\right)\,dx\\ &= \frac{\pi}{2}\left(\frac{\pi}{2}-\frac{\pi}{8}\right) = \frac{3\pi^2}{16}.\end{aligned}$$

4.11.1 Set $f(\mathbf{x}) = |\mathbf{x}|^p\mathbf{1}_D(\mathbf{x})$, where D is the unit disc. Let $A_n \subset \mathbb{R}^2$ be the set $1/n < |\mathbf{x}| \le 1$ and set $f_n = |\mathbf{x}|^p\mathbf{1}_{A_n}$. Then each f_n is L-integrable (in fact, R-integrable), $0 \le f_1 \le f_2 \le \cdots$, and $\lim_{n\to\infty} f_n \underset{L}{=} f$, so f is integrable precisely if the sequence

$$a_n = \int_{\mathbb{R}^2} f_n(\mathbf{x})\,|d^2\mathbf{x}| = \int_{A_n}|\mathbf{x}|^p\,|d^2\mathbf{x}|$$

is bounded. This sequence is easy to compute explicitly by passing to polar coordinates:

$$a_n = \int_0^{2\pi}\left(\int_{1/n}^1 r^p\,r\,dr\right)d\theta = \begin{cases}\frac{2\pi}{p+2}\left(1-\frac{1}{n^{p+2}}\right) & \text{if } p\ne -2\\ 2\pi\ln n & \text{if } p=-2.\end{cases}$$

Thus we see that f is integrable precisely if $p > -2$; in that case, the integral is $2\pi/(p+2)$.

4.11.3 Using the change of variables for polar coordinates, we have the following, where $A = (1,\infty)\times[0,2\pi)$:

$$\begin{aligned}\int_{\mathbb{R}^2-B_1(\mathbf{0})}\left|\begin{pmatrix} x\\ y\end{pmatrix}\right|^p|dx\,dy| &= \int_A\left|\begin{pmatrix} r\cos\theta\\ r\sin\theta\end{pmatrix}\right|^p r\,|dr\,d\theta|\\ &= \int_0^{2\pi}\int_1^\infty r^{p+1}\,dr\,d\theta = 2\pi\int_1^\infty r^{p+1}\,dr.\end{aligned}$$

By Theorem 4.11.20, the existence of this last integral is equivalent to the existence of the original integral. If $p \neq -2$, we can write

$$\int_1^\infty r^{p+1}\,dr = \sum_{n=1}^\infty \int_n^{n+1} r^{p+1}\,dr = \sum_{n=1}^\infty \left[\frac{r^{p+2}}{p+2}\right]_n^{n+1}.$$

Our definition of integrability says that if this series converges, the integral exists. But the series telescopes:

$$\sum_{n=1}^m \left[\frac{r^{p+2}}{p+2}\right]_n^{n+1} = \frac{1}{p+2}\big((m+1)^{p+2}-1\big)$$

and converges if $p+2<0$, i.e., if $p<-2$. Moreover, since $|\mathbf{x}|^p > 0$, the integral is then precisely the sum of the series, so if $p<-2$, we have

$$\int_{\mathbb{R}^2 - B_1(\mathbf{0})} \left|\begin{pmatrix} x \\ y \end{pmatrix}\right|^p |dx\,dy| = -\frac{1}{p+2}.$$

We now need to see that if $p \geq -2$, the integral does not exist. By the monotone convergence theorem, the integral will fail to converge if

$$\sup_{m\to\infty} \int_1^m r^{p+1}\,dr = \infty.$$

There are two cases to consider: if $p \neq -2$, we just computed that integral, to find

$$\int_1^m r^{p+1}\,dr = \frac{1}{p+2}\big((m+1)^{p+2}-1\big),$$

which tends to infinity with m if $p+2>0$. If $p=-2$, then

$$\sup_{m\to\infty} \int_1^m \frac{1}{r}\,dr = \sup_{m\to\infty} \ln m = \infty.$$

4.11.5 Let $A_n = \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; 2^n \leq |\mathbf{x}| < 2^{n+1} \right\}$, so that

$$\mathbb{R}^n - B_1(\mathbf{0}) = A_0 \cup A_1, \cup \cdots.$$

Then

$$|\mathbf{x}|^p \mathbf{1}_{\mathbb{R}^n - B_1(\mathbf{0})} = \sum_{m=0}^\infty |\mathbf{x}|^p \mathbf{1}_{A_m},$$

and so $|\mathbf{x}|^p$ is L-integrable over $\mathbb{R}^n - B_1(\mathbf{0})$ precisely if the series

$$\sum_{m=0}^\infty \int_{A_m} |\mathbf{x}|^p |d^n\mathbf{x}| \quad \text{is convergent.}$$

The mapping $\varphi_m(\mathbf{x}) = 2^m\mathbf{x}$ maps A_1 to A_m, and $\det[\mathbf{D}\varphi_m] = 2^{nm}$. So applying the change of variables formula gives

$$\int_{A_m} |\mathbf{x}|^p |d^n\mathbf{x} = \int_{A_1} 2^{mp}|\mathbf{x}|^p 2^{nm}|d^n\mathbf{x}| = 2^{m(p+n)} \int_{A_1} |\mathbf{x}|^p|d^n\mathbf{x}|.$$

Thus our series is

$$\sum_{m=0}^{\infty} 2^{m(n-p)} \int_{A_1} |\mathbf{x}|^p \, |d^n\mathbf{x}|,$$

where the integral is some constant. It is convergent exactly when

$$n + p < 0, \quad \text{i.e.,} \quad \text{if } p < -n.$$

4.11.7 The situation is rather different for $m \geq 1$ and $m < 1$. If $m \geq 1$, then $1/(x^2+y^2)^p$ is integrable over A_m precisely if it is integrable over $\mathbb{R}^n - B_1(0)$, i.e., if $p > 1$. (See Exercise 4.11.3; note that the present p is half what is denoted p there.)

This is seen as follows: take the eight subsets $A_{m,1}, \dots, A_{m,8}$ obtained by reflecting A_m with respect to both axes and both diagonals. These eight copies cover the complement of the square S where $|x| \leq 1$, $|y| \leq 1$, with some points covered twice when $m > 1$. Thus

$$\begin{aligned}\int_{A_m} \frac{1}{(x^2+y^2)^p} \, |dx\, dy| &\leq \int_{\mathbb{R}^2 - S} \frac{1}{(x^2+y^2)^p} \, |dx\, dy| \\ &\leq 8 \int_{A_m} \frac{1}{(x^2+y^2)^p} \, |dx\, dy|.\end{aligned}$$

Clearly, integrability over $\mathbb{R}^2 - B_1(0)$ and over $\mathbb{R}^2 - S$ are equivalent, since $1/(x^2+y^2)^p$ is always integrable over $S - B_1(0)$.

The case where $m < 1$ is more delicate. In that case, in A_m we have $x^2 \leq x^2 + y^2 \leq 2x^2$, so the integrability of $1/(x^2+y^2)^p$ is equivalent to integrability of $1/x^{2p}$. But this integral can be computed explicitly:

$$\int_1^\infty \left(\int_0^{x^m} \frac{1}{x^{2p}} dy \right) dx = \int_1^\infty x^{m-2p} \, dx,$$

which is finite precisely if $m - 2p < -1$, i.e., if $p > (m+1)/2$.

4.11.9 There are no difficulties associated with infinities in this case:

$$\begin{aligned}\widehat{\mathbf{1}_{[-1,1]}}(\xi) &= \int_{\mathbb{R}} \mathbf{1}_{[-1,1]}(x) e^{i\xi x} dx = \int_{-1}^{1} e^{i\xi x} dx \\ &= \left[\frac{e^{i\xi x}}{i\xi} \right]_{-1}^{1} = \frac{e^{i\xi} - e^{-i\xi}}{i\xi} = 2\frac{\sin \xi}{\xi}.\end{aligned}$$

$$Theonlyfishypartisthecase$$

$\xi = 0$; but note that

Solution 4.11.11: This is a matter of pinning down the fact that exponentials win over polynomials.

$$\lim_{\xi \to 0} 2 \frac{\sin \xi}{\xi} = 2 = \widehat{\mathbf{1}_{[-1,1]}}(0).$$

4.11.11 It is enough to show the result for any monomial $x_1^{k_1} \cdots x_n^{k_n}$. Note that $|x_i^{k_i}| \leq (1 + |\mathbf{x}|)^{k_i}$, so

$$|x_1^{k_1} \dots x_n^{k_n}| \leq (1 + |\mathbf{x}|)^k,$$

where $k = k_1 + \cdots + k_n$. Thus it is enough to show that the integral

$$\int_{\mathbb{R}^n} (1 + |\mathbf{x}|)^k e^{-|\mathbf{x}|^2} \, |d^n\mathbf{x}|$$

exists for every $k \geq 0$. Developing out the power, it is enough to show that

$$\int_{\mathbb{R}^n} |\mathbf{x}|^k e^{-|\mathbf{x}|^2} \, |d^n\mathbf{x}| \quad \text{exists for every } k \geq 0.$$

Let us break up $\mathbb{R}^n$ into the unit ball B and the sets

$$A_m = \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; 2^m < |\mathbf{x}| \leq 2^{m+1} \right\}.$$

We can then write

$$|\mathbf{x}|^k = |\mathbf{x}|^k \mathbf{1}_B(\mathbf{x}) + \sum_{m=0}^{\infty} |\mathbf{x}|^k \mathbf{1}_{A_m}(\mathbf{x}) \overset{\text{def}}{=} g(\mathbf{x}) + \sum_{m=0}^{\infty} f_m(\mathbf{x}),$$

and g and all the f_m are R-integrable. Therefore it is enough to prove that the series of numbers

$$\sum_{m=1}^{\infty} \int_{\mathbb{R}^n} f_m(\mathbf{x}) |d^n\mathbf{x}|$$

is convergent. The change of variable $\phi_m : \mathbf{x} \mapsto 2^m\mathbf{x}$ transforms the region A_1 into the region A_m, and the change of variables formula gives

$$\int_{\mathbb{R}^n} f_m(\mathbf{x}) |d^n\mathbf{x}| = \int_{A_1} 2^{mn} 2^{km} |\mathbf{x}|^k e^{-2^{2m}|\mathbf{x}|^2} |d^n\mathbf{x}| \leq 2^{mn} 2^{km} 2^k e^{-2^{2m}} \overset{\text{def}}{=} a_m.$$

The ratio test, applied to this series, gives

$$\frac{a_{m+1}}{a_m} = \frac{2^{(m+1)n} 2^{k(m+1)} 2^k e^{-2^{2(m+1)}}}{2^{mn} 2^{km} 2^k e^{-2^{2m}}} = \frac{2^{n+k}}{e^{2^{2(m+1)} - 2^{2m}}} = \frac{2^{n+k}}{e^{3\, 2^{2m}}},$$

which clearly tends to 0 as m tends to infinity, so the series converges.

Here is another solution, by Vorrapan Chandee, when a freshman at Cornell:

It is clearly enough to show the result for any monomial $x_1^{k_1} \cdots x_n^{k_n}$. By Fubini's theorem, we have

$$\begin{aligned}
\int_{\mathbb{R}^n} x_1^{k_1} \cdots x_n^{k_n} e^{-|\mathbf{x}|^2} \, |d^n\mathbf{x}| &= \int_{\mathbb{R}^n} x_1^{k_1} e^{-x_1^2} \cdots x_n^{k_n} e^{-x_n^2} \, |d^n\mathbf{x}| \\
&= \left(\int_{\mathbb{R}} x_1^{k_1} e^{-x_1^2} |dx_1| \right) \cdots \left(\int_{\mathbb{R}} x_n^{k_n} e^{-x_n^2} |dx_n| \right).
\end{aligned}$$

The original integral exists if each integral in the above product exists.

So it suffices to show that the one-dimensional integral $\int_{\mathbb{R}} x^k e^{-x^2} |dx|$ exists for $k = 0, 1, 2, \ldots$. By a symmetry argument, we can reduce this to showing that $\int_0^\infty x^k e^{-x^2} dx$ exists. Indeed, $x^k e^{-x^2}$ is eventually dominated by e^{-x}, which is easily seen to be integrable over $[0, \infty)$.

4.11.13 a. For the first sequence, if the f_k converge to anything, they must converge to the function 0. But if you take $\epsilon = 1/2$, then for $x = k + 1/2$ we have $f_k(x) - 0 = 1 > \epsilon$.

For the second example, even if you take $\epsilon = 1$ and $x_k = \frac{1}{2k}$, we have $f_k(x_k) - 0 = k \geq 1$.

For the third, let f_∞ be the function that is 1 on the rationals and 0 on the irrationals. Set $x_k = a_{k+1}$ and $\epsilon = 1/2$. Then

$$|f_k(x_k) - f_\infty(x_k)| = 1 > \epsilon.$$

b. Write $p = a_0 + a_1 x + \cdots + a_m x^m$. Suppose that $k \mapsto p_k$ converges uniformly to p. Choose $\epsilon > 0$; our assumption says that there exists K such that when $k \geq K$, then for any x we have $|p_k(x) - p(x)| < \epsilon$. But $p_k(x) - p(x)$ is a polynomial, and the only bounded polynomials are the constants. Thus $p_k(x) - p(x) = c_k$ for the constant $c_k = a_{0,k} - a_0$ when $k \geq K$. This proves that all the coefficients of the p_k are eventually constant, except perhaps the constant term. Moreover, the inequality

$$|p_k(x) - p(x)| = |a_{0,k} - a_k| < \epsilon$$

clearly implies that $k \mapsto a_{k,0}$ converges to a_k.

In the other direction we must show that if all the nonzero coefficients are constant, and the constant coefficient converges, then the sequence $k \mapsto p_k$ converges uniformly to p. Suppose the coefficients $a_{i,k}$ are eventually constant for $i > 0$, so we can set

$$a_{i,k} = a_i \quad \text{for } k > K \text{ and } i > 0,$$

and suppose $a_{0,k}$ converges to a_0. Set

$$p = a_0 + a_1 x + \cdots + a_m x^m;$$

clearly, if $k > K$, then

$$p(x) - p_k(x) = a_0 - a_{0,k},$$

which converges to 0; therefore, p_k converges uniformly to p.

c. If the two sequences of functions $k \mapsto f_k$ and $k \mapsto g_k$ converge uniformly to f and g, then the sequence $k \mapsto f_k + g_k$ converges uniformly to $f + g$. Thus it is enough to show that if the sequence of numbers $k \mapsto a_k$ converges, say to a_∞, then the sequence of functions $k \mapsto a_k x^i \mathbf{1}_A$ converges uniformly on $\mathbb{R}$. Let M be the maximum of $|x|^i$ on A. Choose $\epsilon > 0$. There exists K such that if $k > K$, then $|a_k - a_\infty| < \epsilon$, and then

$$|a_k x^i - a_\infty x^i| \leq \epsilon M.$$

This can be made arbitrarily small by choosing ϵ sufficiently small.

Solution 4.11.13, part c: The existence of M uses the hypothesis that A is bounded; any polynomial has a maximum on any bounded set.

4.11.15 Since this is a telescoping series, it is enough to show that the last term of any partial sum tends to 0. Since $x(\ln x - 1)$ is an antiderivative of $\ln x$, the last term of

$$\sum_{i=1}^{n-1} \int_{1/2^{i+1}}^{1/2^i} |\ln x|\, dx \quad \text{is} \quad \frac{1}{2^n}\left(\ln\left|\frac{1}{2^n}\right| - 1\right) = -\frac{1}{2^n}(n \ln 2 + 1).$$

This tends to 0 when n tends to infinity: 2^n grows much faster than n.

Solution 4.11.17 illustrates the fact that if we allow improper integrals whose existence depends on cancellations, the change of variables formula is not true. The idea of doing integration theory without being able to make changes of variables is ridiculous.

4.11.17 After change of variables, this integral becomes $\displaystyle\int_0^\infty \frac{1}{u}\sin\frac{1}{u}\,du$. As an improper integral, this should mean

$$\lim_{A\to\infty}\int_0^A \frac{1}{u}\sin\frac{1}{u}\,du;$$

see equation 4.11.52. However, the integral inside the limit above does not exist; the integrand $\frac{1}{u}\sin\frac{1}{u}$ oscillates between the graph of $1/u$ and the graph of $-1/u$; near $u=0$ the functions $1/u$ and $-1/u$ are not integrable.

4.11.19 a. Let $n\mapsto c_n$ be an increasing sequence of numbers converging to b, and set

$$g_n(x) = |f(x)|\mathbf{1}_{[a,c_n]}.$$

Since $\lim_{c\to b}\int_a^c |f(x)|\,dx$ exists, it follows that $\lim_{n\to\infty}\int_a^b g_n(x)\,dx$ exists, so hypothesis 4.11.64 of the monotone convergence theorem is satisfied for g_n. So by that theorem, $|f|$ is L-integrable. Now set

Exercise 4.11.19 brings out the importance of the monotone convergence theorem and the dominated convergence theorem when proving results about Lebesgue integrals.

$$f_n(x) = f(x)\mathbf{1}_{[a,c_n)}(x).$$

Then $|f_n(x)| \le |f(x)|$, so by the dominated convergence theorem, since $f(x)=\lim_{n\to\infty} f_n(x)$, the function f is L-integrable, and

$$\lim_{n\to\infty}\int_a^{c_n} f(x)\,dx. = \lim_{n\to\infty}\int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx.$$

4.11.21 Let $m\le n$. Define $B_i\subset\mathbb{R}^n$ as in equation 4.11.45, and let B_0 be the unit ball. Set $f_i = f\mathbf{1}_{B_i}$ and $g_k=\sum_{i=1}^k f_i$. For $i\ge 1$, we have

$$f_i(\mathbf{x}) \ge \frac{1}{2|\mathbf{x}|^m}\mathbf{1}_{B_i}(\mathbf{x}).$$

The sequence $k\mapsto\int_{\mathbb{R}^n} g_k(\mathbf{x})|d^n\mathbf{x}|$ diverges because

$$\begin{aligned}\int_{\mathbb{R}^n} g_k(\mathbf{x})|d^n\mathbf{x}| &= \sum_{i=1}^k\int_{\mathbb{R}^n} f_i(\mathbf{x})|d^n\mathbf{x}|\\ &\ge \frac{1}{2}\sum_{i=1}^k\int_{B_i}\frac{1}{|\mathbf{x}|^m}|d^n\mathbf{x}| = \frac{1}{2}\sum_{i=1}^k 2^{(n-m)(i-1)}\int_{B_1}\frac{1}{|\mathbf{x}|^m}|d^n\mathbf{x}|.\end{aligned}$$

This is the sequence of partial sums of a divergent geometric series. Thus we are in the situation of part 2 of the monotone convergence theorem (Theorem 4.11.18):

$$f = f_0 + \sup_k g_k \quad\text{and}\quad 0\le g_1\le g_2\le\cdots.$$

Since $\sup_k\int_{\mathbb{R}^n} g_k(\mathbf{x})|d^n\mathbf{x}| = \infty$, we see that f is not integrable over $\mathbb{R}^n$ when $m\le n$.

Solutions for review exercises, chapter 4

4.1 It is enough that $U_N(\mathbf{1}_C) = L_N(\mathbf{1}_C)$ because by Lemma 4.1.9, for all $M \geq N$,

$$U_N(\mathbf{1}_C) \geq U_M(\mathbf{1}_C) \geq L_M(\mathbf{1}_C) \geq L_N(\mathbf{1}_C).$$

Thus the upper sums and lower sums are all equal for $M \geq N$, so the function $\mathbf{1}_C$ is integrable.

But $U_N(\mathbf{1}_C) = L_N(\mathbf{1}_C)$ is obviously true.

4.3 1. False. For instance, multiplication by 2, i.e., the function

$$[2] : \mathbb{Z} \to \mathbb{Z},$$

is not onto; its image is the even integers.

2. True. If A is onto, the image of $\mathbb{Z}^n$ contains the standard basis vectors, so it spans $\mathbb{R}^n$. In particular, the matrix A viewed as a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$, is onto, hence also injective. But any element of the kernel of $A : \mathbb{Z}^n \to \mathbb{Z}^n$ is also an element of the kernel of $A : \mathbb{R}^n \to \mathbb{R}^n$, so it must be $\mathbf{0}$.

3. True. The same argument works as for 2: if we view A as representing a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$, then $\det A \neq 0$ implies that A is injective, and so is its restriction to $\mathbb{Z}^n$.

4. False. The same example as part 1 is a counterexample here also: $\det[2] \neq 0$ but $[2] : \mathbb{Z} \to \mathbb{Z}$ is not onto.

4.5 There are many ways to approach this problem. One is to observe that $\displaystyle \frac{1}{N^2} \sum_{k=1}^{N} \sum_{j=1}^{2N} e^{\frac{k+j}{N}}$ is a Riemann sum (unfortunately not dyadic) for the function

$$f\begin{pmatrix} x \\ y \end{pmatrix} = e^{x+y} \mathbf{1}_{[0,1]\times[0,2]}.$$

By Proposition 4.1.16, this function is integrable, with integral

$$\left(\int_0^2 e^y \, dy \right) \left(\int_0^1 e^x \, dx \right) = (e-1)(e^2-1).$$

With the arguments at hand, this is a little shaky; if we only considered the numbers $N = 2^M$ we would be fine, but the other values of N won't quite enter into our dyadic formalism.

Another possibility is to write

$$\left(\frac{1}{N} \sum_{k=1}^{N} e^{k/N} \right) \left(\frac{1}{N} \sum_{j=1}^{2N} e^{j/N} \right)$$

and to observe that both sums are finite geometric series. Using

$$a + ar + \cdots + ar^n = a\frac{1 - r^{n+1}}{1 - r}$$

(see equation 0.5.4), we see that the first sum is

$$\sum_{k=1}^{N} e^{k/N} = e^{1/N}\frac{1 - e^{(N+1)/N}}{1 - e^{1/N}}.$$

We need to evaluate the limit

$$\lim_{N\to\infty} \frac{1}{N} e^{1/N}\frac{1 - e^{(N+1)/N}}{1 - e^{1/N}}.$$

In this expression, $e^{1/N} \to 1$ and $e^{(N+1)/N} \to e$, so the limit that matters is

Equation 1 uses l'Hôpital's rule.

$$\lim_{N\to\infty} \frac{1}{N(1 - e^{1/N})} = \lim_{x\to 0} \frac{x}{1 - e^x} = \lim_{x\to 0} \frac{1}{-e^x} = -1, \tag{1}$$

giving the limit

$$\lim_{N\to\infty} \frac{1}{N} e^{1/N}\frac{1 - e^{(N+1)/N}}{1 - e^{1/N}} = e - 1.$$

Exactly the same way, we find

$$\lim_{N\to\infty} \sum_{j=1}^{2N} e^{k/N} = e^2 - 1.$$

So we have in all

$$\lim_{N\to\infty} \left(\frac{1}{N}\sum_{k=1}^{N} e^{k/N}\right)\left(\frac{1}{N}\sum_{j=1}^{2N} e^{j/N}\right) = (e-1)(e^2-1).$$

4.7 Set $\mu(x_i) = \mu_1$ if $x_i \in A$, and $\mu(x_i) = \mu_2$ if $x_i \in B$, and compute as follows:

$$\begin{aligned}
\overline{x_i}(C) &= \frac{\int_C x_i \mu(x_i)\,|d^n\mathbf{x}|}{M(C)} = \frac{1}{M(A)+M(B)}\left(\int_A \mu_1 x_i\,|d^n\mathbf{x}| + \int_B \mu_2 x_i\,|d^n\mathbf{x}|\right)\\
&= \frac{1}{M(A)+M(B)}\left(M(A)\frac{\int_A \mu_1 x_i\,|d^n\mathbf{x}|}{M(A)} + M(B)\frac{\int_B \mu_2 x_i\,|d^n\mathbf{x}|}{M(B)}\right)\\
&= \frac{1}{M(A)+M(B)}\Big(M(A)\overline{x_i}(A) + M(B)\overline{x_i}(B)\Big).
\end{aligned}$$

Therefore,

$$\overline{\mathbf{x}}(C) = \frac{M(A)\overline{\mathbf{x}}(A) + M(B)\overline{\mathbf{x}}(B)}{M(A)+M(B)}.$$

4.9 Choose ϵ, and write $X = \cup_{i=1}^{\infty} B_i$, where the B_i are pavable sets with $\sum_{i=1}^{\infty} \mathrm{vol}_n(B_i) < \epsilon/2$. For each i find a dyadic level N_i such that

$$U_{N_i}\mathbf{1}_{B_i} \le \mathrm{vol}_n(B_i) + \frac{\epsilon}{2^{i+1}}.$$

Then at the level N_i, the set B_i is covered by finitely many dyadic cubes with total volume at most $\mathrm{vol}_n(B_i) + \frac{\epsilon}{2^{i+1}}$. Now list first the dyadic cubes that cover B_1, then the dyadic cubes (at a different level) that cover B_2, and so on. The total volume of all these cubes is at most

$$\sum_{i=1}^{\infty}\left(\mathrm{vol}_n(B_i) + \frac{\epsilon}{2^{i+1}}\right) \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

4.11 a. We can write

$$\int_{-\sqrt{2}}^{\sqrt{2}}\left(\int_{x^2}^{2}\sin(x+y)\,dy\right)dx \quad \text{or} \quad \int_0^2\left(\int_{-\sqrt{y}}^{\sqrt{y}}\sin(x+y)\,dx\right)dy.$$

The first leads to

$$\int_{-\sqrt{2}}^{\sqrt{2}}\Big(\cos(x+x^2) - \cos(x+2)\Big)\,dx;$$

the second leads to

$$\int_0^2\Big(\cos(y-\sqrt{y}) - \cos(y+\sqrt{y})\Big)\,dy = \int_0^2 2\sin y\,\sin\sqrt{y}\;dy.$$

We believe that neither can be evaluated in elementary terms. Numerically, using Simpson's rule with 50 subdivisions, the integral comes out to $2.4314344\dots$. MAPLE gives a solution using Fresnel integrals; the numerical value obtained this way is $2.4314328529574138146\dots$.

b. This is much easier: The integral is

$$4\int_1^2\left(\int_1^2(x^2+y^2)\,dx\right)dy = 4\int_1^2\left(\int_1^2(x^2+y^2)\,dy\right)\,dx = \frac{56}{3}.$$

4.13 We have

$$\int_0^1\int_0^2 f\begin{pmatrix}x\\y\end{pmatrix}|dx\,dy| = \int_0^1\left[ax + \frac{bx^2}{2} + cxy\right]_0^2 dy = \int_0^1(2a+2b+2cy)\,dy$$
$$= \left[2ay + 2by + cy^2\right]_0^1 = 2a+2b+c.$$

The other integral is longer to compute:

$$\begin{aligned}\int_0^1\int_0^2\left(f\begin{pmatrix}x\\y\end{pmatrix}\right)^2|dx\,dy| &= \int_0^1\int_0^2(a^2+b^2x^2+c^2y^2+2abx+2acy+2bcxy)\,dx\,dy\\ &= \int_0^1\left(2a^2+\frac{8b^2}{3}+2c^2y^2+4ab+4acy+4bcy\right)dy\\ &= 2a^2+\frac{8b^2}{3}+\frac{2c^2}{3}+4ab+2ac+2bc.\end{aligned}$$

The Lagrange multiplier theorem says that at a minimum of the second integral, constrained so that the first integral is 1, there exists a number λ such that

$$[\,4a+4b+2c \quad \tfrac{16b}{3}+4a+2c \quad \tfrac{4c}{3}+2a+2b\,] = \lambda\,[\,2 \quad 2 \quad 1\,].$$

From this and the constraint we find the system of linear equations

$$\begin{aligned} 4a + 4b + 2c &= \frac{16b}{3} + 4a + 2c \\ 2a + 2b + c &= \frac{4c}{3} + 2a + 2b \\ 2a + 2b + c &= 1. \end{aligned}$$

This could be solved by row reduction, but it is easier to observe that the first equation says that $b = 0$ and the second equation says that $c = 0$. Thus $a = 1/2$ and the minimum is $\int_0^1 \int_0^2 \left(\frac{1}{2}\right)^2 |dx\, dy| = \frac{1}{2}$.

4.15 In order for the equality to be true for all polynomials of degree ≤ 3, it is enough that it should be true for the polynomials 1, x, x^2, x^3. Thus we want the four equations

$$1.\ \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = c(1+1) \qquad 3.\ \int_{-1}^1 x^2 \frac{dx}{\sqrt{1-x^2}} = c(u^2+u^2)$$

$$2.\ \int_{-1}^1 x \frac{dx}{\sqrt{1-x^2}} = c(u-u) \qquad 4.\ \int_{-1}^1 x^3 \frac{dx}{\sqrt{1-x^2}} = c(u^3-u^3).$$

Equations 2 and 4 are true for all c and all u, since both sides are 0. Equation 1 gives

$$2c = \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = [\arcsin x]_{-1}^1 = \pi, \quad \text{so } c = \pi/2.$$

Equation 3: Computing the integral (using integration by parts), we get

$$2cu^2 = \int_{-1}^1 -x \frac{-x\, dx}{\sqrt{1-x^2}} = \left[-x\sqrt{1-x^2}\right]_{-1}^1 + \int_{-1}^1 \sqrt{1-x^2}\, dx = \pi.$$

Substituting $c = \pi/2$ in $2cu^2 = \pi$ gives $\pi u^2 = \pi$, i.e., $u = \pm 1$.

4.17 $$[\mathbf{D}\det(A)]B = [1 \quad -2 \quad -3 \quad 1] \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = a - 2b - 3c + d.$$

$$\begin{aligned} \det A \ \mathrm{tr}(A^{-1}B) &= -5\,\mathrm{tr}\left(\begin{bmatrix} -.2 & .6 \\ .4 & -.2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) \\ &= -5(-.2a + .6c + .4b - .2d) = a - 2b - 3c + d. \end{aligned}$$

Solution 4.17: $[\mathbf{D}\det(A)]$ is not the derivative of $\det A = -5$. It is the derivative at $A = \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$ of the function $\det \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = ad - bc$: i.e., it is the row matrix

$$[d \quad -c \quad -b \quad a],$$

evaluated at $a = 1$, $b = 3$, $c = 2$, $d = 1$.

4.19 The easiest way to compute this integral (at least if one has already done Exercise 4.9.4) is to observe that the volume to be computed is the same as in Exercise 4.9.4, after scaling the x_1 coordinate by n, the x_2 coordinate by $n/2$, etc., until the coordinate x_n is scaled by $n/n = 1$. This gives the volume

$$\frac{1}{n!} \cdot \frac{n}{1} \cdot \frac{n}{2} \cdots \frac{n}{n} = \frac{n^n}{(n!)^2}.$$

Here is the solution to Exercise 4.9.4:

Let us call A_n the answer to the problem; we will try to set up an inductive formula for A_n.

By Fubini's theorem, we find

$$A_1 = \int_0^1 \int_0^{1-x_1} dx_2\, dx_1 = \frac{1}{2}, \quad A_2 = \int_0^1 \int_0^{1-x_1} \int_0^{1-x_1-x_2} dx_3\, dx_2\, dx_1 = \frac{1}{6},$$

and more generally

$$A_n = \int_0^1 \int_0^{1-x_1} \int_0^{1-x_1-x_2} \cdots \int_0^{1-x_1-\cdots-x_{n-1}} dx_n \dots dx_2\, dx_1.$$

We see that if we define

$$B_n(t) = \int_0^t \int_0^{t-x_1} \int_0^{t-x_1-x_2} \cdots \int_0^{t-x_1-\cdots-x_{n-1}} dx_n \dots dx_2\, dx_1, \qquad (1)$$

then

$$A_n = B_n(1).$$

Calculate as above $B_1(t) = t,\ B_2(t) = t^2/2,\ B_3(t) = t^3/6$. A natural guess is that $B_n(t) = t^n/n!$. We will show this by induction: a look at formula 1 will show you that

$$B_{n+1}(t) = \int_0^t B_n(t - x_1)\, dx_1.$$

Thus the following computation does the inductive step:

$$B_{n+1}(t) = \int_0^t B_n(t-x_1)\, dx_1 = -\left[\frac{(t-x_1)^{n+1}}{(n+1)!}\right]_0^t = \frac{t^{n+1}}{(n+1)!}.$$

Thus

$$A_n = B_n(1) = \frac{1}{n!}.$$

4.21 This curve is called a *cardioid* (as in "cardiac arrest"). It is shown in the margin. The area, by the change of variables theorem and Fubini's theorem, is

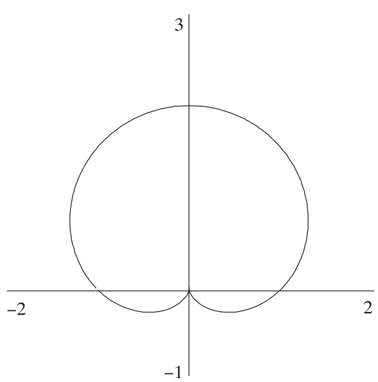


FIGURE FOR SOLUTION 4.21
A cardioid

$$\begin{aligned}\int_0^{2\pi} \left(\int_0^{1+\sin\theta} r\, dr \right) d\theta &= \int_0^{2\pi} \left[\frac{r^2}{2}\right]_0^{1+\sin\theta} d\theta = \int_0^{2\pi} \frac{(1+\sin\theta)^2}{2}\, d\theta \\ &= \left[\frac{\theta}{2} - \cos\theta\right]_0^{2\pi} + \int_0^{2\pi} \frac{(\sin\theta)^2}{2}\, d\theta \\ &= \pi + \frac{\pi}{2} = \frac{3\pi}{2}.\end{aligned}$$

4.23 a. In spherical coordinates, Q corresponds to

$$U = \left\{ \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \middle|\, 0 < r \le 1, \quad 0 \le \theta \le \frac{\pi}{2}, \quad 0 \le \varphi < \frac{\pi}{2} \right\}.$$

With this in mind, we can write our integral as an iterated integral:

$$\int_Q (x+y+z)|d^3\mathbf{x}| = \int_0^1 \left(\int_0^{\pi/2} \left(\int_0^{\pi/2} (r\cos\theta\cos\varphi + r\sin\theta\cos\varphi + r\sin\varphi) r^2 \cos\varphi \, d\varphi \right) d\theta \right) dr.$$

b. We can use a table of integrals to evaluate the more unpleasant sine and cosine portions of the integral, but it is easier to argue that the x, y, and z contributions to the total must be the same (symmetry), and then evaluate the z contribution:

$$\begin{aligned}\int_0^1\left(\int_0^{\pi/2}\left(\int_0^{\pi/2} r^3 \sin\varphi\cos\varphi\, d\varphi\right) d\theta\right) dr &= \int_0^1 \left(\int_0^{\pi/2} r^3 \left[\frac{\sin^2\varphi}{2}\right]_0^{\pi/2} d\theta\right) dr\\ &= \int_0^1 \left(\int_0^{\pi/2} \frac{r^3}{2}\, d\theta\right) dr\\ &= \int_0^1 \frac{\pi r^3}{4}\, dr = \frac{\pi}{16}.\end{aligned}$$

The value of our original integral is three times this: $3\pi/16$.

Solution 4.27, part c: This argument may seem opaque. Why take $x = 1/\sqrt{2}$? Since we were trying to find a point in the complement of a set of measure 0, why did we need a number with special properties?

These are hard questions to answer. It should be clear that the convergence of the series depends on x being poorly approximated by rational numbers, in the sense that if you want to get close to x you will have to use a large denominator. In fact, part b says that most real numbers are like that.

But actually finding one is a different matter. (If you are asked to "pick a number, any number", you will most likely come up with a rational number, which has probability 0.) It often happens that we can prove that some property of numbers (like being transcendental) is shared by almost all numbers, without being able to give a single example, at least easily (see Section 0.6). The approximation of numbers by rationals is called *Diophantine analysis*, and one of its first results is that quadratic irrationals, like $1/\sqrt{2}$, are poorly approximable; in part c we present a proof.

4.25 By symmetry, the center of gravity is on the z-axis. The z-coordinate $\overline{z}$ of the center of gravity is

$$\overline{z} = \frac{\int_A z\,|dx\,dy\,dz|}{\int_A |dx\,dy\,dz|} = \frac{\int_0^1 z(\pi z)\,dz}{\int_0^1 (\pi z)\,dz} = \frac{1/3}{1/2} = \frac{2}{3}.$$

4.27 a. The function $1/\sqrt{|x-a|}$ is L-integrable over $[0,1]$ for every $a \in [0,1]$ (or even outside):

$$\int_0^1 \frac{1}{\sqrt{|x-a|}}\,dx \le \int_{a-1}^{a+1} \frac{1}{\sqrt{|x-a|}}\,dx = \int_{-1}^1 \frac{1}{\sqrt{|x|}}\,dx = 4.$$

Then since

$$\sum_{k=1}^\infty \frac{1}{2^k}\int_0^1 \frac{1}{\sqrt{|x-a_k|}}\,dx \le 4\sum_{k=1}^\infty \frac{1}{2^k} \le 4,$$

Theorem 4.11.17 says that the series of functions $\sum_{k=1}^\infty \frac{1}{2^k}\frac{1}{\sqrt{|x-a_k|}}$ converges for almost all x, and that f is L-integrable on $[0,1]$.

b. We did this in part a: "on a set of measure 0" is the same as "for almost all x". Note, however, that when x is rational, the series does not converge, since we have a 0 in the denominator for one of the terms. So "almost all x" does not include any rational numbers.

c. Let us take the rationals in $[0,1]$ in the following order:

$$a_1 = 0,\ a_2 = 1,\ a_3 = \frac{1}{2},\ a_4 = \frac{1}{3},\ a_5 = \frac{2}{3},\ a_6 = \frac{1}{4},\ a_7 = \frac{3}{4},\ a_8 = \frac{1}{5},\ a_9 = \frac{2}{5},\ldots,$$

taking first all those with denominator 1, then those with denominator 2, then those with denominator 3, etc. Note that if $a_k = p/q$ for p, q integers, then $k \geq 2q-3$, since there are at least two numbers for any denominator except 2 (accounting for the -3).

Now take $x = 1/\sqrt{2}$. Notice (this is the clever step) that

$$2q^2 \left|\frac{1}{\sqrt{2}} - \frac{p}{q}\right| \left|\frac{1}{\sqrt{2}} + \frac{p}{q}\right| = |q^2 - 2p^2| \geq 1$$

since it is a nonzero integer. Hence for any coprime integers p, q with $p/q \in [0,1]$, we have

$$\left|x - \frac{p}{q}\right| = \left|\frac{1}{\sqrt{2}} - \frac{p}{q}\right| \geq \frac{1}{2q^2\left|\frac{1}{\sqrt{2}} + \frac{p}{q}\right|} \geq \frac{1}{4q^2}, \tag{1}$$

The second inequality in equation 1: since

$$1/\sqrt{2} < 1 \quad \text{and} \quad p/q \leq 1,$$

their sum is ≤ 2.

so

$$\frac{1}{\sqrt{|x-a_k|}} \leq 2q.$$

This leads to

$$\sum_{k=1}^{\infty} \frac{1}{2^k} \frac{1}{\sqrt{|x-a_k|}} \leq \sum_{q=1}^{\infty}\sum_{p=0}^{q} \frac{1}{2^{2q-3}}\ 2q = \sum_{q=1}^{\infty} \frac{2q(q+1)}{2^{2q-3}}. \tag{2}$$

Equation 2: the sum $\sum_{p=0}^{q}$ accounts for all the rational numbers with denominator q; for all q except $q=1$, there are at most $q-1$ of them. There are exactly $q-1$ of them if q is prime. This sum becomes multiplication by $q+1$ in the next step. The case where a quantity being summed has no index matching the index of the sum is discussed in a margin note in Section 0.1.

This series is easily seen to be convergent by the ratio test.

4.29 Write $a = a_1 + ia_2$ and $b = b_1 + ib_2$, so that $T(u) = au + b\bar{u}$ can be written

$$\begin{aligned} T(x+iy) &= (a_1+ia_2)(x+iy) + (b_1+ib_2)(x-iy) \\ &= a_1x - a_2y + b_1x + b_2y + i(a_1y + a_2x - b_1y + b_2x). \end{aligned} \tag{3}$$

we are identifying $\mathbb{C}$ with $\mathbb{R}^2$ in the standard way, with $x+iy$ written $\begin{pmatrix} x \\ y \end{pmatrix}$, so equation 3 is equivalent to the matrix multiplication

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a_1+b_1 & -a_2+b_2 \\ a_2+b_2 & a_1-b_1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Thus

$$\det T = (a_1+b_1)(a_1-b_1) - (a_2+b_2)(-a_2+b_2) = a_1^2 - b_1^2 + a_2^2 - b_2^2 = |a|^2 - |b|^2.$$

Now we show the result for the norm. Any complex number of absolute value 1 can be written $|e^{i\theta}| = 1$. This, with the triangle inequality, gives

$$\|T\| = \sup_{\theta} \left|ae^{i\theta} + be^{-i\theta}\right| = |a+b| \leq |a| + |b|.$$

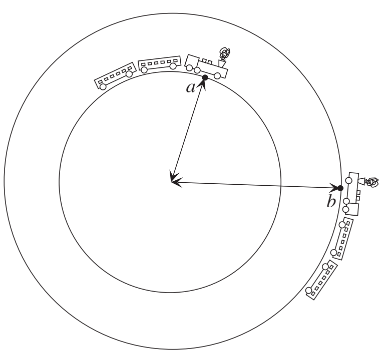


FIGURE FOR SOLUTION 4.29
One train represents $ae^{i\theta}$, the other $be^{-i\theta}$. There exists a value of θ for which the trains line up, i.e., for which $ae^{i\theta}$ and $be^{-i\theta}$ are multiples of each other by a positive real number.

Think of two circles centered at the origin, one with radius $|a|$, the other with radius $|b|$: $ae^{i\theta}$ travels around its circle in one direction, and $be^{-i\theta}$ travels around its circle in the other direction, as illustrated in the figure at left.

Clearly there is a value of θ where the two are on the same half-line from the origin through the circles; for this value of θ, we see that $ae^{i\theta}$ and $be^{-i\theta}$ are multiples of each other by a positive real number, giving

$$\|T\| = |a+b| = |a| + |b|.$$

4.31 a. Set $\begin{cases} u = x \\ v = xyz \\ w = y/(xz), \end{cases}$
so that A becomes the parallelepiped

$$1 \le u \le b, \quad 1 \le v \le a, \quad 1 \le w \le c.$$

b. We need to express x, y, z in terms of u, v, w:

$$\begin{aligned} x &= u \\ y &= \sqrt{vw} \\ z &= \frac{v}{xy} = \frac{v}{u\sqrt{vw}} = \frac{1}{u}\sqrt{\frac{v}{w}}. \end{aligned}$$

Set

$$\Phi\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} u \\ \sqrt{vw} \\ \frac{1}{u}\sqrt{\frac{v}{w}} \end{pmatrix}.$$

Then

$$\left[\mathbf{D}\Phi\begin{pmatrix} u \\ v \\ w \end{pmatrix}\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2}\sqrt{\frac{w}{v}} & \frac{1}{2}\sqrt{\frac{v}{w}} \\ -\frac{1}{u^2}\sqrt{\frac{v}{w}} & \frac{1}{2u}\sqrt{\frac{1}{vw}} & -\frac{1}{2u}\frac{\sqrt{v}}{w^{3/2}} \end{bmatrix}$$

and

$$\det\left[\mathbf{D}\Phi\begin{pmatrix} u \\ v \\ w \end{pmatrix}\right] = -\frac{1}{4u}\sqrt{\frac{w}{v}}\frac{\sqrt{v}}{w^{2/3}} - \frac{1}{4u}\sqrt{\frac{v}{w}}\sqrt{\frac{1}{vw}} = -\frac{1}{2uw}.$$

c.

$$\int_1^b \int_1^a \int_1^c \left| -\frac{1}{2uw} \right| dw\, dv\, du = \frac{1}{2}(a-1)\left(\int_1^b \frac{du}{u}\right)\left(\int_1^c \frac{dw}{w}\right) = \frac{1}{2}(a-1)\ln b \ln c.$$

4.33 Let a_i be a list of the rationals in $[0,1]$: for instance

$$0,\ 1,\ 1/2,\ 1/3,\ 2/3, \dots,$$

and let X be the open set in $\mathbb{R}^2$

$$X = \bigcup_{i=1}^{\infty}\left(a_i - \frac{1}{2^{10+i}},\ a_i + \frac{1}{2^{10+i}}\right).$$

The set X is open, hence a countable disjoint union of intervals. and it contains all the rationals, so it is dense. But the sum of the lengths of the intervals is $< 1/2^9$, hence tiny.

Let $f : X \to \mathbb{R}$ be a continuous function that maps each connected component of X bijectively to $(-1, 1)$.

We claim that the graph of f does not have volume.

Indeed, for any N the set of cubes of $\mathcal{D}_N(\mathbb{R})$ in $[0,1]$ that completely contain one of the intervals of X has total length at least $1-1/2^9$. To cover the part of the graph above these cubes requires $2\cdot 2^N$ squares of $\mathcal{D}_N(\mathbb{R}^2)$. Thus the upper sum of $\mathbf{1}_{\Gamma_f}$ is at least $2(1-1/2^9)$, but the lower sum is 0, since every square of $\mathcal{D}_N(\mathbb{R}^2)$ contains points not on the graph. So the graph does not have volume.

It does have measure 0, since it is the countable union of the graphs of the functions defined on each component of X, and each of these does have measure (in fact, volume) 0.

4.35 Set

$$A_k \stackrel{\text{def}}{=} \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; \frac{1}{2^{k+1}} \le |\mathbf{x}| < \frac{1}{2^k} \right\}$$

Then

$$|\mathbf{x}|^p \mathbf{1}_{B_1(\mathbf{0})} = \sum_{k=0}^{\infty} |\mathbf{x}|^p \mathbf{1}_{A_k},$$

where the maps $|\mathbf{x}|^p\mathbf{1}_{A_k}$ are Riemann integrable and nonnegative. To know that $|\mathbf{x}|^p$ is integrable for $B_1(\mathbf{0})$, we need to know that the series of integrals of the $|\mathbf{x}|^p\mathbf{1}_{A_k}$ converges. The map $\Phi_k : \mathbf{x} \mapsto 2^{-k}\mathbf{x}$ maps A_0 to A_k, and $\det[\mathbf{D}\Phi_k] = 2^{-nk}$. Thus

$$\int_{A_k} |\mathbf{x}|^p \, |d^n\mathbf{x}| = 2^{-nk} \int_{A_0} |2^{-k}\mathbf{x}|^p \, |d^n\mathbf{x}| = 2^{-k(n+p)} \int_{A_0} |\mathbf{x}|^p |d^n\mathbf{x}|.$$

Thus

$$\int_{B_1(\mathbf{0})} |\mathbf{x}|^p |d^n\mathbf{x}| = \left(\sum_{k=0}^{\infty} 2^{-k(n+p)} \right) \int_{A_0} |\mathbf{x}|^p |d^n\mathbf{x}|.$$

The series is a geometric series and converges if and only if $n+p>0$, i.e., if $p>-n$.

Solutions for Chapter 5

5.1.1 Set $T = [\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3]$; then $T^\top T = \begin{bmatrix} 3 & 2 & 3 \\ 2 & 6 & 4 \\ 3 & 4 & 6 \end{bmatrix}$ and $\det\big(T^\top T\big) = 30$, so the 3-dimensional volume of the parallelogram is $\sqrt{30}$.

5.1.3 Here are two solutions.

First solution

Set $T = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k]$. Since the vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k$ are linearly dependent, $\operatorname{rank} T < k$. Further, $\operatorname{img} T^\top T \subset \operatorname{img} T^\top$, so

$$\operatorname{rank} T^\top T \le \operatorname{rank} T^\top \underbrace{=}_{\text{Prop. 2.5.11}} \operatorname{rank} T < k.$$

Since $T^\top T$ is a $k \times k$ matrix with rank $< k$, it is not invertible, hence its determinant is 0, so

$$\operatorname{vol}_k P(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k) = \sqrt{\det T^\top T} = 0.$$

Second solution

Since the vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k$ are linearly dependent, there exist numbers $a_1, \dots, a_k$ not all 0 such that $a_1\vec{\mathbf{v}}_1 + \dots + a_k\vec{\mathbf{v}}_k = \mathbf{0}$, i.e.,

$$[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k] \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = \mathbf{0}.$$

Set $T = [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k]$. Since $T^\top T\mathbf{a} = \mathbf{0}$, it follows that $\ker\,(T^\top T) \ne \mathbf{0}$, so $T^\top T$ is not invertible, and $\det(T^\top T) = 0$, so

$$\operatorname{vol}_k P(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_k) = \sqrt{\det T^\top T} = 0.$$

Solution 5.1.5: Here is another approach. By equation 5.1.9, the matrix $T^\top T$ is symmetric, so equation 1 shows that $T^\top T$ is a symmetric matrix representing a quadratic form taking only nonnegative values, i.e., its signature is $(l, 0)$ for some $l \le k$. So $T^\top T$ has l positive eigenvalues, and the others are zero, by Theorem 3.7.16. Since the determinant of a triangular matrix is the product of the eigenvalues (Theorem 4.8.9), and a symmetric matrix is diagonalizable (the spectral theorem), and the determinant is basis independent (Theorem 4.8.6), it follows that $\det T^\top T \ge 0$.

5.1.5 Note first that if $\vec{\mathbf{v}}$ is a nonzero vector in $\mathbb{R}^k$,

$$(T^\top T\vec{\mathbf{v}}) \cdot \vec{\mathbf{v}} = \big(T^\top (T\vec{\mathbf{v}})\big)^\top \vec{\mathbf{v}} = (T\vec{\mathbf{v}})^\top T\vec{\mathbf{v}} = T\vec{\mathbf{v}} \cdot T\vec{\mathbf{v}} \ge 0. \tag{1}$$

Denote by A the $k \times k$ matrix $T^\top T$ and let I be the $k \times k$ identity matrix. Consider the matrix $tA + (1-t)I$ for $t \in [0,1]$, which we can think of as A (when $t = 1$) being transformed to I (when $t = 0$). For $\vec{\mathbf{v}} \ne \mathbf{0}$ and $0 \le t < 1$, we have

$$\big(tA + (1-t)I\big)\vec{\mathbf{v}} \cdot \vec{\mathbf{v}} = t\underbrace{A\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}_{\ge 0} + \underbrace{(1-t)}_{>0}\underbrace{\vec{\mathbf{v}} \cdot \vec{\mathbf{v}}}_{>0} > 0.$$

This implies that, for $0 < t \le 1$, $\ker\big(tA + (1-t)I\big) = \mathbf{0}$ and thus that $\det\big(tA + (1-t)I\big)$ is never 0 when $0 \le t < 1$. Since when $t = 0$, we have $\det\big(tA + (1-t)I\big) = 1$, and when $t = 1$, $\det\big(tA + (1-t)I\big) = \det A$, it follows that $\det A \geq 0$. (This uses the continuity of determinant, which is a polynomial function.)

Nathaniel Schenker suggests a third solution to Exercise 5.1.5:

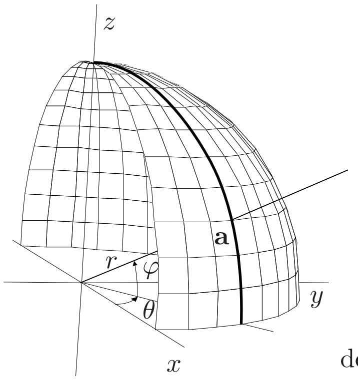
By the singular value decomposition (Theorem 3.8.1), $T = PDQ^\top$, where P and Q are, respectively $n \times n$ and $k \times k$ orthogonal matrices, and D is an $n \times k$ nonnegative rectangular diagonal matrix. Then

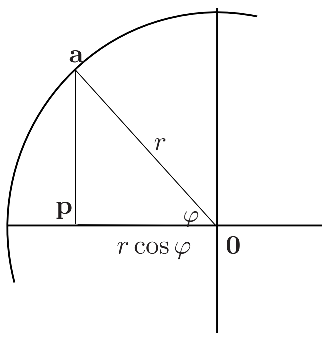
$$T^\top T = QD^\top P^\top PDQ^\top = QD^\top I_n DQ^\top = QD^\top DQ^\top,$$

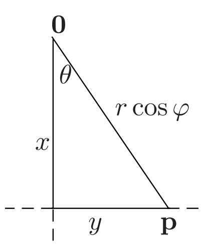
where the first equality is Theorem 1.2.17 and I_n is the $n \times n$ identity matrix. Then

$$\det(T^\top T) = \det Q \det(D^\top D) \det Q^\top = \det Q \det(D^\top D)(\det Q)^{-1} = \det(D^\top D),$$

where the first equality is Theorem 4.8.4 and the second is Corollary 4.8.5. Since $D^\top D$ is diagonal with nonnegative entries, Theorem 4.8.9 implies that $\det(D^\top D) \geq 0$.







TOP: In spherical coordinates, a point **a** is specified by its distance from the origin (r), its longitude (θ), and its latitude (φ). MIDDLE: This triangle lives in a plane perpendicular to the (x, y)-plane; you might think of slicing an orange in half vertically. BOTTOM: This triangle lives in the (x, y)-plane.

Note on parametrizations

A number of exercises in Sections 5.2–5.4 involve finding parametrizations. The "small catalog of parametrizations" in Section 5.2 should help. In many cases we adapt one of the standard parametrizations (spherical, polar, cylindrical) to the problem at hand. Note also the advice given in Solution 5.1: *To find a parametrization for some region, we ask, "how should we describe where we are at some point in the region?"* Thus a first step is often to draw a picture.

If you memorized the spherical, cylindrical, and polar coordinates maps without thinking about why those are good descriptions of "where we are at some point", this is a good time to think about how these standard change of coordinates maps follow from high school trigonometry. Consider, for instance, the spherical coordinates map

$$S : \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta\cos\varphi \\ r\sin\theta\cos\varphi \\ r\sin\varphi \end{pmatrix}$$

where r is the radius of the sphere, θ is longitude, and φ is latitude, as shown in the top figure at left (which you already saw as Figure 4.10.4).

The middle figure at left shows the triangle whose vertices are a point **a** on the sphere, the origin, and the projection **p** of **a** onto the (x, y)-plane. Since $\cos\varphi$ is adjacent side over hypotenuse, the length of the side between **0** and **p** is $r\cos\varphi$. In the bottom triangle, $\cos\theta$ is adjacent side over hypotenuse $r\cos\varphi$, so the adjacent side (the side lying on the x-axis) is $r\cos\theta\cos\varphi$. This gives the first entry in the spherical coordinates map; it is the "x-coordinate" of the point **p**, which is the same as the x-coordinate of the point **a**. Since $\sin\theta$ is opposite side over hypotenuse, the "y-coordinate" of **p** and of **a** is $r\sin\theta\cos\varphi$. The z-coordinate of **a** should be the distance between **a** and **p**, as shown in the middle figure; that distance is indeed $r \sin\varphi$.

Trigonometry was invented by the ancient Greeks for the purpose of doing these computations.

5.2.1 Condition 3 of Definition 5.2.3 requires that $\gamma : (U - X) \to M$ be one to one, of class C^1, with locally Lipschitz derivative. To show that it is one to one, we must show that if $\gamma\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix} = \gamma\begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix}$, then $\begin{pmatrix} r_1 \\ \theta_1 \end{pmatrix} = \begin{pmatrix} r_2 \\ \theta_2 \end{pmatrix}$. Set

$$\begin{aligned} r_1 \cos\theta_1 &= r_2 \cos\theta_2 \\ r_1 \sin\theta_1 &= r_2 \sin\theta_2 \\ r_1 &= r_2. \end{aligned}$$

Since $r_1 = r_2$ by the third equation, $\cos\theta_1 = \cos\theta_2$ and $\sin\theta_1 = \sin\theta_2$. Since we are restricting θ to $0 < \theta < 2\pi$, this implies that $\theta_1 = \theta_2$. So γ is one to one on $U - X$. Moreover, it is clearly not just C^1 but C^∞, and Proposition 2.8.9 says that the derivative of a C^2 mapping is Lipschitz, so γ has locally Lipschitz derivative.

The mapping γ is one to one on $U - X$ but not on U.

The map γ is C^∞: we can differentiate it as many times as we like.

For condition 4, we must show that $[\mathbf{D}\gamma(\mathbf{u})]$ is one to one for all $\mathbf{u}$ in $U - X$. We have

$$\left[\mathbf{D}\gamma\begin{pmatrix} r \\ \theta \end{pmatrix}\right] = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \\ 1 & 0 \end{bmatrix}.$$

It is one to one if $x = 0$, $y = 0$ is the only solution to

$$\left[\mathbf{D}\gamma\begin{pmatrix} r \\ \theta \end{pmatrix}\right]\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - yr\sin\theta \\ x\sin\theta + yr\cos\theta \\ x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Since $x = 0$, we have $yr\cos\theta = yr\sin\theta = 0$; since in $U - X$ we have $r > 0$, we can divide by r to get $y\cos\theta = y\sin\theta = 0$. Since $\sin\theta$ and $\cos\theta$ do not simultaneously vanish, this implies that $y = 0$.

5.2.3 The map S is not a parametrization by Definition 3.1.18, but it is a parametrization by Definition 5.2.3.

It is not difficult to show that S is equivalent to the spherical coordinates map using latitude and longitude (Definition 4.10.6). Any point $S(\mathbf{x})$ is on a sphere of radius r, since

$$r^2 \sin^2\varphi \cos^2\theta + r^2 \sin^2\varphi \sin^2\theta + r^2 \cos^2\varphi = r^2.$$

Since $0 \le r \le \infty$, the mapping S is onto.

The angle φ tells what latitude a point is on. Going from 0 to π, it covers every possible latitude. The polar angle θ tells what longitude the point is on; going from 0 to 2π, it covers every possible longitude.

However, S is not a parametrization by Definition 3.1.18, because it is not one to one. Trouble occurs in several places. If $r = 0$, then $S(\mathbf{x})$ is the origin, regardless of the values of θ and φ. For θ, trouble occurs when $\theta = 0$ and $\theta = 2\pi$; for fixed r and φ, these two values of θ give the same point.

For φ, trouble occurs at 0 and π; if $\varphi = \pi$, the point is the south pole of a sphere of radius r, regardless of θ, and if $\varphi = 0$, the point is the north pole of a sphere of radius r, regardless of θ.

By Definition 5.2.3, S is a parametrization, since the trouble occurs only on a set of 3-dimensional volume 0. In the language of Definition 5.2.3,

$$M = \mathbb{R}^3, \qquad U = \overbrace{[0,\infty)}^{\text{for } r} \times \overbrace{[0,2\pi]}^{\text{for } \theta} \times \overbrace{[0,\pi]}^{\text{for } \varphi},$$

$$X = \Big(\underbrace{\{0\} \times [0,2\pi] \times [0,\pi]}_{\text{trouble when } r=0}\Big) \cup \Big(\underbrace{[0,\infty) \times \{0,2\pi\} \times [0,\pi]}_{\text{trouble when } \theta=0 \text{ or } \theta=2\pi}\Big)$$
$$\cup \Big(\underbrace{[0,\infty] \times [0,2\pi] \times \{0,\pi\}}_{\text{trouble when } \varphi=0 \text{ or } \varphi=\pi}\Big).$$

Trouble occurs on a union of five surfaces, which you can think of as five sides of a box whose sixth side is at infinity. The base of the box lies in the plane where $r = 0$; the sides of the base have length π and 2π. The base represents the set "trouble when $r = 0$." Two parallel sides of the box stretching to infinity represent "trouble when $\theta = 0$ or $\theta = 2\pi$". The other set of parallel sides represents "trouble when $\varphi = 0$ or $\varphi = \pi$". By Proposition 4.3.5 and Definition 5.2.1, these surfaces have $\text{vol}_3 = 0$.

By the dimension formula, the derivative is invertible if it is one to one.

Next let us check (condition 4 of Definition 5.2.3), that the derivative is one to one for all $\mathbf{u}$ in $U - X$. This will be true (Theorem 4.8.3) if and only if the determinant of the derivative is not 0. The determinant is

$$\det \left[\mathbf{D}S \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \right] = \det \begin{bmatrix} \sin\varphi\cos\theta & -r\sin\varphi\sin\theta & r\cos\varphi\cos\theta \\ \sin\varphi\sin\theta & r\sin\varphi\cos\theta & r\cos\varphi\sin\theta \\ \cos\varphi & 0 & -r\sin\varphi \end{bmatrix} = -2r^2\sin\varphi,$$

which is 0 only if $r = 0$, or $\varphi = 0$, or $\varphi = \pi$, which are not in $U - X$.

Polynomials are continuous by Corollary 1.5.31. Theorem 1.5.29 discusses combining continuous functions

Next, we will show that $S : (U - X) \to \mathbb{R}^3$ is of class C^1 with locally Lipschitz derivative. We know the first derivatives exist, since we just computed the derivative. They are continuous, since they are all polynomials in r, sine, and cosine, which are continuous. In fact, $S : (U - X) \to \mathbb{R}^3$ is C^∞, since its derivatives of all order are polynomials in r, sine, and cosine, and are thus continuous. We do not need to check anything about Lipschitz conditions because Proposition 2.8.9 says that the derivative of a C^2 function is Lipschitz.

Finally, we need to show that $S(X)$ has 3-dimensional volume 0. If $r = 0$, then $S\begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix}$ is the origin, whatever the values of θ and φ. If $\theta = 0$ we have $S\begin{pmatrix} r \\ 0 \\ \varphi \end{pmatrix} = \begin{pmatrix} r\sin\varphi \\ 0 \\ r\cos\varphi \end{pmatrix}$, which is a surface in the (x,z)-plane. If $\theta = 2\pi$, we get the same surface. If $\varphi = 0$, we get $\begin{pmatrix} 0 \\ 0 \\ r \end{pmatrix}$, the z-axis going from 0 to ∞; if $\varphi = \pi$ $\begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix}$, we get the z-axis going from 0 to $-\infty$.

5.2.5 The set

$$U = \bigcup_{i=1}^{\infty} \left(a_i - \frac{1}{2^{i+k}}, a_i - \frac{1}{2^{i+k}} \right)$$

is an open subset of $\mathbb{R}$, hence indeed a 1-dimensional manifold. The pathological property of U is that its boundary has positive length, so that its indicator function is not integrable.[5]

But U is simply a countable union of disjoint open intervals. (Every open subset of $\mathbb{R}$ is a countable union of open intervals.) List these intervals $I_1, I_2, \dots$ in order of length (which is *not* the order in which they appear in $\mathbb{R}$). Let the length of I_i be l_i, and let

$$J_i = \left(\sum_{j=1}^{i-1} l_i, \sum_{j=1}^{i} l_i \right),$$

so that J_i is another interval of length l_i. Now let $J = \cup_{i=1}^{\infty} J_i$, and define $\gamma : J \to U$ to be the map that translates J_i onto I_i. This is a parametrization according to Definition 5.2.3. Indeed, let X be empty, and check the conditions:

1. $\gamma(J) = U$, by definition;
2. $\gamma(J - X) = U$, since X is empty;
3. $\gamma : J - X \to U$ is one to one, of class C^1, with Lipschitz derivative. These are local conditions, and γ is locally a translation.
4. $\gamma'(x) \neq 0$ for all $x \in J$, since $\gamma'(x) = 1$ everywhere.
5. $\gamma(X)$ has 1-dimensional volume 0, since X is empty.

Solution 5.2.5: The point of showing that U can be parametrized is that we will be able to compute integrals over U using the parametrization. We can't integrate a function f over U, since the upper and lower sums

$$U_N(f\mathbf{1}_U) \text{ and } L_N(f\mathbf{1}_U)$$

do not have a common limit. But we can integrate $f \circ \gamma$ over J, since the boundary of J is the increasing sequence $i \mapsto a_i$ of numbers (including its limit) defined by

$$a_i = \sum_{j=1}^{i} l_j, \; i = 0, 1, \dots, \infty.$$

Thus the boundary of J has 1-dimensional volume 0.

5.2.7 a. Let $\pi : \mathbb{R}^n \to \mathbb{R}^k$ denote the projection of $\mathbb{R}^n$ using the corresponding k coordinate functions. Choose $\epsilon > 0$, and N so large that

$$\sum_{\substack{C \in \mathcal{D}_N(\mathbb{R}^n) \\ C \cap X \neq \phi}} \Big(\underbrace{\frac{1}{2^N}}_{\text{sidelength of } C} \Big)^k < \epsilon.$$

[5] The boundary of U is the complement of U in $[0,1]$: by Definition 1.5.10, a point x is in the boundary if each neighborhood of x intersects both U and its complement. Since the length of $[0,1]$ is 1, and the sum of the lengths of the intervals making up U is ϵ, the boundary $[0,1] - U$ has positive length.

Then

$$\epsilon > \sum_{\substack{C \in \mathcal{D}_N(\mathbb{R}^n) \\ C \cap X \neq \emptyset}} \left(\frac{1}{2^N}\right)^k \geq \sum_{\substack{C_1 \in \mathcal{D}_N(\mathbb{R}^k) \\ C_1 \cap \pi(X) \neq \emptyset}} \left(\frac{1}{2^N}\right)^k,$$

since for every k-dimensional cube $C_1 \in \mathcal{D}_N(\mathbb{R}^k)$ such that $C_1 \cap \pi(X) \neq \emptyset$, there is at least one n-dimensional cube

$$C \in \mathcal{D}_N(\mathbb{R}^n) \text{ with } C \subset \pi^{-1}(C_1) \text{ such that } C \cap X \neq \emptyset.$$

Thus $\mathrm{vol}_k(\pi(X)) < \epsilon$ for any $\epsilon > 0$.

b. If you tried to find an unbounded subset of $\mathbb{R}^2$ of length 0 whose projection onto the x-axis has positive length, you will have had trouble. What can be found is such a subset whose projection does not have length. For example, the subset of $\mathbb{R}^2$ consisting of points with x-coordinate p/q and y-coordinate q, with p and q integers satisfying $0 \leq \frac{p}{q} \leq 1$, is unbounded, since q can be arbitrarily large; it has length 0, since it consists of isolated points. Its projection onto the x-axis is the set of rational numbers in $[0,1]$, which (as we saw in Example 4.4.3) does not have defined volume.

5.3.1 a. Going back to Cartesian coordinates, set $x(t) = r(t)\cos\theta(t)$ and $y(t) = r(t)\sin\theta(t)$:

$$\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} r(t)\cos\theta(t) \\ r(t)\sin\theta(t) \end{pmatrix}.$$

So the formula $l = \int_a^b |\vec{\gamma}'(t)|\,dt$ (equation 5.3.4) gives

$$\begin{aligned} l &= \int_a^b \sqrt{\big((r\cos\theta)'(t)\big)^2 + \big((r\sin\theta)'(t)\big)^2}\,dt \\ &= \int_a^b \sqrt{\Big(r'(t)\cos(\theta(t)) - r(t)\sin(\theta(t))\theta'(t)\Big)^2 + \Big(r'(t)\sin(\theta(t)) + r\cos(\theta(t))\theta'(t)\Big)^2}\,dt \\ &= \int_a^b \sqrt{\big(r'(t)\big)^2 + \big(r(t)\big)^2\big(\theta'(t)\big)^2}\,dt. \end{aligned}$$

When computing the first entry of $\gamma'(t)$, remember that the derivative of $\cos\theta(t)$ is the derivative at t of the composition $\cos\circ\theta$, computed using the chain rule: $(\cos\circ\theta)'(t) = \cos'(\theta(t))\theta'(t) = -\sin(\theta(t))\theta'(t)$. Similarly, for the second entry, the derivative of $\sin\theta(t)$ is the derivative at t of the composition $\sin\circ\theta$, giving

$$(\sin\circ\theta)'(t) = \sin'(\theta(t))\theta'(t) = \cos(\theta(t))\theta'(t).$$

b. The spiral is the curve parametrized by $\gamma(t) = \begin{pmatrix} e^{-\alpha t}\cos t \\ e^{-\alpha t}\sin t \end{pmatrix}$, whose derivative is $\vec{\gamma}'(t) = \begin{bmatrix} -e^{-\alpha t}\sin t - \alpha e^{-\alpha t}\cos t \\ e^{-\alpha t}\cos t - \alpha e^{-\alpha t}\sin t \end{bmatrix}$, so the length between

$t = 0$ and $t = b$ is

$$\int_0^b |\vec{\gamma}'(t)|\, dt = \int_0^b \sqrt{\alpha^2 e^{-2\alpha t} + e^{-2\alpha t}}\, dt = \int_0^b \sqrt{1+\alpha^2} e^{-\alpha t}\, dt$$
$$= \sqrt{\alpha^2+1}\, \frac{(e^{-\alpha b} - 1)}{-\alpha}.$$

Solution 5.3.1, part b: How do we know this curve is a spiral? Each point of the curve is known by polar coordinates; at time t the distance from the point to the origin is $e^{-\alpha t}$ and the polar angle is t. The beginning point, at $t = 0$, is the point $\begin{pmatrix}1\\0\end{pmatrix}$ in the (x, y)-plane – the point with coordinates

$$x = e^0 \cos 0, \quad y = e^0 \sin 0.$$

As t increases, the polar angle increases and the distance from the origin decreases; as $t \to \infty$, we have $e^{-\alpha t} \to 0$.

The limit of this length as $\alpha \to 0$ is b: expressing $e^{-\alpha b}$ as its Taylor polynomial (equation 3.4.5) gives

$$\begin{aligned}\lim_{\alpha\to 0} \sqrt{\alpha^2+1}\, \frac{(e^{-\alpha b} - 1)}{-\alpha} &= \lim_{\alpha\to 0} \sqrt{\alpha^2+1}\, \frac{\big((1 - b\alpha + \frac{b^2\alpha^2}{2} - \cdots) - 1\big)}{-\alpha}\\ &= \lim_{\alpha\to 0} \sqrt{\alpha^2+1}\, \Big(b + \alpha(\text{higher degree terms})\Big)\\ &= b.\end{aligned}$$

At $\alpha = 0$, the distance from the origin is 1, so the curve is a circle of radius 1 rather than a narrowing spiral.

c. For every 2π increment in t the spiral turns around the origin once: $\theta(t) = \theta(t + 2\pi)$. So the spiral turns infinitely many times about the origin as $t \to \infty$. The length does not tend to infinity, since

$$\lim_{b\to\infty} l = \lim_{b\to\infty} \frac{\sqrt{\alpha^2+1}}{\alpha}\,(1 - e^{-\alpha b}) = \frac{\sqrt{\alpha^2+1}}{\alpha}.$$

5.3.3 a. Since

$$x(t) = r(t)\cos\varphi(t)\cos\theta(t),\ \ y(t) = r(t)\cos\varphi(t)\sin\theta(t),$$
$$z(t) = r(t)\sin\varphi(t),$$

Solution 5.3.3: In the exercise, the order in which we listed the variables θ and φ in the domain is reversed, compared to Definition 4.10.6. You should think that here we have

$$S : \begin{pmatrix} r\\ \varphi\\ \theta\end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta\cos\varphi\\ r\sin\theta\cos\varphi\\ r\sin\varphi\end{pmatrix}.$$

we have

$$\begin{aligned} l = \int \Big(&(r'\cos\varphi\cos\theta - r\sin\varphi\cos\theta\ \varphi' - r'\cos\varphi\sin\theta\ \theta')^2\\ &+ (r'\cos\varphi\sin\theta - r\sin\varphi\sin\theta\ \varphi' + r'\cos\varphi\cos\theta\ \theta')^2\\ &+ (r'\sin\varphi - r\cos\varphi\ \varphi')^2\Big)^{1/2} dt\\ = \int &\sqrt{(r')^2 + r^2\ (\varphi')^2 + r^2\cos^2\varphi\ (\theta')^2}\, dt.\end{aligned}$$

b. We get the integral

$$\int_0^a \sqrt{(-\sin t)^2 + (\cos t)^2 + (\cos t)^2(\cos t)^2 \frac{1}{(\cos t)^4}}\, dt = \int_0^a \sqrt{2}\ dt = a\sqrt{2}.$$

Solution 5.3.3, part b: Even though $\theta(t) = \tan t$ tends to infinity, the curve does not become long.

5.3.5 a. The map

$$\gamma : \begin{pmatrix} r\\ \theta\end{pmatrix} \mapsto \begin{pmatrix} 2r\cos\theta\\ 3r\sin\theta\\ r^2\end{pmatrix}$$

parametrizes the locus given by the equation $z = \frac{x^2}{4} + \frac{y^2}{9}$ (which is an elliptic paraboloid)[6]; it maps the part of the (r,θ)-plane where $0 \le \theta \le 2\pi$, $r \le a$ to the region of the paraboloid where $z \le a^2$. Thus the surface area is given by the integral

$$\int_0^{2\pi}\int_0^a \sqrt{\det\left(\left[\mathbf{D}\gamma\begin{pmatrix} r\\ \theta\end{pmatrix}\right]^\top \left[\mathbf{D}\gamma\begin{pmatrix} r\\ \theta\end{pmatrix}\right]\right)}\,dr\,d\theta = \int_0^{2\pi}\int_0^a \sqrt{\det\begin{bmatrix} 2\cos\theta & 3\sin\theta & 2r\\ -2r\sin\theta & 3r\cos\theta & 0\end{bmatrix}\begin{bmatrix} 2\cos\theta & -2r\sin\theta\\ 3\sin\theta & 3r\cos\theta\\ 2r & 0\end{bmatrix}}\,dr\,d\theta$$

$$= \int_0^{2\pi}\int_0^a 2r\sqrt{9 + r^2(4\sin^2\theta + 9\cos^2\theta)}\;dr\,d\theta.$$

This integral cannot be evaluated in terms of elementary functions, but it can be evaluated in terms of elliptic functions, using MAPLE (at least, version 7 or higher):

$$\int_0^{2\pi}\int_0^a 2r\sqrt{9 + r^2(4\sin^2\theta + 9\cos^2\theta)}\;dr\,d\theta = \frac{2}{3\sqrt{1+a^2}}\left(36\sqrt{\frac{1+a^2}{9+4a^2}}a^2\,\mathrm{ellipticK}\sqrt{-5\frac{a^2}{9+4a^2}}\right.$$

$$+81\sqrt{\frac{1+a^2}{9+4a^2}}\mathrm{elliptic}\pi\left(\frac{-5}{4}, \sqrt{-5\frac{a^2}{9+4a^2}}\right) + 16\sqrt{\frac{1+a^2}{9+4a^2}}a^4\mathrm{ellipticE}\sqrt{-5\frac{a^2}{9+4a^2}}$$

$$\left.+36\sqrt{\frac{1+a^2}{9+4a^2}}a^2\mathrm{ellipticE}\sqrt{-5\frac{a^2}{9+4a^2}} - 9\pi\sqrt{1+a^2}\right).$$

The elliptic functions "ellipticK," "ellipticE," and "ellipticπ" are tabulated functions; tables with these functions can be found, for example, in *Handbook of Mathematical Functions*, edited by Milton Abramowitz and Irene Stegun (Dover Publications, Inc.).

The graph of the area as a function of a is shown in the margin.

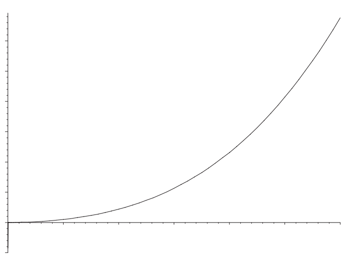


FIGURE FOR SOLUTION 5.3.5

The area of the surface as a function of a.

b. It is easier to begin by finding the volume of the region outside the elliptic paraboloid, and then subtract it from the volume of the elliptic cylinder. Call U_a the ellipse

$$\frac{x^2}{4} + \frac{y^2}{9} \le a^2.$$

Then the volume of the region above this ellipse in the plane $z = 0$ and beneath the paraboloid is

$$\int_{U_a}\left(\frac{x^2}{4} + \frac{y^2}{9}\right)|dx\,dy| = \int_0^{2\pi}\int_0^a \det\begin{bmatrix} 2\cos\theta & -2r\sin\theta\\ 3\sin\theta & 3r\cos\theta\end{bmatrix} r^2\,dr\,d\theta = 3a^4\pi.$$

The elliptic cylinder has total volume $\pi(6a^2)a^2 = 6\pi a^4$, and the difference of the two is $3\pi a^4$.

[6]How did we find this parametrization? The surface cut horizontally at height z is an ellipse, so we were guided by the parametrization $t \mapsto \begin{pmatrix} a\cos t\\ b\sin t\end{pmatrix}$ for the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$; see Example 4.10.19. We got the r^2 by noting that then $t \mapsto \begin{pmatrix} ra\cos t\\ rb\sin t\end{pmatrix}$ parametrizes the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = r^2$.

Here is another solution to part b, provided by Nathaniel Schenker and using Definition 5.3.2. In it the elliptic paraboloid is sliced horizontally:

For z fixed, the area of the ellipse $\frac{x^2}{4}+\frac{y^2}{9}\le z$ is $f\begin{pmatrix}x\\y\\z\end{pmatrix}=6\pi z$. Let M denote the z-axis between $z=0$ and $z=a^2$, parametrized as : $t\mapsto\begin{pmatrix}0\\0\\t\end{pmatrix}$, $0\le t\le a^2$. Then the volume of the region in question is

$$\begin{aligned}\int_M f\begin{pmatrix}x\\y\\z\end{pmatrix}\left|d^1\begin{pmatrix}x\\y\\z\end{pmatrix}\right| &= \int_{[0,a^2]} f(\gamma(t))\sqrt{\det([\mathbf{D}\gamma(t)]^\top[\mathbf{D}\gamma(t)]}\,|dt|\\ &= \int_0^{a^2} 6\pi t\sqrt{1}\,dt = 3\pi a^4\end{aligned}$$

Solution 5.3.7: This parametrization is inspired by the spherical change of coordinates

$$\begin{pmatrix}\theta\\\varphi\end{pmatrix}\mapsto\begin{pmatrix}\cos\theta\cos\varphi\\\sin\theta\cos\varphi\\\sin\varphi\end{pmatrix}.$$

for the unit sphere. Scaling a sphere in three directions with three scaling factors a,b,c gives an ellipsoid.

5.3.7 a. The obvious way to parametrize the surface of the ellipsoid is

$$\gamma\begin{pmatrix}\varphi\\\theta\end{pmatrix}=\begin{pmatrix}a\cos\varphi\cos\theta\\b\cos\varphi\sin\theta\\c\sin\varphi\end{pmatrix},\qquad -\frac{\pi}{2}<\varphi<\frac{\pi}{2},\ 0\le\theta<2\pi.$$

b. Using equation 5.3.33, we find that the area is

$$\begin{aligned}&\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\left|\left(\overrightarrow{D_1}\gamma\begin{pmatrix}\varphi\\\theta\end{pmatrix}\right)\times\left(\overrightarrow{D_2}\gamma\begin{pmatrix}\varphi\\\theta\end{pmatrix}\right)\right|d\varphi\,d\theta\\ &=\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\left|\begin{bmatrix}-a\sin\varphi\cos\theta\\-b\sin\varphi\sin\theta\\c\cos\varphi\end{bmatrix}\times\begin{bmatrix}-a\cos\varphi\sin\theta\\b\cos\varphi\cos\theta\\0\end{bmatrix}\right|d\varphi\,d\theta\\ &=\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\sqrt{b^2c^2\cos^4\varphi\cos^2\theta+a^2c^2\cos^4\varphi\sin^2\theta+a^2b^2\cos^2\varphi\sin^2\varphi}\ d\varphi\,d\theta.\end{aligned}$$

5.3.9 a. We parametrize S by $\gamma\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x\\y\\x^3+y^4\end{pmatrix}$. We could use our standard approach and set up the integral as

$$\begin{aligned}&\int_S(x+y+\overbrace{x^3+y^4}^{z})\sqrt{\det[\mathbf{D}\gamma(\mathbf{x})]^\top[\mathbf{D}\gamma(\mathbf{x})]}|dx\,dy|\\ &=\int_{-1}^1\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}(x+y+x^3+y^4)\sqrt{\det\left(\begin{bmatrix}1&0&3x^2\\0&1&4y^3\end{bmatrix}\begin{bmatrix}1&0\\0&1\\3x^2&4y^3\end{bmatrix}\right)}\,dx\,dy,\end{aligned}$$

but since we are in $\mathbb{R}^3$ it is probably easier computationally to use the length of the cross product to give the area of the parallelogram spanned by the partial derivatives $\overrightarrow{D_1}\gamma$ and $\overrightarrow{D_2}\gamma$ (see Proposition 1.4.20 or equation

5.3.33). Thus we have

$$\int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}(x+y+\overbrace{x^3+y^4}^{z})\left|\left[\begin{matrix}1\\0\\3x^2\end{matrix}\right]\times\left[\begin{matrix}0\\1\\4y^3\end{matrix}\right]\right|dx\,dy$$
$$=\int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}(x+y+x^3+y^4)\sqrt{(9x^4+16y^6+1}\,dx\,dy.$$

Solution 5.3.9: In part b, we use polar coordinates, but the result will not be computable in elementary terms in any case.

b. It is easier to pass to polar coordinates, so that the integral will be over a rectangle. We find

$$\int_0^{2\pi}\int_0^1(r\cos\theta+r\sin\theta+r^3\cos^3\theta+r^4\sin^4\theta)\sqrt{9r^4\cos^4\theta+16r^6\sin^6\theta+1}\;r\,dr\,d\theta.$$

We have no idea how to compute this integral symbolically. To evaluate it numerically, we need to choose a numerical method. The following sequence of MATLAB instructions performs a midpoint Riemann sum, breaking the domain into 100 intervals on each side, and evaluating the function at the center of each rectangle.

First, define the function:

```
function prob545=integrand (x,y)
prob545 = x^2*(cos(y)+sin(y)+x^2* (cos(y))^3 +x^3* (sin(y))^4)* sqrt(9*(x*cos(y))^4
    +16*(x*sin(y))^6+1);
```

Then create a matrix A whose values are the values of this function at the points $\begin{pmatrix}(2i-1)/200\\2\pi(2j-1)/200\end{pmatrix}$, for $i,j=1,\dots,200$, using the following commands:

MATLAB has a command "dblquad" that should compute double integrals, but we could not get it to work, even on MATLAB's own example.

```
N = 100
for I = 1 : N,
        for J = 1 : N,
            A(I, J) = integrand((2*I-1)/(2*N),2*pi*(2*J-1)/(2*N));
        end,
end;
```

Now sum the elements of the matrix, and multiply by $2\pi/(100^2)$, the area of the small rectangles:

```
sum(sum(A(1:N,1:N)))*2*pi/(N^2)
```

The answer the computer comes up with is 0.96117719450855.

If you want a more precise answer, you might try the 2-dimensional Simpson's method. The following MATLAB program will do this. It yields the estimate 0.96143451150993. If you repeat the procedure with $N = 201$, you will find 0.96143438060954. It seems likely that the digits 0.9614343 are accurate; this is in agreement with the estimate (from Theorem 4.6.2) that says the error should be a constant times $1/N^4$, where the constant is computed from the partials of f up to order 4.

```
N=101
S(1)=1;
S(N)=1;
for i=1:(N-1)/2, S(2*i)=4; end
for i=1:(N-3)/2, S(2*i+1)=2; end
for i=1:(N-3)/2, S(2*i+1)=2; end
for i=1:N,
   for j=1:N,
            SS(i,j) = S(i)*S(j);
   end
end

for I = 1:N,
      for J = 1:N,
         B(I,J) = integrand((I-1)/(N-1),2*pi*(J-1)/(N-1));
      end,
end;
(sum(sum(SS(1:N,1:N).*B(1:N,1:N))))*2*pi/(9*(N-1)^2)
```

5.3.11 a. The Gaussian curvature K of the sphere S of radius R is $1/R^2$. This computation is very similar to that in Solution 3.9.3. Any point can be rotated to the point $\mathbf{a} = \begin{pmatrix} 0 \\ 0 \\ R \end{pmatrix}$, so it is enough to compute the curvatures there. Near $\mathbf{a}$ the "best coordinates" are $X = x$, $Y = y$, $Z = z - R$, so that

Solution 5.3.11, part a: More generally, the same sort of computation shows that when a surface is scaled by a, its Gaussian curvature is scaled by $1/a^2$. The mean curvature is scaled by $1/a$.

$$Z = \overbrace{\sqrt{R^2 - X^2 - Y^2}}^{z} - R = R\left(\sqrt{1 - \frac{X^2}{R^2} - \frac{Y^2}{R^2}} - 1\right).$$

As in Solution 3.9.3, we use equation 3.4.9 to compute

$$Z \approx R\left(1 + \frac{1}{2}\left(-\frac{X^2}{R^2} - \frac{Y^2}{R^2}\right) + \cdots - 1\right).$$

The quadratic terms of the Taylor polynomial are $-\frac{1}{2}\left(\frac{X^2}{R} + \frac{Y^2}{R}\right)$, so

$$A_{1,1} = 0, \quad A_{2,0} = A_{0,2} = 1/R, \quad \text{giving} \quad K = 1/R^2.$$

By Definition 3.9.8, the Gaussian curvature K of a surface at a point $\mathbf{a}$ is

$$K(\mathbf{a}) = A_{2,0}A_{0,2} - A_{1,1}^2.$$

The area of the sphere of radius R is $4\pi R^2$, so its total curvature $\mathbf{K}(S)$ is 4π, independent of R:

$$\mathbf{K}(S) = \int_S |K(\mathbf{x})|\,|d^2\mathbf{x}| = \frac{1}{R^2}4\pi R^2 = 4\pi.$$

b. The hyperboloid of equation $z = x^2 - y^2$ is parametrized by x and y. We computed the Gaussian curvature of this surface in Example 3.9.12, and found it to be

The parametrization in Solution 5.3.11, part b is an example of parametrizing as a graph:

$$\gamma\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ x^2 - y^2 \end{pmatrix}.$$

$$K\begin{pmatrix} x \\ y \\ x^2 - y^2 \end{pmatrix} = \frac{-4}{(1+4x^2+4y^2)^2}.$$

Thus our total curvature is

$$\int_{\mathbb{R}^2} \frac{|-4|}{(1+4x^2+4y^2)^2} \left| \begin{bmatrix} 1 \\ 0 \\ 2x \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ -2y \end{bmatrix} \right| |dx\,dy| = 4\int_{\mathbb{R}^2} \frac{|dx\,dy|}{(1+4x^2+4y^2)^{3/2}}.$$

This integral exists (as a Lebesgue integral), and can be computed in polar coordinates (this uses Theorems 4.11.21 and 4.11.20), to yield

$$4\int_0^{2\pi}\left(\int_0^\infty \frac{r}{(1+4r^2)^{3/2}}\,dr\right)d\theta = \pi\int_0^\infty \frac{8r}{(1+4r^2)^{3/2}}\,dr = 2\pi.$$

Solution 5.3.13, part a: This result was first due to Archimedes, who was so proud of it that he arranged for a sphere inscribed in a cylinder to adorn his tomb, in Sicily. One hundred thirty-seven years later, in 75 BC, Cicero (then the Roman questor of Sicily) found the tomb, "completely surrounded and hidden by bushes of brambles and thorns. I remembered having heard of some simple lines of verse which had been inscribed on his tomb, referring to a sphere and cylinder . . . . And so I took a good look round . . . . Finally I noted a little column just visible above the scrub: it was surmounted by a sphere and a cylinder."

5.3.13 a. Using cylindrical coordinates, the mapping from the cylinder to the sphere can be written

$$f\begin{pmatrix} \theta \\ z \end{pmatrix} = \begin{pmatrix} \sqrt{1-z^2}\cos\theta \\ \sqrt{1-z^2}\sin\theta \\ z \end{pmatrix};$$

the cross product of the two partials is the vector

$$\begin{bmatrix} -\sqrt{1-z^2}\sin\theta \\ \sqrt{1-z^2}\cos\theta \\ 0 \end{bmatrix} \times \begin{bmatrix} -z\cos\theta/\sqrt{1-z^2} \\ -z\sin\theta/\sqrt{1-z^2} \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{1-z^2}\cos\theta \\ \sqrt{1-z^2}\sin\theta \\ z \end{bmatrix}.$$

This vector has length 1, so by equation 5.3.33 and Definition 5.3.1, for any subset $C \subset S_1$ with volume,

$$\begin{aligned} \mathrm{vol}_2 f(C) &= \int_C \sqrt{\det[\mathbf{D}\gamma(\mathbf{u})]^\top[\mathbf{D}\gamma(\mathbf{u})]}\,d^2\mathbf{u} \\ &= \int_C 1\,d^2\mathbf{u} = \int_{\mathbb{R}^2} \mathbf{1}_C\,d^2\mathbf{u} = \mathrm{vol}_2\, C. \end{aligned}$$

b. The area of the polar caps and of the tropics is exactly the same as the area of the corresponding parts of the cylinder:

$$\text{Area of polar caps} = 2(40000)^2(1-\cos 23°27')/(2\pi) \approx 42.06\ \times 10^6\ \text{km}^2$$

$$\text{Area of tropics} = 2(40000)^2(\sin 23°27')/(2\pi) \approx 203\ \times 10^6\ \text{km}^2.$$

Thus the tropics have roughly five times the area of the polar caps.

c. As explained by the figure in the margin, the area in question is

$$\begin{aligned} A(r) &= 2\pi R^2\left(1-\cos\frac{r}{R}\right) = 2\pi R^2\left(\frac{r^2}{2R^2} - \frac{r^4}{24R^4} + \cdots\right) \\ &= \pi r^2 - \frac{\pi}{12R^2}r^4 + \cdots. \end{aligned}$$

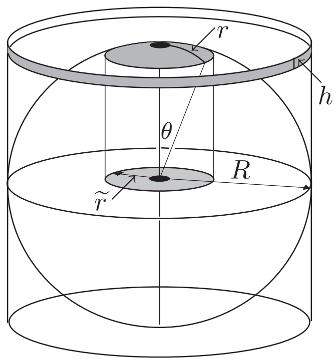


Figure for Solution 5.3.13.

Part c: A sphere of radius R inscribed in a cylinder. The angle θ at the center of the sphere corresponding to an arc of great circle of length r on a sphere of radius R is r/R. Thus the part of the cylinder that projects to it has length $2\pi R$ and height

$$h = R\left(1 - \cos\frac{r}{R}\right),$$

hence area

$$2\pi R\left(R\left(1-\cos\frac{r}{R}\right)\right).$$

By part a, this equals the area of the polar cap of radius r.

5.3.15 a. For the point $\begin{pmatrix} a\\ b\\ c\\ d\end{pmatrix} \in \mathbb{R}^4$ to be on the unit sphere, we must have $a^2+b^2+c^2+d^2=1$, so a first step in showing that γ parametrizes S^3 is to show that the sum of the squares of the coordinates is 1:

$$\begin{aligned}&\cos^2\psi\cos^2\varphi\cos^2\theta + \cos^2\psi\cos^2\varphi\sin^2\theta + \cos^2\psi\sin^2\varphi + \sin^2\psi\\ &\quad = \cos^2\psi\cos^2\varphi + \cos^2\psi\sin^2\varphi + \sin^2\psi = \cos^2\psi + \sin^2\psi = 1.\end{aligned}$$

Next we must check the domains. Except possibly on sets of 3-dimensional volume 0 (individual points, for example), is γ one to one and onto? Think of the fourth coordinate, $\sin\psi$, as being the "height" of S^3, analogous to the z-coordinate for the unit sphere in $\mathbb{R}^3$. Clearly $\sin\psi$ must run through $[-1,1]$, and does run through these values exactly once as ψ varies from $-\pi/2$ to $\pi/2$.

For any fixed value of $\psi \in (-\pi/2, \pi/2)$, the other coordinates satisfy

$$a^2+b^2+c^2+(\sin\psi)^2 = 1, \quad \text{i.e.,} \quad a^2+b^2+c^2 = (\cos\psi)^2, \qquad (1)$$

so $\begin{pmatrix} a\\ b\\ c\end{pmatrix}$ runs through the 2-sphere in $\mathbb{R}^3$ of radius $\cos\psi$. For fixed $\sin\psi$, the first three coordinates are:

$$\begin{pmatrix} \cos\psi\cos\theta\cos\varphi\\ \cos\psi\sin\theta\cos\varphi\\ \cos\psi\sin\varphi\end{pmatrix},$$

which is the parametrization of the sphere in $\mathbb{R}^3$ with radius $\cos\psi$ by spherical coordinates (Definition 4.10.6). Thus γ is a parametrization of the 3-sphere in $\mathbb{R}^4$.

Equations 1: $\sin^2\varphi$ and $(\sin\varphi)^2$ mean the same thing, as do $\cos^2\varphi$ and $(\cos\varphi)^2$; the second is better notation but the first is more traditional.

b. To compute $\mathrm{vol}_3(S^3)$ we use equation 5.3.3. First we compute

$$\left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right] \quad\text{and}\quad \left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right]^\top \left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right]:$$

$$\left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right] = \begin{bmatrix} -\cos\psi\cos\varphi\sin\theta & -\cos\psi\sin\varphi\cos\theta & -\sin\psi\cos\varphi\cos\theta\\ \cos\psi\cos\varphi\cos\theta & -\cos\psi\sin\varphi\sin\theta & -\sin\psi\cos\varphi\sin\theta\\ 0 & \cos\psi\cos\varphi & -\sin\psi\sin\varphi\\ 0 & 0 & \cos\psi\end{bmatrix}.$$

$$\left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right]^\top = \begin{bmatrix} -\cos\psi\cos\varphi\sin\theta & \cos\psi\cos\varphi\cos\theta & 0 & 0\\ -\cos\psi\sin\varphi\cos\theta & -\cos\psi\sin\varphi\sin\theta & \cos\psi\cos\varphi & 0\\ -\sin\psi\cos\varphi\cos\theta & -\sin\psi\cos\varphi\sin\theta & -\sin\psi\sin\varphi & \cos\psi\end{bmatrix},$$

giving

Computing this product is a little easier if you remember that $A^\top A$ is always symmetric (see Exercise 1.2.16 and equation 5.1.9).

$$\left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right]^\top \left[\mathbf{D}\gamma\begin{pmatrix}\theta\\ \varphi\\ \psi\end{pmatrix}\right] = \begin{bmatrix} \cos^2\psi\cos^2\varphi & 0 & 0\\ 0 & \cos^2\psi & 0\\ 0 & 0 & 1\end{bmatrix}.$$

The determinant of this matrix is $\cos^4\psi\cos^2\varphi$, so applying equation 5.3.3 to our problem gives

$$\begin{aligned}\operatorname{vol}_3 S^3 &= \int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2}\int_0^{2\pi} \cos^2\psi\cos\varphi\, d\theta\, d\varphi\, d\psi\\ &= 2\pi\int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2}\cos^2\psi\cos\varphi\, d\varphi\, d\psi\\ &= 2\pi\int_{-\pi/2}^{\pi/2}[\sin\varphi]_{-\pi/2}^{\pi/2}\cos^2\psi\, d\psi = 4\pi\int_{-\pi/2}^{\pi/2}\cos^2\psi\, d\psi = 4\pi\frac{\pi}{2}\\ &= 2\pi^2.\end{aligned}$$

To integrate $\cos^2\psi$ we used the fact that $\cos^2\psi$ has average $1/2$ over one full period, which for $\cos^2$ is π.

You may notice that this result is the same as the value for $\operatorname{vol}_3 S^3$ given in Table 5.3.3.

5.3.17 It's pretty easy to write down the appropriate integrals, but one of them is easy to evaluate in elementary terms, one is hard, and one is very hard.

Denote by $\overline{x}$ and $\overline{y}$ the x and y components of the center of gravity. Then (by Definition 4.2.1)

$$\overline{x} = \frac{\int_C x\,|d\mathbf{x}|}{\int_C |d\mathbf{x}|} \quad\text{and}\quad \overline{y} = \frac{\int_C y\,|d\mathbf{x}|}{\int_C |d\mathbf{x}|}.$$

The curve is parametrized by $x \mapsto \begin{pmatrix} x\\ x^2\end{pmatrix}$, so the integrals become

$$\begin{aligned}\int_0^a x|d^1\mathbf{x}|\begin{bmatrix}1\\2x\end{bmatrix} dx &= \int_0^a x\sqrt{1+4x^2}\,dx\\ \int_0^a y|d^1\mathbf{x}|\begin{bmatrix}1\\2x\end{bmatrix} dx &= \int_0^a x^2\sqrt{1+4x^2}\,dx \qquad (1)\\ \int_0^a |d^1\mathbf{x}|\begin{bmatrix}1\\2x\end{bmatrix} dx &= \int_0^a \sqrt{1+4x^2}\,dx.\end{aligned}$$

In equation 1, the element of length $|d^1\mathbf{x}|$ is being evaluated on the vector $\begin{bmatrix}1\\2x\end{bmatrix}$. We could also write this as $\left|\begin{bmatrix}1\\2x\end{bmatrix}\right|$.

The first of these is easy, setting $4x^2 = u$, so $8x\,dx = du$. This leads to

$$\int_0^a x\sqrt{1+4x^2}\,dx = \frac{1}{12}\left((1+4a^2)^{3/2} - 1\right).$$

The third is a good bit harder. Substitute $2x = \tan\theta$, so that $2\,dx = d\theta/\cos^2\theta$. Our integral becomes

$$\int\sqrt{1+4x^2}\,dx = \frac{1}{2}\int\frac{1}{\cos^3\theta}\,d\theta = \frac{1}{2}\int\frac{\cos\theta}{(1-\sin^2\theta)^2}\,d\theta.$$

Now substitute $\sin\theta = t$, so that $\cos\theta\,d\theta = dt$, and the integral becomes

$$\frac{1}{2}\int\frac{dt}{(1-t^2)^2} = \frac{1}{8}\left(\int\frac{dt}{1-t} + \int\frac{dt}{1+t} + \int\frac{dt}{(1-t)^2} + \int\frac{dt}{(1+t)^2}\right),$$

using partial fractions. This is easy to integrate, giving

$$\frac{1}{4}\frac{t}{1-t^2} + \frac{1}{8}\ln\left|\frac{1+t}{1-t}\right|.$$

Substituting back leads to

$$\int_0^a \sqrt{1+4x^2}\,dx = \frac{1}{2}\left(a\sqrt{1+4a^2} + \frac{1}{2}\ln\bigl|2a+\sqrt{1+4a^2}\bigr|\right). \tag{2}$$

The second integral is trickier yet. Begin by integrating by parts:

$$\underbrace{\int x^2\sqrt{1+4x^2}\,dx}_{A} = \int \frac{x}{8}\cdot 8x\sqrt{1+4x^2}\,dx = \frac{x}{12}(1+4x^2)^{3/2} - \frac{1}{12}\int(1+4x^2)\sqrt{1+4x^2}\,dx$$

$$= \frac{x}{12}(1+4x^2)^{3/2} - \frac{1}{12}\underbrace{\int\sqrt{1+4x^2}\,dx}_{\text{see equation 2}} - \frac{4}{12}\underbrace{\int x^2\sqrt{1+4x^2}\,dx}_{A}.$$

Now insert the value computed in equation 2, and add the term on the far right to both sides of the equation, to get

$$\frac{4}{3}\int x^2\sqrt{1+4x^2}\,dx = \frac{x}{12}(1+4x^2)^{3/2} - \frac{1}{24}\left(x\sqrt{1+4x^2} + \frac{1}{2}\ln\bigl|2x+\sqrt{1+4x^2}\bigr|\right).$$

Finally, we find

$$\int_0^a x^2\sqrt{1+4x^2}\,dx = \frac{a}{16}(1+4a^2)^{3/2} - \frac{1}{32}\left(a\sqrt{1+4a^2} + \frac{1}{2}\ln\bigl|2a+\sqrt{1+4a^2}\bigr|\right).$$

This gives

We do not consider an ability to compute such complicated integrals an important mathematical skill. In general, it is far better to use MAPLE or similar software to carry out such computations.

$$\overline{x} = \frac{1}{6}\frac{\bigl((1+4a^2)^{3/2}-1\bigr)}{a\sqrt{1+4a^2}+\frac{1}{2}\ln|2a+\sqrt{1+4a^2}|}$$

$$\overline{y} = \frac{\frac{a}{8}(1+4a^2)^{3/2} - \frac{1}{16}\Bigl(a\sqrt{1+4a^2}+\frac{1}{2}\ln|2a+\sqrt{1+4a^2}|\Bigr)}{a\sqrt{1+4a^2}+\frac{1}{2}\ln|2a+\sqrt{1+4a^2}|}.$$

5.3.19 To go from the third to the last line of equation 5.3.49, first we do an integration by parts, with

$$du = \sqrt{1+r^2}\,r\,dr \text{ and } v = r, \quad \text{so that} \quad u = \frac{(1+r^2)^{3/2}}{3} \text{ and } dv = dr.$$

This gives

$$4\pi\int_0^R \sqrt{1+r^2}\,r^2\,dr = \frac{4\pi}{3}\Bigl[(1+r^2)^{3/2}r\Bigr]_0^R - \frac{4\pi}{3}\int_0^R (1+r^2)^{3/2}\,dr$$

$$= \frac{4\pi}{3}R(1+R^2)^{3/2} - \frac{4\pi}{3}\left(\int_0^R \sqrt{1+r^2} + \sqrt{1+r^2}\,r^2\,dr\right).$$

Multiplying each side by $3/(4\pi)$ gives

$$3\int_0^R \sqrt{1+r^2}\,r^2\,dr \;=\; R(1+R^2)^{3/2} - \int_0^R \sqrt{1+r^2}\,dr - \int_0^R \sqrt{1+r^2}\,r^2\,dr,$$

and adding the last term on the right to both sides gives

$$4\int_0^R \sqrt{1+r^2}\,r^2\,dr \;=\; R(1+R^2)^{3/2} - \int_0^R \sqrt{1+r^2}\,dr. \tag{1}$$

Now we fiddle with the last term on the right side. To go from the first to the second line below, we set $\tan\theta = r$, so that $dr = \frac{d\theta}{\cos\theta}$, and $\sin\theta = \tan\theta\sqrt{\frac{1}{1+\tan^2\theta}} = r\sqrt{\frac{1}{1+r^2}}$:

$$\begin{aligned}
-\int_0^R \sqrt{1+r^2}\,dr &= -\int_0^R \frac{1+r^2}{\sqrt{1+r^2}}\,dr = \overbrace{-\int_0^R \frac{1}{\sqrt{1+r^2}}\,dr}^{\text{1st}} \overbrace{-\int_0^R \frac{r^2}{\sqrt{1+r^2}}\,dr}^{\text{2nd}} \\
&= \underbrace{-\int_0^{\arctan R} \frac{\cos\theta}{\cos^2\theta}\,d\theta}_{\substack{\text{from 1st term in braces, using}\\ \text{trig. as described above}}} \underbrace{-\left[r\sqrt{1+r^2}\right]_0^R + \int_0^R \sqrt{1+r^2}\,dr}_{\substack{\text{from 2nd term in braces,}\\ \text{by integration by parts}}}
\end{aligned}$$

So, subtracting from both sides the last term on the second line, we get

$$\begin{aligned}
-2\int_0^R \sqrt{1+r^2}\,dr &= -\int_0^{\arctan R} \frac{\cos\theta}{\cos^2\theta}\,d\theta - R\sqrt{1+R^2} \\
&= -\frac{1}{2}\int_0^{\arctan R} \frac{2\cos\theta}{1-\sin^2\theta}\,d\theta - R\sqrt{1+R^2} \\
&= -\frac{1}{2}\int_0^{\arctan R} \left(\frac{\cos\theta}{1+\sin\theta} + \frac{\cos\theta}{1-\sin\theta}\right)d\theta - R\sqrt{1+R^2} \\
&= -\frac{1}{2}\left[\ln\left(\frac{1+\sin\theta}{1-\sin\theta}\right)\right]_0^{\arctan R} - R\sqrt{1+R^2} \\
&= -\frac{1}{2}\left[\ln \frac{\frac{r+\sqrt{1+r^2}}{\sqrt{1+r^2}}}{\frac{\sqrt{1+r^2}-r}{\sqrt{1+r^2}}}\right]_0^R - R\sqrt{1+R^2} \\
&= -\frac{1}{2}\ln\frac{(R+\sqrt{1+R^2})^2}{1+R^2-R^2} - R\sqrt{1+R^2} \\
&= -\ln(R+\sqrt{1+R^2}) - R\sqrt{1+R^2}.
\end{aligned}$$

Thus

$$-\int_0^R \sqrt{1+r^2}\,dr = -\frac{1}{2}\ln\Big(R+\sqrt{1+R^2}\Big) - \frac{R\sqrt{1+R^2}}{2}. \qquad (2)$$

Now we multiply equation 1 by π, replacing the second term on the right side by its equivalent value in equation 2, to get

$$4\pi\int_0^R \sqrt{1+r^2}\,r^2\,dr = \pi\left(R(1+R^2)^{3/2} - \frac{1}{2}\ln\Big(R+\sqrt{1+R^2}\Big) - \frac{R\sqrt{1+R^2}}{2}\right).$$

5.3.21 We will use equation 5.3.3 for computing the k-dimensional volume of a k-dimensional manifold in $\mathbb{R}^n$, in this case a 2-dimensional manifold in $\mathbb{C}^3$, which we can think of as being $\mathbb{R}^6$.

Let us write $z = s(\cos\theta + i\sin\theta)$, so that $s = |z|$. (This uses equations 0.7.9 and 0.7.10.) Then using de Moivre's formula

$$\big(s(\cos\theta + i\sin\theta)\big)^p = s^p(\cos p\theta + i\sin p\theta),$$

we have

$$\gamma\begin{pmatrix} s\\ \theta\end{pmatrix} = \begin{pmatrix} s^p\cos p\theta\\ s^p\sin p\theta\\ s^q\cos q\theta\\ s^q\sin q\theta\\ s^r\cos r\theta\\ s^r\sin r\theta\end{pmatrix} \quad\text{and}\quad \left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right] = \begin{bmatrix} ps^{p-1}\cos p\theta & -ps^p\sin p\theta\\ ps^{p-1}\sin p\theta & ps^p\cos p\theta\\ qs^{q-1}\cos q\theta & -qs^q\sin q\theta\\ qs^{q-1}\sin q\theta & qs^q\cos q\theta\\ rs^{r-1}\cos r\theta & -rs^r\sin r\theta\\ rs^{r-1}\sin r\theta & rs^r\cos r\theta\end{bmatrix}.$$

Computing $\left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right]^\top\left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right]$ and taking the determinant gives

$$\begin{aligned}\det\overbrace{\begin{bmatrix} p^2s^{2p-2}+q^2s^{2q-2}+r^2s^{2r-2} & 0\\ 0 & p^2s^{2p}+q^2s^{2q}+r^2s^{2r}\end{bmatrix}}^{\left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right]^\top\left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right]}\\ = \left(p^2s^{2p-1}+q^2s^{2q-1}+r^2s^{2r-1}\right)^2,\end{aligned}$$

so taking the square root of the determinant is easy. (This is the kind of thing that happens with complex numbers.) Thus if we call our surface S, equation 5.3.3 gives

$$\begin{aligned}\operatorname{vol}_k S &= \int_{|z|\le 1}\sqrt{\det\left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right]^\top\left[\mathbf{D}\gamma\begin{pmatrix} s\\ \theta\end{pmatrix}\right]}\,ds\,d\theta\\ &= \int_0^{2\pi}\int_0^1\left(p^2s^{2p-1}+q^2s^{2q-1}+r^2s^{2r-1}\right)\,ds\,d\theta\\ &= 2\pi\left(\left[\frac{p^2s^{2p}}{2p}\right]_0^1+\left[\frac{q^2s^{2q}}{2q}\right]_0^1+\left[\frac{r^2s^{2r}}{2r}\right]_0^1\right)\\ &= 2\pi\left(\frac{p}{2}+\frac{q}{2}+\frac{r}{2}\right) = \pi(p+q+r).\end{aligned}$$

5.4.1 a. The image of the hyperboloid under the Gauss map is the part of the sphere where $|\cos\varphi| \ge \dfrac{1}{\sqrt{1+a^2}}$. Set $\phi_a = \arccos 1/\sqrt{1+a^2}$. Using spherical coordinates, we find the area of this region to be

$$\int_0^{2\pi}\int_{-\phi_a}^{\phi_a}\cos\varphi\,d\varphi\,d\theta = 2\pi(2\sin\phi_a) = 4\pi\frac{a}{\sqrt{1+a^2}}.$$

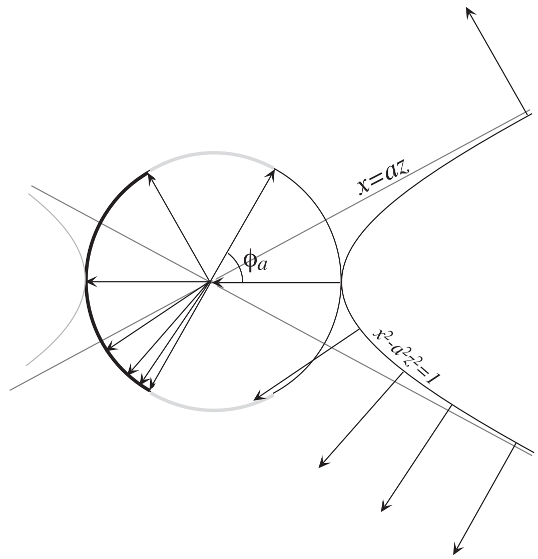


FIGURE FOR SOLUTION 5.4.1, part a: At right, we see the branch of the hyperbola of equation

$$x^2 + a^2z^2 = 1$$

where $x > 0$, together with representative normals, and the same normals anchored at the origin. These normals fill the arc

$$\pi - \frac{a}{\sqrt{a^2+1}} < \varphi < \pi + \frac{a}{\sqrt{a^2+1}}.$$

The image of the hyperboloid under the Gauss map is obtained by rotating this arc around the z-axis.

b. In this case, we are rotating the branch of the hyperbola of equation $a^2z^2 = x^2 + 1$ where $z > 0$ around the z-axis. The image of the Gauss map is also obtained by revolving the set of normals to the hyperbola around the z-axis. The normals (choosing the upward-pointing normal) satisfy $\arcsin \frac{a}{\sqrt{1+a^2}} \le \varphi \le \pi/2$. Thus the image of the Gauss map has area

$$\int_0^{2\pi} \int_{\arcsin \frac{a}{\sqrt{1+a^2}}}^{\frac{\pi}{2}} \cos\varphi \, d\varphi \, d\theta = 2\pi \left[\sin\varphi\right]_{\arcsin \frac{a}{\sqrt{1+a^2}}}^{\frac{\pi}{2}} = 2\pi \left(1 - \frac{a}{\sqrt{1+a^2}}\right).$$

5.4.3 The map

$$\gamma : \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ \theta \end{pmatrix}, \quad 0 \le r \le R, \ 0 \le \theta < a,$$

parametrizes our surface. Indeed,

$$x\cos z + y \sin z = -r\sin\theta\cos\theta + r\sin\theta\cos\theta = 0,$$

so the image of γ is part of the surface, and moreover γ is invertible:

$$\gamma^{-1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sqrt{x^2+y^2} \\ z \end{pmatrix}.$$

We have

$$\det\left(\left[\mathbf{D}\gamma\begin{pmatrix} r \\ \theta \end{pmatrix}\right]\right)^\top \left(\left[\mathbf{D}\gamma\begin{pmatrix} r \\ \theta \end{pmatrix}\right]\right) = \det \begin{bmatrix} -\sin\theta & \cos\theta & 0 \\ -r\cos\theta & -r\sin\theta & 1 \end{bmatrix} \begin{bmatrix} -\sin\theta & -r\cos\theta \\ \cos\theta & -r\sin\theta \\ 0 & 1 \end{bmatrix} = 1 + r^2.$$

We saw in Example 3.9.13 that the curvature depends only on r, and

$$K(r) = \frac{-1}{(1+r^2)^2}.$$

Therefore, by Definition 5.3.2,

$$\int_{X_{a,R}} |K(\mathbf{x})||d^2\mathbf{x}| = \int_0^a \left(\int_0^R \frac{1}{(1+r^2)^2}\sqrt{1+r^2}dr\right) d\theta = \frac{aR}{\sqrt{1+R^2}}.$$

The inner integral is calculated by setting

$$r = \tan t,\ dr = \frac{dt}{\cos^2 t} \quad \text{and} \quad 1 + r^2 = \frac{1}{\cos^2 t},$$

leading to

$$\int_0^R \frac{dr}{(1+r^2)^{3/2}} = \int_0^{\arctan R} \frac{\cos^3 t\ dt}{\cos^2 t} = \Big[\sin t\Big]_0^{\arctan R} = \frac{R}{\sqrt{1+R^2}}.$$

5.4.5 Choose an orthonormal basis $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ of $T_{\mathbf{x}_0}M$, and define coordinate functions $X_1, \dots, X_{n+1} : \mathbb{R}^{n+1} \to \mathbb{R}$ such that for any $\mathbf{x} \in \mathbb{R}^{n+1}$ we have

$$\mathbf{x} = \mathbf{x}_0 + X_1(\mathbf{x})\vec{\mathbf{v}}_1 + \cdots + X_n(\mathbf{x})\vec{\mathbf{v}}_n + X_{n+1}(\mathbf{x})\vec{\mathbf{n}}. \qquad (1)$$

There is a formula for these functions:

$$X_i(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_0) \cdot \vec{\mathbf{v}}_i,\ i = 1, \dots, n; \quad X_{n+1}(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_0) \cdot \vec{\mathbf{n}}; \qquad (2)$$

in particular, the functions exist and are unique.

Since $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n$ is an orthonormal basis of $T_{\mathbf{x}_0}M$, the vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_n, \vec{\mathbf{n}}$ form an orthonormal basis of $\mathbb{R}^{n+1}$. So it follows from equation 1 that

$$\begin{aligned}(\mathbf{x} - \mathbf{x}_0) \cdot \vec{\mathbf{v}}_i &= \Big(X_1(\mathbf{x})\vec{\mathbf{v}}_1 + \cdots + X_n(\mathbf{x})\vec{\mathbf{v}}_n\Big) \cdot \vec{\mathbf{v}}_i \\ &\quad + X_{n+1}(\mathbf{x})\vec{\mathbf{n}} \cdot \vec{\mathbf{v}}_i \\ &= X_i(\mathbf{x}).\end{aligned}$$

and

$$\begin{aligned}(\mathbf{x} - \mathbf{x}_0) \cdot \vec{\mathbf{n}} &= \Big(X_1(\mathbf{x})\vec{\mathbf{v}}_1 + \cdots + X_n(\mathbf{x})\vec{\mathbf{v}}_n\Big) \cdot \vec{\mathbf{n}} \\ &\quad + X_{n+1}(\mathbf{x})\vec{\mathbf{n}} \cdot \vec{\mathbf{n}} \\ &= X_{n+1}(\mathbf{x}).\end{aligned}$$

The manifold M is then the graph of a function $\mathbf{f}$ expressing X_{n+1} in terms of $X_1, \dots, X_n$. This function vanishes at the origin, and the graph of its linear terms is the tangent space $T_{\mathbf{x}_0}M$, which in these coordinates has equation $X_{n+1} = 0$; i.e., the linear terms vanish, and the function $\mathbf{f}$ has a Taylor polynomial starting with quadratic terms.

Write

$$\mathbf{f}(\mathbf{X}) = \frac{1}{2}Q(\mathbf{X}) + o(|\mathbf{X}|^2),$$

so that Q is a quadratic form on $\mathbb{R}^n$. Choose an orthonormal basis $\vec{\mathbf{w}}_1, \dots, \vec{\mathbf{w}}_n$ of $\mathbb{R}^n$ that diagonalizes this quadratic form:

$$Q(c_1\vec{\mathbf{w}}_1 + \cdots + c_n\vec{\mathbf{w}}_n) = a_1c_1^2 + \cdots + a_nc_n^2,$$

where the a_i are the eigenvalues of the symmetric matrix representing Q.

We described the vectors $\vec{\mathbf{w}}_i$ as being in $\mathbb{R}^n$, i.e.,

$$\vec{\mathbf{w}}_i = p_{1,i}\vec{\mathbf{e}}_1 + \cdots + p_{n,i}\vec{\mathbf{e}}_n,$$

where the $p_{1,i}, \dots, p_{n,i}$ are coordinates of the change of basis matrix (Proposition and Definition 2.6.20). But they should really be thought of as linear combinations of the $\vec{\mathbf{v}}_j$. Set

$$\vec{\mathbf{u}}_i = p_{1,i}\vec{\mathbf{v}}_1 + \cdots + p_{n,i}\vec{\mathbf{v}}_n,$$

so the $\vec{\mathbf{u}}_i,\ i = 1, \dots, n$ form a new orthonormal basis of $T_{\mathbf{x}_0}M$. With respect to these coordinates, i.e, if we set

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n+1}\end{pmatrix} = \mathbf{x}_0 + Y_1(\mathbf{x})\vec{\mathbf{u}}_1 + \cdots + Y_n(\mathbf{x})\vec{\mathbf{u}}_n + Y_{n+1}(\mathbf{x})\vec{\mathbf{n}},$$

the equation of M near $\mathbf{x}_0$ is of the form

$$Y_{n+1} = \frac{1}{2}\left(a_1 Y_1^2 + \cdots + a_n Y_n^2\right) + o(|\mathbf{Y}|^2).$$

b. There are only two unit normals at $\mathbf{x}_0$: $\vec{\mathbf{n}}$ and $-\vec{\mathbf{n}}$. Exchanging $\vec{\mathbf{n}}$ for $-\vec{\mathbf{n}}$ changes the sign of $H(\mathbf{x}_0)$, so the vector $H(\mathbf{x}_0)\vec{\mathbf{n}}$ is not changed.

5.4.7 This is almost identical to the proof of Theorem 5.4.4. Let $U \subset \mathbb{R}^n$ be a subset such that $\gamma : U \to \mathbb{R}^{n+1}$ is a parametrization of M. Set $\vec{W}(\mathbf{u}) = \vec{\mathbf{w}}(\gamma(\mathbf{u}))$, and define

$$\gamma_t(\mathbf{u}) = \gamma(\mathbf{u}) + t\vec{\mathbf{w}}(\mathbf{u})).$$

Then the area of M_t is

$$\begin{aligned}&\int_U \sqrt{\det\Big([\mathbf{D}\gamma_t(\mathbf{u})]^\top [\mathbf{D}\gamma_t(\mathbf{u})]\Big)}|d^n\mathbf{u}|\\ =&\int_U \sqrt{\det\Big([\mathbf{D}\gamma(\mathbf{u})]^\top [\mathbf{D}\gamma(\mathbf{u})]\Big) + t\big([\mathbf{D}\gamma(\mathbf{u})]^\top[\mathbf{D}\vec{W}(\mathbf{u})] + [\mathbf{D}\vec{W}(\mathbf{u})]^\top[\mathbf{D}\gamma(\mathbf{u})]\big)\Big)}|d^n\mathbf{u}| + o(t).\end{aligned}$$

The exercise asks you to show that this integral is equal to

$$\int_U \sqrt{\det\Big([\mathbf{D}\gamma(\mathbf{u})]^\top [\mathbf{D}\gamma(\mathbf{u})]\Big)}\Big(1 - 2t\vec{H}(\gamma(\mathbf{u}))\cdot\vec{W}(\mathbf{u})\Big)|d^n\mathbf{u}| + o(t).$$

Again, we will show that these integrals are equal in the best way, by showing that the integrands are equal. This is a pointwise computation, so we may assume that we are in best coordinates, i.e., that we are at the origin and that M is the graph of a map

$$\mathbf{f}(\mathbf{x}) = \frac{1}{2}\Big(a_1x_1^2 + \cdots + a_nx_n^2\Big) + o(|\mathbf{x}|^2).$$

We use the parametrization

$$\widetilde{\mathbf{f}} : \mathbf{x} \mapsto \begin{pmatrix}\mathbf{x}\\ \mathbf{f}(\mathbf{x})\end{pmatrix}.$$

The vector $\vec{W}(\mathbf{x})$ is orthogonal to M at $\mathbf{x}$. Thus up to terms of higher order it can be written

$$\vec{W}(\mathbf{x}) = \alpha(\mathbf{x})\begin{bmatrix}-a_1x_1\\ \vdots\\ -a_nx_n\\ 1\end{bmatrix} + o(|\mathbf{x}|,$$

since

Equation 1: The vector

$$\begin{bmatrix} \vdots \\ \vec{\mathbf{e}}_i \\ \vdots \\ a_i x_i \end{bmatrix}$$

is $n+1$ high. The standard basis vector $\vec{\mathbf{e}}_i \in \mathbb{R}^n$ gives the first n entries; the last entry is $a_i x_i$.

$$\begin{bmatrix} -a_1x_1 \\ \vdots \\ -a_nx_n \\ 1 \end{bmatrix} \cdot D_i\widetilde{\mathbf{f}} = \begin{bmatrix} -a_1x_1 \\ \vdots \\ -a_nx_n \\ 1 \end{bmatrix} \cdot \left(\begin{bmatrix} \vdots \\ \vec{\mathbf{e}}_i \\ \vdots \\ a_i x_i \end{bmatrix} + o(|\mathbf{x}|) \right) = o(|\mathbf{x}|) \qquad (1)$$

so the vector $\begin{bmatrix} -a_1x_1 \\ \vdots \\ -a_nx_n \\ 1 \end{bmatrix}$ is (up to terms of high order) orthogonal to a basis of $T_{\mathbf{x}}M$. This gives

$$[\mathbf{D}\vec{W}(\mathbf{x})] = [\mathbf{D}\alpha(\mathbf{x})] \begin{bmatrix} -a_1x_1 \\ \vdots \\ -a_nx_n \\ 1 \end{bmatrix} + \alpha(\mathbf{x}) \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \\ 0 & \dots & 0 \end{bmatrix} + o(1).$$

In the last equality of equation 2 we use the fact (Theorem 4.8.15) that

$$[\mathbf{D}\det(I)]B = \operatorname{tr} B,$$

so

$$\det(I+tB) = 1 + t\operatorname{tr} B + o(t).$$

and finally

$$[\mathbf{D}\vec{W}(\mathbf{0})] = \begin{bmatrix} \alpha(\mathbf{0})a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha(\mathbf{0})a_n \\ D_1\alpha(\mathbf{0}) & \dots & D_n\alpha(\mathbf{0}) \end{bmatrix}.$$

We can now evaluate the first form of the integrand at the origin:

$$\begin{aligned} &\sqrt{\det\left([\mathbf{D}\gamma(\mathbf{0})]^\top[\mathbf{D}\gamma(\mathbf{0})] + t\left([\mathbf{D}\gamma(\mathbf{0})]^\top[\mathbf{D}\vec{W}(\mathbf{0})] + [\mathbf{D}\vec{W}(\mathbf{0})]^\top[\mathbf{D}\gamma(\mathbf{0})]\right)\right)} \\ &= \sqrt{\det\left(\begin{bmatrix} I \\ 0 \end{bmatrix}^\top \begin{bmatrix} I \\ 0 \end{bmatrix} + t\alpha(\mathbf{0})\left(\begin{bmatrix} I \\ 0 \end{bmatrix}^\top [\mathbf{D}\vec{W}(\mathbf{0})] + [\mathbf{D}\vec{W}(\mathbf{0})]^\top \begin{bmatrix} I \\ 0 \end{bmatrix}\right)\right)} + o(t) \qquad (2) \\ &= \sqrt{\det\left(I + 2t\alpha(\mathbf{0}) \begin{bmatrix} a_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_n \end{bmatrix}\right)} + o(t) \underbrace{=}_{\text{Thm. 4.8.15}} 1 + 2t\alpha(\mathbf{0})(a_1 + \dots + a_n) + o(t). \end{aligned}$$

This last expression is indeed what we want, since

$$\vec{\mathbf{H}}(\mathbf{0}) = (a_1 + \dots + a_n)\vec{\mathbf{e}}_{n+1} \quad \text{and} \quad \vec{W}(\mathbf{0}) = \alpha(\mathbf{0})\vec{\mathbf{e}}_{n+1}.$$

5.5.1 a. In base 3, the first digit after the decimal point corresponds to thirds, the second to ninths, the third to twenty-sevenths, and so on. So removing the open middle third of $[0,1]$ removes all numbers greater than $1/3$ and less than $2/3$, which are all numbers that in base 3 start with 0.1. The number $1/3$ itself remains, but it can be written $0.0222\dots$.

At the next step, removing the open middle third of $[0,1/3]$ corresponds to removing all x such that $1/9 < x < 2/9$, which correspond to numbers that in base 3 start with 0.01. Again, the number $1/9$ itself remains, but can be written $0.00222\dots$. Similarly, removing the open middle third of

$[2/3, 1]$ corresponds to removing all x such that $7/9 < x < 8/9$, which corresponds to numbers that in base 3 start with 0.21. The number 7/9 remains, but it can be written $0.20222\ldots$.

Thus going one level deeper into the procedure corresponds to looking at one more digit after the decimal, when numbers are written in base 3. Removing the second third of $[0, 1]$ corresponds to removing numbers that begin with 0.1, and removing the second and eighth ninth corresponds to removing numbers that begin with 0.01 and 0.21 respectively.

We can show that C is an uncountable set, i.e., that the elements of C cannot be put in one-to-one correspondence with the positive integers, by noting that by changing every 2 to 1, any element of C becomes a number written in base 2; conversely, any number written in base 2 can be turned into an element of C by changing 1 to 2. Since any real number can be written in base 2, and the reals are uncountable (see Section 0.6), the elements of C are uncountable.

Solution 5.5.1: Although the rational numbers 1/3, 2/3, 1/9, 2/9, 4/9, 5/9, 7/9, 8/9, $1/27, \ldots$ are elements of C, there are elements of C not in this list.

We could also use the same procedure as we used to show that the set of real numbers is uncountable (equation 0.6.3). Make any list of elements of C, for example

$$\begin{aligned}&0.0000000200\ldots\\&0.2000000200\ldots\\&0.0000000200\ldots\\&0.0000200200\ldots\\&0.0000020200\ldots\\&\ldots\end{aligned}$$

Now change every digit along the diagonal so that 2 becomes 0 and 0 becomes 2, and use those digits to form the number $222222\ldots$. This number cannot be the nth number on the list, because it doesn't have the same nth decimal.

b. For C to be pavable, its indicator function $\mathbf{1}_C$ must be integrable (Definition 4.1.18). Theorem 4.3.9 says that a bounded function with bounded support is integrable if it is continuous except on a set of volume 0. Certainly $\mathbf{1}_C$ is bounded, since $\mathbf{1}_C(x)$ can only equal 0 or 1, and it has bounded support, since its support is $[0, 1]$. Since $\mathbf{1}_C$ equals 0 everywhere except on C, and C is closed, $\mathbf{1}_C(x)$ is continuous except on C, so if we can show that C has length 0, we will have also shown that C is pavable.

By Proposition 4.1.23, we can show that C has length 0 by showing that for every $\epsilon > 0$, we can put every element of C in a 1-dimensional dyadic cube such that the total length of the cubes is $\le \epsilon$. We will actually show that we can do this using non-dyadic cubes, and will not bother to spell out how we could convert that argument to dyadic cubes.

Consider the sequence $C_1, C_2, C_3, \ldots$, where C_1 is the interval $[0, 1]$ with the open middle third removed, C_2 is the interval with the middle third, second ninth, and eighth ninth removed, and so on. Thus $\cap_{j\to\infty} C_j = C$.[7] Further, C_1 has length 2/3, C_2 has length $4/9 = (2/3)^2$, etc. These sets include every element of C, and

$$\operatorname{vol}_1 C = \lim_{j\to\infty} \operatorname{vol}_1 C_j = \lim_{j\to\infty} \left(\frac{2}{3}\right)^j = 0.$$

c. To determine for what if any k the set C has k-dimensional volume that is not 0 or infinity, we can use the procedure of Example 5.5.1. Note that C consists of two equal parts, the first third and last third of $[0, 1]$. Denote by A the first third. Then $\operatorname{vol}_k C = 2 \operatorname{vol}_k A$. But imagine that A is a rubber band; if you stretch it by a factor of 3 so that it fills the interval $[0, 1]$, it will be identical to C. Thus $\operatorname{vol}_k C = 3^k \operatorname{vol}_k A$.

Part c: $\operatorname{vol}_k C$ is $3^k \operatorname{vol}_k A$, not $3 \operatorname{vol}_k A$ because we are assuming that C is a k-dimensional object, and multiplying one dimension of a k-dimensional object while leave other dimensions unchanged will not produce a self-similar object. To enlarge a square while keeping it square, one must multiply side and length by the same factor.

[7] Rather than speak of the intersection of the C_j as $j \to \infty$ it might seem more natural to think of this as the limit of the C_j as $j \to \infty$, but we have not defined the limit of a sequence of sets.

The only value of k (other than 0 or infinity) that satisfies

$$\begin{aligned}\operatorname{vol}_k C &= 2\operatorname{vol}_k A\\ \operatorname{vol}_k C &= 3^k \operatorname{vol}_k A\end{aligned}$$

Here we use the formula

$$\ln(a^k) = k\ln a.$$

is $\quad 2 = 3^k, \qquad$ i.e., $\quad \ln 2 = k\,\ln 3, \quad$ or $\quad k = \dfrac{\ln 2}{\ln 3} \approx 0.63093.$

You might want to compute the $\frac{\ln 2}{\ln 3}$-dimensional volume of C:

$$\operatorname{vol}_{\frac{\ln 2}{\ln 3}} C = \lim_{j\to\infty} 2^j \left(\frac{1}{3}\right)^{j\frac{\ln 2}{\ln 3}}. \tag{1}$$

In addition to $\ln(a^k) = k\ln a$, equation 2 uses the formula

$$\ln(ab) = \ln a + \ln b.$$

The 2 gets exponent j instead of $j\frac{\ln 2}{\ln 3}$ because it is the *number* of pieces contained in C: C_1 has two pieces, C_2 has four pieces, etc., so C_j has 2^j pieces. Now take the log of each side of equation 1:

$$\begin{aligned}\ln\left(\operatorname{vol}_{\frac{\ln 2}{\ln 3}} C\right) &= \ln\left(\lim_{j\to\infty} 2^j\left(\frac{1}{3}\right)^{j\frac{\ln 2}{\ln 3}}\right) = \lim_{j\to\infty} j\ln 2 + j\frac{\ln 2}{\ln 3}(-\ln 3)\\ &= \lim_{j\to\infty}(j\ln 2 - j\ln 2) = 0,\end{aligned} \tag{2}$$

so $\operatorname{vol}_{\frac{\ln 2}{\ln 3}} C = 1$.

SOLUTIONS FOR REVIEW EXERCISES, CHAPTER 5

5.1 To find a parametrization for some region, we ask, "how should we describe where we are at some point in the region?" Using the trigonometric formulas discussed in "Note on parametrizations" (immediately before Solution 5.2.1), we see that the map

$$f : u \mapsto \begin{pmatrix} x = R + r\cos u \\ y = 0 \\ z = r\sin u \end{pmatrix}$$

parametrizes the circle in the (x, z)-plane of radius r centered at $\begin{pmatrix} R \\ 0 \\ 0 \end{pmatrix}$. Imagine that circle (the small circle at the top of the figure in the margin) traveling in a perfect circle around the z-axis, to form a torus. The shaded annulus at the bottom of the figure is the projection of this torus onto the (x, y)-plane.

As the point

$$\begin{pmatrix} R + r\cos u \\ r\sin u \end{pmatrix}$$

in the little circle turns by angle v around the z-axis, its position is given by

$$\begin{pmatrix} (R + r\cos u)\cos v \\ (R + r\cos u)\sin v \\ r\sin u \end{pmatrix}.$$

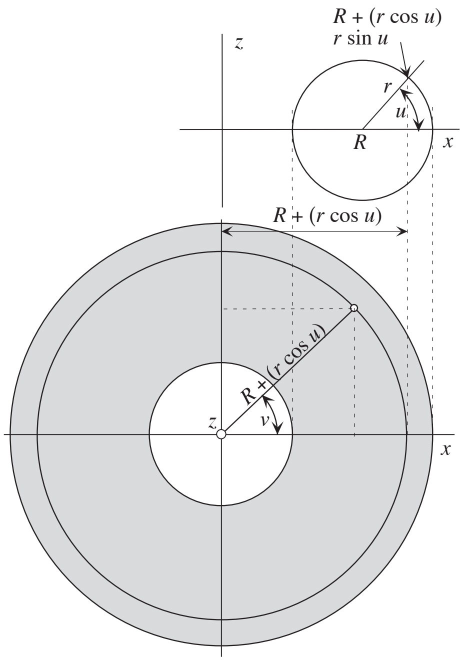


FIGURE FOR SOLUTION 5.1.

TOP: In the (x, z)-plane, we travel on a circle of radius r centered at R. Polar coordinates describe where we are on that circle. As this small circle orbits the z-axis, it forms a torus. BOTTOM: The shaded annulus is the projection of that torus onto the (x, y)-plane.

5.3 As we did in Solution 5.3.1, we return to Cartesian coordinates, setting $x = r(t)\cos\theta(t)$ and $y = r(t)\sin\theta(t)$; since $r(t) = 1/t^\alpha$ and $\theta(t) = t$, this gives

$$\gamma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \frac{1}{t^\alpha}\cos t \\ \frac{1}{t^\alpha}\sin t \end{pmatrix}.$$

So the length of the spiral is

$$\begin{aligned} \int_1^\infty |\gamma'(t)|\,dt &= \int_1^\infty \left| \begin{bmatrix} -t^{-\alpha}\sin t - \alpha t^{-\alpha-1}\cos t \\ t^{-\alpha}\cos t - \alpha t^{-\alpha-1}\sin t \end{bmatrix} \right| dt \\ &= \int_1^\infty \sqrt{t^{-2\alpha} + \alpha^2 t^{-2\alpha-2}}\,dt \\ &= \int_1^\infty \frac{1}{t^{\alpha+1}}\sqrt{\alpha^2 + t^2}\,dt. \end{aligned}$$

For t sufficiently large,

$$\sqrt{\alpha^2 + t^2} = t\sqrt{1 + \frac{\alpha^2}{t^2}} \quad \text{satisfies} \quad \frac{1}{2}t < t\sqrt{1 + \frac{\alpha^2}{t^2}} < 2t;$$

i.e., it is approximately t. Thus for t sufficiently large, the integrand $\frac{1}{t^{\alpha+1}}\sqrt{\alpha^2+t^2}$ behaves like $\frac{t}{t^{\alpha+1}} = \frac{1}{t^{\alpha}}$, so if $\alpha > 1$, the length is finite; if $\alpha \le 1$, it is not.

5.5 a. The surface is parametrized by $\gamma : \begin{pmatrix} x \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} x \\ f(x)\cos\theta \\ f(x)\sin\theta \end{pmatrix}$ (see equation 5.2.9).

b. The surface area is given by

$$\int_0^{2\pi}\int_a^b \underbrace{\left| \begin{bmatrix} 1 \\ f'(x)\cos\theta \\ f'(x)\sin\theta \end{bmatrix} \times \begin{bmatrix} 0 \\ -f(x)\sin\theta \\ f(x)\cos\theta \end{bmatrix} \right|}_{\sqrt{|[\mathbf{D}\gamma(\mathbf{u})]^\top[\mathbf{D}\gamma(\mathbf{u})]|}\ \text{(equation 5.3.33)}} dx\, d\theta = 2\pi \int_a^b f(x)\sqrt{1+(f'(x))^2}\, dx.$$

5.7 Recall that the total curvature $\mathbf{K}(S)$ of a surface $S \subset \mathbb{R}^3$ is defined in Exercise 5.3.11 to be $\mathbf{K}(S) = \int_S |K(\mathbf{x})||d^2\mathbf{x}|$, where K is the Gaussian curvature. The helicoid is parametrized by

$$\begin{pmatrix} r \\ z \end{pmatrix} \mapsto \begin{pmatrix} r\cos z \\ r\sin z \\ z \end{pmatrix}.$$

Its Gaussian curvature was computed in Example 3.9.13 to be $-1/(1+r^2)^2$. Thus our total curvature is

$$\begin{aligned}\int_0^a\int_0^\infty \frac{1}{(1+r^2)^2}\left| \begin{bmatrix} \cos z \\ \sin z \\ 0 \end{bmatrix} \times \begin{bmatrix} -r\sin z \\ r\cos z \\ 1 \end{bmatrix} \right| |dr\, dz| &= \int_0^a\int_0^\infty \frac{1}{(1+r^2)^{3/2}}\, dr\, dz \\ &= a\int_0^\infty \frac{dr}{(1+r^2)^{3/2}} = a\left[\frac{r}{\sqrt{1+r^2}}\right]_0^\infty \\ &= a.\end{aligned}$$

5.9 a. In base 5, the number $1/5$ is written 0.1, $2/5$ is written 0.2, and so on, so removing the open middle fifth is equivalent to removing the numbers that start with 0.2 in base 5, except for 0.2 itself, which can be written as the limit of the sequence $0.14, 0.144, 0.1444, \dots$. Removing the open middle fifths of the four remaining fifths is equivalent to removing the numbers that can only be written beginning $0.02, 0.12, 0.32$, or 0.42, and so on. One way to show that C is an uncountable set is to throw out all the numbers in C containing the digits 3 or 4. What remains is all the real numbers, written in base 2. Since the reals are uncountable (see Section 0.6), the elements of C are uncountable.

b. For C to be pavable, its indicator function $\mathbf{1}_C$ must be integrable (Definition 4.1.18). Theorem 4.3.9 says that a bounded function with bounded support is integrable if it is continuous except on a set of volume 0. Certainly $\mathbf{1}_C$ is bounded, since $\mathbf{1}_C(x)$ can only equal 0 or 1, and its support is

$[0,1]$. Since $\mathbf{1}_C$ equals 0 everywhere except on C, and C is closed, $\mathbf{1}_C(x)$ is continuous except on C, so if we can show that C has length 0, we will have also shown that C is pavable.

By Proposition 4.1.23, we can do this by showing that for every $\epsilon > 0$, we can put every element of C in a 1-dimensional dyadic cube such that the total length of the cubes is $\leq \epsilon$. We will actually show that we can do this using non-dyadic cubes, and will not bother to spell out how we could convert that argument to dyadic cubes.

Consider the sequence $i \mapsto C_i$, where C_1 is the interval $[0,1]$ with the open middle fifth removed, C_2 is the interval at the next level, and so on. Thus $\cap_{j\to\infty} C_j = C$. Further, C_1 has length $4/5$, while C_2 has length $16/25 = (4/5)^2$, and so on. These sets include every element of C, and

$$\operatorname{vol}_1 C = \lim_{j\to\infty} \operatorname{vol}_1 C_j = \lim_{j\to\infty} \left(\frac{4}{5}\right)^j = 0.$$

c. Note that C is made up of two identical segments, from 0 to $2/5$ and from $3/5$ to 1. But stretching the first segment by a factor of $5/2$ would make it identical to C. So

$$\operatorname{vol}_k C = 2 \operatorname{vol}_k A$$
$$\operatorname{vol}_k C = (5/2)^k \operatorname{vol}_k A.$$

Note that this dimension is a little bit bigger than the dimension of the triadic Cantor set in Exercise 5.5.1, which was about .63; this set is a little bit "thicker," more like a line.

Therefore,

$$\left(\frac{5}{2}\right)^k = 2, \quad \text{i.e.,} \quad k = \frac{\log 2}{\log 5/2} = \frac{\log 2}{\log 5 - \log 2} \approx 0.756471.$$

Solutions for Chapter 6

6.1.1 a. On $\mathbb{R}^3$ there are three elementary 1-forms, dx, dy, and dz; three elementary 2-forms, $dx \wedge dy$, $dx \wedge dz$, and $dy \wedge dz$; one elementary 0-form, the number 1; and one elementary 3-form, $dx \wedge dy \wedge dz$.

b. On $\mathbb{R}^4$ there are

1 elementary 0-form:	the number 1.
4 elementary 1-forms:	dx_1, dx_2, dx_3, and dx_4.
6 elementary 2-forms:	$dx_1 \wedge dx_2$, $dx_1 \wedge dx_3$, $dx_1 \wedge dx_4$, $dx_2 \wedge dx_3$, $dx_2 \wedge dx_4$, and $dx_3 \wedge dx_4$.
4 elementary 3-forms:	$dx_1 \wedge dx_2 \wedge dx_3$, $dx_1 \wedge dx_2 \wedge dx_4$, $dx_1 \wedge dx_3 \wedge dx_4$, and $dx_2 \wedge dx_3 \wedge dx_4$.
1 elementary 4-form:	$dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4$.

c. Since $\binom{m}{k} = \binom{5}{4} = \dfrac{5!}{4!1!} = 5$, there are five elementary 4-forms on $\mathbb{R}^5$:

$$dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \qquad dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_5 \qquad dx_1 \wedge dx_2 \wedge dx_4 \wedge dx_5$$

$$dx_1 \wedge dx_3 \wedge dx_4 \wedge dx_5 \qquad dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_5$$

6.1.3 a. $\det \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = 0$

b. $\det \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} + 2 \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 1 + 2 \cdot (-1) = -1.$

c. $\det \begin{bmatrix} 2 & 2 \\ 0 & -3 \end{bmatrix} = -6.$

d. The answer is 0; no need to compute $\det \begin{bmatrix} 1 & -2 & 2 \\ 3 & 1 & 2 \\ 3 & 1 & 2 \end{bmatrix} = 0$; exchanging the two dx_2 changes the sign of the answer while leaving it unchanged.

6.1.5 a. The y-component is -3.

b. The x_2, x_4-component of signed volume is

$$dx_2 \wedge dx_4 \begin{bmatrix} 2 & -2 \\ 1 & 1 \\ 0 & 2 \\ 4 & -3 \end{bmatrix} = \det \begin{bmatrix} 1 & 1 \\ 4 & -3 \end{bmatrix} = -7.$$

6.1.7 a. On $\mathbb{R}^4$ there are

$$\text{1 elementary 0-form, since } \binom{4}{0} = \frac{4!}{0!4!} = 1$$

$$\text{4 elementary 1-forms, since } \binom{4}{1} = \frac{4!}{1!3!} = 4$$

$$\text{6 elementary 2-forms, since } \binom{4}{2} = \frac{4!}{2!2!} = 6$$

$$4 \text{ elementary 3-forms, since } \binom{4}{3} = \frac{4!}{3!1!} = 4$$

$$1 \text{ elementary 4-form, since } \binom{4}{4} = \frac{4!}{4!0!} = 1$$

b. On $\mathbb{R}^5$ there are

$$5 \text{ elementary 1-forms, since } \binom{5}{1} = \frac{5!}{1!4!} = 5$$

$$10 \text{ elementary 2-forms, since } \binom{5}{2} = \frac{5!}{2!3!} = 10$$

$$10 \text{ elementary 3-forms, since } \binom{5}{3} = \frac{5!}{3!2!} = 10$$

$$5 \text{ elementary 4-forms, since } \binom{5}{4} = \frac{5!}{4!1!} = 5$$

$$1 \text{ elementary 5-form, since } \binom{5}{5} = \frac{5!}{5!0!} = 1.$$

Solution 6.1.7, part b: the 5 elementary 1-forms on $\mathbb{R}^5$ are

$$dx_1,\ dx_2,\ dx_3,\ dx_4,\ dx_5.$$

The 10 elementary 2-forms on $\mathbb{R}^5$ are

$dx_1 \wedge dx_2$; $dx_1 \wedge dx_3$; $dx_1 \wedge dx_4$;

$dx_1 \wedge dx_5$;

$dx_2 \wedge dx_3$; $dx_2 \wedge dx_4$; $dx_2 \wedge dx_5$;

$dx_3 \wedge dx_4$; $dx_3 \wedge dx_5$;

$dx_4 \wedge dx_5$.

The 10 elementary 3-forms on $\mathbb{R}^5$ are

$dx_1 \wedge dx_2 \wedge dx_3$;

$dx_1 \wedge dx_2 \wedge dx_4$;

$dx_1 \wedge dx_2 \wedge dx_5$;

$dx_1 \wedge dx_3 \wedge dx_4$

$dx_1 \wedge dx_3 \wedge dx_5$

$dx_1 \wedge dx_4 \wedge dx_5$

$dx_2 \wedge dx_3 \wedge dx_4$;

$dx_2 \wedge dx_3 \wedge dx_5$

$dx_2 \wedge dx_4 \wedge dx_5$

$dx_3 \wedge dx_4 \wedge dx_5$.

The 5 elementary 4-forms on $\mathbb{R}^5$ are

$dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4$;

$dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_5$;

$dx_1 \wedge dx_2 \wedge dx_4 \wedge dx_5$;

$dx_1 \wedge dx_3 \wedge dx_4 \wedge dx_5$

$dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_5$.

The single elementary 5-form on $\mathbb{R}^5$ is

$$dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4 \wedge dx_5.$$

6.1.9 Note that the second index on the $v_{i,j}$ is the vector. Thus $v_{i,1}$ in equation 6.1.23 is v_i in the language of equation 6.1.22 and $v_{j,2}$ is w_j in the language of equation 6.1.22. For $k = 2$ and $n = 3$ we have

$$\begin{aligned}
&\sum_{1 \le i_1 < \cdots < i_k \le n} \Big(\sum_{\sigma \in \operatorname{Perm}(k)} (\operatorname{sgn}\sigma)\ v_{i_1,\sigma(1)} \cdots v_{i_k,\sigma(k)} \varphi(\vec{\mathbf{e}}_{i_1}, \dots, \vec{\mathbf{e}}_{i_k}) \Big) \\
&= +\big(\underbrace{v_{1,1}v_{2,2}}_{v_1 w_2} - \underbrace{v_{1,2}v_{2,1}}_{w_1 v_2}\big)\varphi(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2) + \big(\underbrace{v_{2,1}v_{3,2}}_{v_2 w_3} - \underbrace{v_{2,2}v_{3,1}}_{w_2 v_3}\big)\varphi(\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3) \\
&\quad + \big(\underbrace{v_{1,1}v_{3,2}}_{v_1 w_3} - \underbrace{v_{1,2}v_{3,1}}_{w_1 v_3}\big)\varphi(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_3).
\end{aligned}$$

(We first consider $i_1 = 1$ and $i_k = 2$, then $i_1 = 2$ and $i_k = 3$, then $i_1 = 1$ and $i_k = 3$.)

6.1.11 If φ and ψ are 2-forms, then

$$\begin{aligned}
\varphi \wedge \psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4) = {} & \varphi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)\psi(\vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4) - \varphi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_3)\psi(\vec{\mathbf{v}}_2, \vec{\mathbf{v}}_4) \\
& + \varphi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_4)\psi(\vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3) + \varphi(\vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)\psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_4) \\
& - \varphi(\vec{\mathbf{v}}_2, \vec{\mathbf{v}}_4)\psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_3) + \varphi(\vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4)\psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2).
\end{aligned}$$

6.1.13 1. *Distributivity.* This is just busy work. Let $S(k,l) \subset \operatorname{Perm}_{(k+l)}$ be the set of (k,l)-shuffles, i.e., the set of permutations σ of $1, \dots, k+l$ such that

$$\sigma(1) < \cdots < \sigma(k),\ \sigma(k+1) < \cdots < \sigma(k+l).$$

If $\varphi \in A^k_{const}(E)$ and $\psi_1, \psi_2 \in A^l_{const}(E)$, then

$$\begin{aligned}
\varphi \wedge (\psi_1 + \psi_2)(\mathbf{v}_1, \dots, \mathbf{v}_{k+l}) &= \sum_{\sigma \in S(k,l)} \operatorname{sgn}(\sigma)\varphi(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k)})(\psi_1 + \psi_2)(\mathbf{v}_{\sigma(k+1)}, \dots, \mathbf{v}_{\sigma(k+l)}) \\
&= \sum_{\sigma \in S(k,l)} \operatorname{sgn}(\sigma)\varphi(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k)})\Big(\psi_1(\mathbf{v}_{\sigma(k+1)}, \dots, \mathbf{v}_{\sigma(k+l)}) + \psi_2(\mathbf{v}_{\sigma(k+1)}, \dots, \mathbf{v}_{\sigma(k+l)}\Big) \\
&= \sum_{\sigma \in S(k,l)} \operatorname{sgn}(\sigma)\varphi(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k)})\psi_1(\mathbf{v}_{\sigma(k+1)}, \dots, \mathbf{v}_{\sigma(k+l)}) \\
&\qquad + \sum_{\sigma \in S(k,l)} \operatorname{sgn}(\sigma)\varphi(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k)})\psi_2(\mathbf{v}_{\sigma(k+1)}, \dots, \mathbf{v}_{\sigma(k+l)}) \\
&= (\varphi \wedge \psi_1)(\mathbf{v}_1, \dots, \mathbf{v}_{k+l}) + (\varphi \wedge \psi_2)(\mathbf{v}_1, \dots, \mathbf{v}_{k+l}).
\end{aligned}$$

2. *Associativity.* The proof of associativity is not quite so easy. Let us introduce the notation $S(k_1, k_2, k_3) \subset \operatorname{Perm}_{(k_1+k_2+k_3)}$ for the triple shuffles, i.e., permutations $\sigma \in \operatorname{Perm}_{(k_1+k_2+k_3)}$ such that $\sigma(i) < \sigma(j)$ whenever

$$1 \le i < j \le k_1, \text{ or } k_1 + 1 \le i < j \le k_1 + k_2, \text{ or } k_1 + k_2 + 1 \le i < j \le k_1 + k_2 + k_3.$$

Imagine you have a pack of 39 cards, say the spades, hearts, and diamonds of a standard deck, originally arranged with first the diamonds, then the hearts, then the spades, each arranged from lowest to highest. Then the set of shuffles $S(13, 13, 13)$ consists of the arrangements of the deck such that a lower spade comes sooner than a higher spade, and similarly for the hearts and diamonds, but nothing is said about the position of a heart relative to a spade.

The key to associativity is the observation that any such arrangement can be achieved uniquely by first shuffling together the hearts and spades, and then shuffling that pack with the diamonds, or equivalently, first shuffling the diamonds and hearts, and then shuffling that pack with the pack of spades.

To help translate this into mathematics, let us further invent the notation

$$S(\widehat{k_1}, k_2, k_3) \subset S(k_1, k_2, k_3)$$

to denote the subset of the triple shuffles σ such that $\sigma(i) = i$ when $i \le k_1$, and similarly $S(k_1, k_2, \widehat{k_3}) \subset S(k_1, k_2, k_3)$ the triple shuffles such that $\sigma(i) = i$ when $k_1 + k_2 + 1 \le i \le k_1 + k_2 + k_3$, like the shuffles of the diamonds and hearts which don't affect the spades, and the shuffles of hearts and spades that don't affect the diamonds.

The principle above can be restated in mathematical terms:

Lemma 1. *Every element $\sigma \in S(k_1, k_2, k_3)$ can be written uniquely as*

$$\sigma = \tau_2 \circ \tau_1$$

with $\tau_1 \in S(\widehat{k_1}, k_2, k_3)$ and $\tau_2 \in S(k_1, k_2 + k_3)$, and also as

$$\sigma = \tau_2' \circ \tau_1'$$

with $\tau_1' \in S(k_1, k_2, \widehat{k_3})$ and $\tau_2 \in S(k_1 + k_2, k_3)$.

Using this notation and Lemma 1, associativity is the following:

$$\begin{aligned}
&\varphi_1 \wedge (\varphi_2 \wedge \varphi_3)(\mathbf{v}_1, \dots, \mathbf{v}_{k_1+k_2+k_3})\\
&= \sum_{\sigma \in S(k_1, k_2+k_3)} \operatorname{sgn}(\sigma)\varphi_1(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k_1)})\\
&\quad \left(\sum_{\tau \in S(\widehat{k}_1, k_2, k_3)} \operatorname{sgn}(\tau)\Big(\varphi_2(\mathbf{v}_{\sigma(\tau(k_1+1))}, \dots, \mathbf{v}_{\sigma(\tau(k_1+k_2))})\Big)\Big(\varphi_3(\mathbf{v}_{\sigma(\tau(k_1+k_2+1))}, \dots, \mathbf{v}_{\sigma(\tau(k_1+k_2+k_3))})\Big)\right)\\
&= \sum_{\sigma \in S(k_1, k_2+k_3)} \sum_{\tau \in S(\widehat{k}_1, k_2, k_3)} \operatorname{sgn}(\tau)\operatorname{sgn}(\sigma)\varphi_1(\mathbf{v}_{\sigma(\tau(1))}, \dots, \mathbf{v}_{\sigma(\tau(k_1))})\\
&\quad \left(\Big(\varphi_2(\mathbf{v}_{\sigma(\tau(k_1+1))}, \dots, \mathbf{v}_{\sigma(\tau(k_1+k_2))})\Big)\Big(\varphi_3(\mathbf{v}_{\sigma(\tau(k_1+k_2+1))}, \dots, \mathbf{v}_{\sigma(\tau(k_1+k_2+k_3))})\Big)\right)\\
&= \sum_{\alpha \in S(k_1, k_2, k_3)} \operatorname{sgn}(\alpha)\varphi_1(\mathbf{v}_{\alpha(1)}, \dots, \mathbf{v}_{\alpha(k_1)})\varphi_2(\mathbf{v}_{\alpha(k_1+1)}, \dots, \mathbf{v}_{\alpha(k_1+k_2)})\varphi_3(\mathbf{v}_{\alpha(k_1+k_2+1)}, \dots, \mathbf{v}_{\alpha(k_1+k_2+k_3)})
\end{aligned}$$

The second equality uses the fact that $\sigma(\tau(i)) = \sigma(i)$ when $1 \le i \le k_1$, since $\tau \in S(\widehat{k}_1, k_2, k_3)$. The third uses Lemma 1, i.e., that the $\sigma \circ \tau$ are exactly running through the elements of $S(k_1, k_2, k_3)$, and the fact that $\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma \circ \tau)$.

An exactly parallel development shows that

$$\big((\varphi_1 \wedge \varphi_2) \wedge \varphi_3\big)(\mathbf{v}_1, \dots, \mathbf{v}_{k_1+k_2+k_3})$$

is also the same sum over the triple shuffles. This completes the discussion of associativity.

3. *Skew commutativity.* For skew commutativity, we want again to look carefully at permutations. Let $\alpha \in \operatorname{Perm}_{(k_1+k_2)}$ be the permutation that puts the last indices at the beginning and the first indices at the end, i.e.,

$$\alpha(i) = \begin{cases} i + k_1 & \text{if } i \le k_2\\ i - k_2 & \text{if } i > k_2. \end{cases}$$

Then every permutation in $\tau \in S(k_2, k_1)$ can be written uniquely $\tau = \sigma \circ \alpha$ with $\sigma \in S(k_1, k_2)$.

Using this, we see that

$$(\varphi \wedge \psi)(\mathbf{v}_1, \dots, \mathbf{v}_{k_1+k_2}) = \sum_{\sigma \in S(k_1,k_2)} \operatorname{sgn}(\sigma)\varphi(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k_1)})\psi(\mathbf{v}_{\sigma(k_1+1)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)}),$$

whereas

$$\begin{aligned}
(\psi \wedge \varphi)(\mathbf{v}_1, \dots, \mathbf{v}_{k_1+k_2}) &= \sum_{\tau \in S(k_2,k_1)} \psi(\mathbf{v}_{\tau(1)}, \dots, \mathbf{v}_{\tau(k_2)})\varphi(\mathbf{v}_{\tau(k_2+1)}, \dots, \mathbf{v}_{\tau(k_1+k_2)})\\
&= \sum_{\sigma \in S(k_1,k_2)} \operatorname{sgn}(\sigma \circ \alpha)\psi(\mathbf{v}_{\sigma\circ\alpha(1)}, \dots, \mathbf{v}_{\sigma\circ\alpha(k_2)})\psi(\mathbf{v}_{\sigma\circ\alpha(k_2+1)}, \dots, \mathbf{v}_{\sigma\circ\alpha(k_1+k_2)})\\
&= \operatorname{sgn}(\alpha) \sum_{\sigma \in S(k_1,k_2)} \operatorname{sgn}(\sigma)\psi(\mathbf{v}_{\sigma(k_1+1)}, \dots, \mathbf{v}_{\sigma(k_1+k_2)})\psi(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{k_1}).
\end{aligned}$$

So we need to show that $\operatorname{sgn}(\alpha) = (-1)^{k_1 k_2}$. We can do this by counting transpositions: imagine switching k_1+1 first with k_1, then with k_1-1, and so forth, until it arrives in the first position. We will have made k_1 transpositions. Now move k_1+2 into second position; this will require another k_1 transpositions. Now move k_1+3 into third position, ..., and finally k_1+k_2 into position k_2. This will realize α as $k_2 k_1$ transpositions.

6.1.15 a. $\sin(x_4)\det\begin{bmatrix} 3 & -1 \\ 2x_1 & x_4 \end{bmatrix} = \sin x_4(3x_4 + 2x_1)$

b. $e^x \det[2] = 2e^x$ c. $(-x_1)^2 e^{x_2}\det\begin{bmatrix} 3 & -1 & -1 \\ 2 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} = 7x_1^2 e^{x_2}$

6.2.1 a.

$$\int_{\gamma(I)} x\,dy + y\,dz = \int_I (x\,dy + y\,dz)\left(P_{\begin{pmatrix}\sin t\\ \cos t\\ t\end{pmatrix}}\left(\begin{bmatrix}\cos t\\ -\sin t\\ 1\end{bmatrix}\right)\right)dt$$

$$= \int_{-1}^{1}\big(-(\sin t)^2 + \cos t\big)dt = \int_{-1}^{1}\left(\frac{1}{2}\cos 2t - \frac{1}{2} + \cos t\right)dt$$

$$= \left[\frac{\sin 2t}{4} - \frac{1}{2}t + \sin t\right]_{-1}^{1} = \frac{2\sin 2}{4} - 1 + 2\sin 1$$

$$= \frac{\sin 2}{2} + 2\sin 1 - 1.$$

To get the second equality in the second line, we use

$$\begin{aligned}\cos 2t &= \cos^2 t - \sin^2 t\\ &= 2\cos^2 t - 1\\ &= 1 - 2\sin^2 t.\end{aligned}$$

In the second line, the variable x is replaced by $\sin t$, and y is replaced by $\cos t$, since the parallelogram is anchored at $\begin{pmatrix}\sin t\\ \cos t\\ t\end{pmatrix}$.

b. To get the last equality below, we used MAPLE:

$$\int_{\gamma(U)} x_1\,dx_2\wedge dx_3 + x_2\,dx_3\wedge dx_4$$

$$= \int_U x_1\,dx_2\wedge dx_3 + x_2\,dx_3\wedge dx_4\left(P_{\begin{pmatrix}uv\\ u^2+v^2\\ u-v\\ \ln(u+v+1)\end{pmatrix}}\left(\begin{bmatrix}v\\ 2u\\ 1\\ 1/(u+v+1)\end{bmatrix},\begin{bmatrix}u\\ 2v\\ -1\\ 1/(u+v+1)\end{bmatrix}\right)\right)du\,dv$$

$$= \int_U\left(uv\det\begin{bmatrix}2u & 2v\\ 1 & -1\end{bmatrix} + (u^2+v^2)\det\begin{bmatrix}1 & -1\\ 1/(u+v+1) & 1/(u+v+1)\end{bmatrix}\right)|du\,dv|$$

$$= \int_U\left(-uv(2u+2v) + 2\frac{u^2+v^2}{1+u+v}\right)du\,dv$$

$$= \int_0^2\left(\int_0^{2-u}(-2u^2v - 2uv^2)\,dv\right)du + 2\int_0^2\int_0^{2-u}\left(\frac{u^2+v^2}{1+u+v}\,dv\right)du = \frac{64}{45} - \frac{4}{3}\ln 3$$

It is possible to compute the integral by hand, but it requires a lot of computing, the kind of thing that computers are more reliable at than people.

6.2.3 a.

$$\int_{\gamma(U)} (x_1 + x_4)\, dx_2 \wedge dx_3 = \int_U (e^u + \sin v) \det \begin{bmatrix} 0 & -e^{-v} \\ -\sin u & 0 \end{bmatrix} |du\, dv|$$
$$= -\int_0^1 \int_{-u}^{u} (e^u + \sin v)(e^v \sin u)\, dv\, du$$

b.

$$\int_U (u-v)(w-v) dx_1 \wedge dx_3 \wedge dx_4 \left(\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right) |du\, dv\, dw|$$
$$= \int_U (u-v)(w-v) \det \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} |du\, dv\, dw|$$
$$= \int_U (u-v)(w-v) 2\, |du\, dv\, dw|$$
$$= \int_0^1 \left(\int_{-(1-w)}^{1-w} \left(\int_{-\sqrt{(w-1)^2 - v^2}}^{\sqrt{(w-1)^2-v^2}} 2(u-v)(w-v)\, du \right) dv \right) dw.$$

Solution 6.2.3: Actually, the problem only asks you to go to the next-to-last line of both computations.

In part b, the final step is better done passing to cylindrical coordinates.

6.2.5 We will identify a point in $\mathbb{C}^3$ with the point $\begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \end{pmatrix} \in \mathbb{R}^6$. Then the point $\begin{pmatrix} z \\ z^2 \\ z^3 \end{pmatrix} \in \mathbb{C}^3$ is identified to $\begin{pmatrix} x \\ y \\ x^2 - y^2 \\ 2xy \\ x^3 - 3xy^2 \\ 3x^2y - y^3 \end{pmatrix}$, since the real and imaginary parts of z^2 are $x^2 - y^2$ and $2xy$ respectively, and the real and imaginary parts of z^3 are $x^3 - 3xy^2$ and $3x^2y - y^3$ respectively. Then

$$\overrightarrow{D_x}\gamma(z) = \begin{bmatrix} 1 \\ 0 \\ 2x \\ 2y \\ 3x^2 - 3y^2 \\ 6xy \end{bmatrix} \quad \text{and} \quad \overrightarrow{D_y}\gamma(z) = \begin{bmatrix} 0 \\ 1 \\ -2y \\ 2x \\ -6xy \\ 3x^2 - 3y^2 \end{bmatrix}.$$

If we set

$$\varphi = dx_1 \wedge dy_1 \;+\; dx_2 \wedge dy_2 \;+\; dx_3 \wedge dy_3,$$

Definition 6.2.1 then gives

$$\begin{aligned}\int_{[\gamma(S)]}\varphi &= \int_S \varphi\left(P_{\gamma(z)}(\overrightarrow{D_x}\gamma(z), \overrightarrow{D_y}\gamma(z))\right)|dx|\,|dy|\\ &= \int_S \det\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \det\begin{bmatrix}2x & -2y\\ 2y & 2x\end{bmatrix} + \det\begin{bmatrix}3x^2-3y^2 & -6xy\\ 6xy & 3x^2-3y^2\end{bmatrix}|dx|\,|dy|\\ &= \int_S \big(1+4x^2+4y^2+9x^4+9y^4+18x^2y^2\big)|dx|\,|dy|.\end{aligned}$$

Since S is a disc, we now pass to polar coordinates. This gives

$$\begin{aligned}\int_S &\big(1+4(x^2+y^2)+9(x^2+y^2)^2\big)|dx|\,|dy|\\ &= \int_0^{2\pi}\left(\int_0^1 \Big(1+4r^2(\cos^2\theta+\sin^2\theta)+9r^4(\cos^2\theta+\sin^2\theta)\Big)r\,dr\right)d\theta\\ &= \int_0^{2\pi}\left[\frac{r^2}{2}+r^4+\frac{9r^6}{6}\right]_0^1 d\theta\\ &= \int_0^{2\pi}\left(\frac{1}{2}+1+\frac{9}{6}\right)d\theta = 6\pi.\end{aligned}$$

6.2.7 The parallelogram $P_{\mathbf{x}}(h\vec{\mathbf{v}}_1, \dots, h\vec{\mathbf{v}}_k)$ is parametrized by

$$\gamma : \begin{pmatrix}t_1\\ \vdots\\ t_k\end{pmatrix} \mapsto \mathbf{x} + t_1\vec{\mathbf{v}}_1 + \cdots + t_k\vec{\mathbf{v}}_k, \quad 0 \le t_i \le h.$$

So

$$\int_{[P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)]}\varphi = \int_0^h \cdots \int_0^h \varphi\left(P_{\gamma(\mathbf{t})}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\right)\,dt_1\dots dt_k,$$

where on the right the $\vec{\mathbf{v}}_i$ are the partial derivatives of γ.

This proof is valid even if φ is just continuous. If φ is of class C^1, the $o(1)$ can be improved to $O(h)$, and $o(h^k)$ can be improved to $O(h^{k+1})$. The notation $O(h)$ is discussed in Appendix A.11.

Let us see that

$$\varphi\Big(P_{\gamma(\mathbf{t})}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big) - \varphi\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big)$$

tends to 0 when $t \to 0$.

Indeed,

$$\varphi = \sum a_{i_1\dots i_k}dx_{i_1}\wedge\cdots\wedge dx_{i_k}$$

and

$$\lim_{\mathbf{t}\to\mathbf{0}} a_{i_1\dots i_k}\Big(\mathbf{x}+\sum t_i\vec{\mathbf{v}}_i\Big) = a_{i_1\dots i_k}(\mathbf{x}),$$

i.e.,

$$a_{i_1\dots i_k}\left(\mathbf{x}+\sum t_i\vec{\mathbf{v}}_i\right) - a_{i_1\dots i_k}(\mathbf{x}) \in o(1).$$

Thus

$$\int_{[P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)]}\varphi = \int_0^h \cdots \int_0^h \Big(\overbrace{\varphi\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big)}^{\text{constant}} + o(1)\Big)|dt_1\dots dt_k|$$
$$= h^k\varphi\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big) + o(h^k).$$

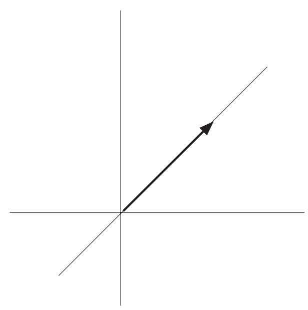


FIGURE FOR SOLUTION 6.3.1
Orientation of the line $x-y=0$ by the vector field $\begin{bmatrix}1\\1\end{bmatrix}$.

6.3.1 The vector field $\begin{bmatrix}1\\1\end{bmatrix}$ is not tangent to the line of equation $x+y=0$, so it does not orient that line. But it does orient the line of equation $x-y=0$, since it is a nonvanishing vector field tangent to that line.

Remark. We could have defined orientation of curves in the plane by *transverse* vector fields; with that definition, $\begin{bmatrix}1\\1\end{bmatrix}$ does orient $x+y=0$, since it is transverse to that line. Indeed, transverse vector fields $\vec{\mathbf{v}}$ don't just orient surfaces in $\mathbb{R}^3$, as described in Proposition 6.3.4; they also orient $(n-1)$-dimensional manifolds M in $\mathbb{R}^n$ for any n, by the formula

$$\Omega^{\vec{\mathbf{v}}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}) = \operatorname{sgn}\det[\vec{\mathbf{v}}(\mathbf{x}),\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}]$$

for any $\mathbf{x}\in M$ and $\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}\in\mathcal{B}(T_{\mathbf{x}}M)$.

In particular, transverse vector fields orient lines in $\mathbb{R}^2$. Let L be the line of equation $x+y=0$, and set $\vec{\mathbf{v}}=\begin{bmatrix}1\\1\end{bmatrix}$. Then for all $\mathbf{x}=\begin{pmatrix}x\\y\end{pmatrix}\in L$ and $a>0$, the vector $\begin{bmatrix}a\\-a\end{bmatrix}$ is a basis of $T_{\mathbf{x}}L$. Then

$$\Omega^{\vec{\mathbf{v}}}\left(\begin{bmatrix}a\\-a\end{bmatrix}\right) = \operatorname{sgn}\det\begin{bmatrix}1 & a\\1 & -a\end{bmatrix} = -1.$$

Thus $\begin{bmatrix}a\\-a\end{bmatrix}$ is indirect for this orientation. $\triangle$

6.3.3 No, there is no constant vector field that defines an orientation of the unit sphere in $\mathbb{R}^3$. Let U be an open subset of $\mathbb{R}^n$ and $f:U\to\mathbb{R}$ a function of class C^1. Suppose that $M\subset\mathbb{R}^n$ is the $(n-1)$-dimensional manifold defined by $f(\mathbf{x})=0$, and that $[\mathbf{D}f(\mathbf{x})]\neq[0]$ for all $\mathbf{x}\in M$. Then a vector field $\vec{F}$ on U is transversal to M if $[\mathbf{D}f(\mathbf{x})]\vec{F}(\mathbf{x})\neq 0$ for all $\mathbf{x}\in M$.

The problem concerns the case where M is the 2-sphere $S^2\subset\mathbb{R}^3$, the function f is defined by $f\begin{pmatrix}x\\y\\z\end{pmatrix}=x^2+y^2+z^2-1$, and $\vec{F}=\vec{\mathbf{v}}$ is a constant vector field.

The constant vector field $\vec{\mathbf{v}}$ is never transversal to S^2, since

$$[\mathbf{D}f(\mathbf{x})]\vec{\mathbf{v}} = 2\vec{\mathbf{x}}\cdot\vec{\mathbf{v}},$$

If $\vec{\mathbf{v}}$ points toward Polaris (the North Star), it is not transverse on the equator; it is tangent to the earth on the equator.

which is 0 if $\mathbf{x}\in\vec{\mathbf{v}}^{\perp}\cap S^2$. The intersection $\vec{\mathbf{v}}^{\perp}\cap S^2$ is nonempty; it is a great circle on the sphere, cutting S^2 into two parts. On the part containing $\frac{\vec{\mathbf{v}}}{|\vec{\mathbf{v}}|}$, the orientation $\Omega^{\vec{\mathbf{v}}}$ is the orientation by the outward-pointing normal, and on the other part (containing $-\frac{\vec{\mathbf{v}}}{|\vec{\mathbf{v}}|}$), it is the orientation by the inward-pointing normal.

6.3.5 A vector $\vec{\mathbf{v}}$ defines an orientation of a plane P if and only if $\vec{\mathbf{v}}$ is not in P. In our case, all four vectors do not belong to P, so they all define orientations. Choose for instance the basis $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ of P. Then

$$\operatorname{sgn}\det\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \operatorname{sgn}(1+1+1) = +1$$

$$\operatorname{sgn}\det\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \operatorname{sgn}(-1+1+1) = +1$$

$$\operatorname{sgn}\det\begin{bmatrix} -1 & 1 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \operatorname{sgn}(-1-1+1) = -1$$

$$\operatorname{sgn}\det\begin{bmatrix} -1 & 1 & 1 \\ -1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix} = \operatorname{sgn}(-1-1-1) = -1$$

Thus $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ define one orientation of P, and $\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$ define the other.

6.3.7 Since $\vec{\mathbf{w}}_1 = 2\vec{\mathbf{v}}_1 - 3\vec{\mathbf{v}}_2$ and $\vec{\mathbf{w}}_2 = \vec{\mathbf{v}}_1 + 2\vec{\mathbf{v}}_2$, we have the change of basis matrix

$$[P_{\{\vec{\mathbf{w}}\}\to\{\vec{\mathbf{v}}\}}] = \begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix}, \quad \text{with} \quad \det[P_{\{\vec{\mathbf{w}}\}\to\{\vec{\mathbf{v}}\}}] = 7 > 0.$$

Since the determinant is positive, the two bases $\vec{\mathbf{w}}_1, \vec{\mathbf{w}}_2$ and $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2$ define the same orientation of V.

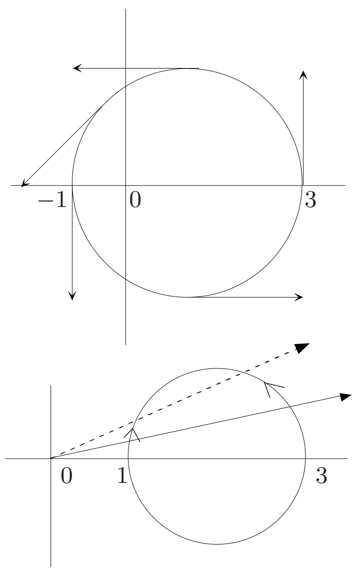


FIGURE FOR SOLUTION 6.3.9. TOP: Part a. The vector field $\vec{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} -y \\ x-1 \end{bmatrix}$ describes the orientation "increasing polar angle" for the circle of equation $(x-1)^2 + y^2 = 4$. BOTTOM: Part b. The circle of equation $(x-2)^2 + y^2 = 1$. As the polar angle increases (from the solid arrow to the dotted arrow), we move simultaneously counterclockwise on part of the circle and clockwise on another part of the circle.

6.3.9 a. The vector field $\vec{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} -y \\ x-1 \end{bmatrix}$ describes the orientation "increasing polar angle" for the circle of equation $(x-1)^2 + y^2 = 4$, shown at left (top). Again, we can confirm this by taking the dot product of this vector and a vector going from the center of the circle to a point on the circle: $\begin{bmatrix} -y \\ x-1 \end{bmatrix} \cdot \begin{bmatrix} x-1 \\ y \end{bmatrix} = 0$. To get the desired unit vector field, we divide $\vec{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} -y \\ x-1 \end{bmatrix}$ by its length, which is $\sqrt{(x-1)^2 + y^2} = 2$, so the answer is $\frac{1}{2}\begin{bmatrix} -y \\ x-1 \end{bmatrix}$.

b. Polar angle is defined from the origin, and every point of the circle of equation $(x-2)^2+y^2=1$ is to the right of the origin. As the polar angle increases, we move simultaneously counterclockwise on part of the circle and clockwise on another part of the circle; see the figure in the margin. Polar angle defined from the origin orients the circle if the origin is inside the circle but does not orient it if the origin is outside the circle.

6.3.11 a. The locus $S \subset \mathbb{R}^4$ is given by the equation $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, where

$$\mathbf{f}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1^2 - x_2^2 - x_3 \\ 2x_1x_2 - x_4 \end{pmatrix}. \text{ The derivative } [\mathbf{Df}(\mathbf{x})] = \begin{bmatrix} 2x_1 & -2x_2 & -1 & 0 \\ 2x_2 & 2x_1 & 0 & -1 \end{bmatrix}$$

is onto $\mathbb{R}^2$ for all values of x_1 and x_2, so S is a surface by Theorem 3.1.10.

b. At the origin, the orientation given by Proposition 6.3.9 is

$$\Omega(\vec{\mathbf{v}}, \vec{\mathbf{w}}) = \operatorname{sgn}\det \begin{bmatrix} 0 & 0 & v_1 & w_1 \\ 0 & 0 & v_2 & w_2 \\ -1 & 0 & v_3 & w_3 \\ 0 & -1 & v_4 & w_4 \end{bmatrix},$$

where $\vec{\mathbf{v}}, \vec{\mathbf{w}}$ is a basis for $T_{\mathbf{0}}S$. The tangent space $T_{\mathbf{0}}S$ is the kernel of $[\mathbf{Df}(\mathbf{0})]$, so any two linearly independent vectors in $\mathbb{R}^4$ whose last two entries vanish are a basis for $T_{\mathbf{0}}S$. For instance, $\vec{\mathbf{e}}_1$ and $\vec{\mathbf{e}}_2$ form such a basis. It is direct for the orientation Ω because

$$\operatorname{sgn}\det \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} = +1.$$

$$B_t = \begin{bmatrix} \cos \pi t & 0 & -\sin \pi t \\ 0 & 1 & 0 \\ \sin \pi t & 0 & \sin \pi t \end{bmatrix}$$

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

SOLUTION 6.3.13, part a: The matrix B_t gives rotation around the y-axis, and the matrix D gives rotation by 45° in the (x,y)-plane. Computing $D^{-1}B_tD$ puts rotation around the y-axis in the "basis" of rotation by 45° in the (x,y)-plane. We could just as well use the matrix giving rotation around the x-axis instead of the matrix B_t giving around the y-axis.

6.3.13 a. Rotate by πt, for $0 \le t \le 1$, around the line $x = y$, $z = 0$. To find this rotation matrix, compute

$$D^{-1}B_tD = \begin{bmatrix} \frac{1}{2}\cos \pi t + \frac{1}{2} & -\frac{1}{2}\cos \pi t + \frac{1}{2} & -\frac{\sqrt{2}}{2}\sin \pi t \\ -\frac{1}{2}\cos \pi t + \frac{1}{2} & \frac{1}{2}\cos \pi t + \frac{1}{2} & -\frac{\sqrt{2}}{2}\sin \pi t \\ \frac{\sqrt{2}}{2}\sin \pi t & -\frac{\sqrt{2}}{2}\sin \pi t & \cos \pi t \end{bmatrix},$$

where B_t is the matrix (shown in the margin) giving rotation around the y-axis and D (also shown in the margin) gives rotation by 45° in the (x,y)-plane. Then $\mathbf{v}(t) = \begin{bmatrix} \frac{1}{2}\cos \pi t + \frac{1}{2} \\ -\frac{1}{2}\cos \pi t + \frac{1}{2} \\ \frac{\sqrt{2}}{2}\sin \pi t \end{bmatrix}$ and $\mathbf{w}(t) = \begin{bmatrix} -\frac{1}{2}\cos \pi t + \frac{1}{2} \\ \frac{1}{2}\cos \pi t + \frac{1}{2} \\ -\frac{\sqrt{2}}{2}\sin \pi t \end{bmatrix}$ are the desired maps.

Here is a different solution to part a.

a. Set $v_1(t) = w_2(t) = 1 - t$ and $v_2(t) = w_1(t) = t$ for $0 \le t \le 1$. Then the functions $t \mapsto \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$ and $t \mapsto \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix}$ are linearly independent except at $t = 1/2$ since

$$\det \begin{bmatrix} v_1(t) & w_1(t) \\ v_2(t) & w_2(t) \end{bmatrix} = \det \begin{bmatrix} 1-t & t \\ t & 1-t \end{bmatrix} = (1-t)^2 - t^2 = 1 - 2t;$$

the determinant is 0 only at $t = 1/2$. Thus if we choose $v_3(t)$ and $w_3(t)$ to satisfy

$$v_3(0) = v_3(1) = w_3(0) = w_3(1) = 0 \quad \text{and} \quad v_3(1/2) \neq w_3(1/2)$$

we will be done. One such choice is $v_3(t) = t(1-t)$ and $w_3(t) = 2t(1-t)$:

$$\mathbf{v} = \begin{bmatrix} 1-t \\ t \\ t(1-t) \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} t \\ 1-t \\ 2t(1-t) \end{bmatrix}.$$

b. If there were such maps $t \mapsto \mathbf{v}(t)$ and $t \mapsto \mathbf{w}(t)$, then

$$\det[\mathbf{v}(0), \mathbf{w}(0)] = \det[\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2] = 1 \text{ and } \det[\mathbf{v}(1), \mathbf{w}(1)] = \det[\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_1] = -1,$$

so by the intermediate value theorem there would exist $t \in [0,1]$ with $\det[\mathbf{v}(t), \mathbf{w}(t)] = 0$.

c. Choose a basis $\vec{\mathbf{u}}_1$, $\vec{\mathbf{u}}_2$, $\vec{\mathbf{u}}_3$, ... , $\vec{\mathbf{u}}_n$ of $\mathbb{R}^n$, and let $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ satisfy $\vec{\mathbf{u}}_i \mapsto \vec{\mathbf{e}}_i$. Then the first two columns of the matrix $\Phi^{-1} T_t \Phi$, where T_t is the matrix in the margin, give the desired maps.

$$T_t = \begin{bmatrix} D^{-1} B_t D & [0] \\ [0] & I \end{bmatrix}$$

Matrix for Solution 6.3.13, part c. The submatrix $D^{-1}B_tD$ is 3×3, the zero matrix in the first column is $(n-3) \times 3$, the other zero matrix is $3 \times (n-3)$, and the identity submatrix is $(n-3) \times (n-3)$.

6.3.15 a. Any vector $\mathbf{v}$ can be written uniquely as

$$\begin{aligned} \mathbf{v} &= c_1 \mathbf{v}_1 + \cdots + c_n \mathbf{v}_n = (a_1 + ib_1)\mathbf{v}_1 + \cdots + (a_n + ib_n)\mathbf{v}_n \\ &= a_1 \mathbf{v}_1 + b_1(i\mathbf{v}_1) + \cdots + a_n \mathbf{v}_n + b_n(i\mathbf{v}_n). \end{aligned}$$

Thus the vectors $\mathbf{v}_1, i\mathbf{v}_1, \dots, \mathbf{v}_n, i\mathbf{v}_n$ span $\mathbb{C}^n$ as a real vector space. The same computation shows that they are linearly independent over $\mathbb{R}$: if $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathbb{R}^n$, and

$$\sum_{j=1}^n (a_j \mathbf{v}_j + b_j(i\mathbf{v}_j)) = \sum_{j=1}^n (a_j + ib_j)\mathbf{v}_j = \vec{\mathbf{0}},$$

then $a_j + ib_j = 0$ for each j so $\vec{\mathbf{a}} = \vec{\mathbf{b}} = \vec{\mathbf{0}}$.

b. Suppose $n = 1$, that the nonzero vectors $\mathbf{v}$, $\mathbf{w}$ are two bases of E, and that $\mathbf{w} = c\mathbf{v} = (a + ib)\mathbf{v}$, where $C = [c]$ is the complex change of basis matrix. Then

$$\mathbf{w} = a\mathbf{v} + b(i\mathbf{v}), \quad i\mathbf{w} = a(i\mathbf{v}) - b\mathbf{v}, \tag{1}$$

so the real change of basis matrix $\widetilde{C}$ is $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, with determinant $\det \widetilde{C} = a^2 + b^2 = |c|^2 = |\det C|^2$.

Now suppose that $n = 2$, and that $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{w}_1, \mathbf{w}_2$ are two bases, with change of basis matrix

$$[P_{\mathbf{w}\to\mathbf{v}}] = C = \begin{bmatrix} c_{1,1} & c_{1.2} \\ c_{2.1} & c_{2,2} \end{bmatrix}$$

so that $\mathbf{w}_1 = c_{1,1}\mathbf{v}_1 + c_{2,1}\mathbf{v}_2$ and $\mathbf{w}_2 = c_{1,2}\mathbf{v}_1 + c_{2,2}\mathbf{v}_2$.

If we write $c_{k,l} = a_{k,l} + ib_{k,l}$, this leads to

$$\begin{aligned} \mathbf{w}_1 &= (a_{1,1} + ib_{1,1})\mathbf{v}_1 + (a_{2,1} + ib_{2,1})\mathbf{v}_2 \\ &= a_{1,1}\mathbf{v}_1 + b_{1,1}(i\mathbf{v}_1) + a_{2,1}\mathbf{v}_2 + b_{2,1}(i\mathbf{v}_2) \\ i\mathbf{w}_1 &= i(a_{1,1} + ib_{1,1})\mathbf{v}_1 + i(a_{2,1} + ib_{2,1})\mathbf{v}_2 \\ &= -b_{1,1}\mathbf{v}_1 + a_{1,1}(i\mathbf{v}_1) - b_{2,1}\mathbf{v}_2 + a_{2,1}(i\mathbf{v}_2). \end{aligned}$$

A similar computation for $\mathbf{w}_2$ and $i\mathbf{w}_2$ leads to the change of basis matrix $\widetilde{C} : \mathbb{R}^4 \to \mathbb{R}^4$

$$\widetilde{C} = \begin{bmatrix} a_{1,1} & -b_{1,1} & a_{1,2} & -b_{1,2} \\ b_{1,1} & a_{1,1} & b_{1,2} & a_{1,2} \\ a_{2,1} & -b_{2,1} & a_{2,2} & -b_{2,2} \\ b_{2,1} & a_{2,1} & b_{2,2} & a_{2,2} \end{bmatrix}, \tag{1}$$

where the first column corresponds to the coefficients with which we have written $\mathbf{w}_i$, the second to the coefficients with which we have written $i\mathbf{w}_i$, and so on.

In the general case, the change of matrix $\widetilde{C} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is

$$\widetilde{C} = \begin{bmatrix} a_{1,1} & -b_{1,1} & \dots & a_{1,n} & -b_{1,n} \\ b_{1,1} & a_{1,1} & \dots & b_{1,n} & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & -b_{n,1} & \dots & a_{n,n} & -b_{n,n} \\ b_{n,1} & a_{n,1} & \dots & b_{n,n} & a_{n,n} \end{bmatrix}.$$

Part b: For $n = 1$, the real change of basis matrix $\widetilde{C}$ that writes $\mathbf{w}$ and $i\mathbf{w}$ in terms of $\mathbf{v}$ and $i\mathbf{v}$ is the 2×2 matrix

$$\begin{bmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{bmatrix},$$

where

$$\begin{aligned} \mathbf{w} &= p_{1,1}\mathbf{v} + p_{2,1}(i\mathbf{v}) \\ i\mathbf{w} &= p_{1,2}\mathbf{v} + p_{2,2}(i\mathbf{v}). \end{aligned}$$

(This uses the notation for the change of basis matrix given in Section 2.6.) So by equation (1), the first column of $\widetilde{C}$ is $\begin{bmatrix} a \\ b \end{bmatrix}$ and the second is $\begin{bmatrix} -b \\ a \end{bmatrix}$. Note the similarity of $\widetilde{C} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ with the matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ giving clockwise rotation by θ. If we write a complex number z as

$$a + ib = r(\cos\theta + i\sin\theta),$$

then multiplying by z (which is a linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$) rotates by θ and expands by a factor of r.

Note that each complex entry $c = a+ib$ has been replaced by the 2×2 real matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

c. We have already done the case $n = 1$. If you try to do the case $n = 2$ by direct computation, developing the determinant of (1) according to the first column or row, you will find that $\det \widetilde{C}$ is a sum of 24 terms, which do come out to be $|\det C|^2$, but the verification of this is very messy. Clearly no such direct approach will work in higher dimensions.

The easiest way to see the general case is by row reduction, and this is easiest in terms of multiplication by elementary matrices. But first we need to show two generalities about the tilde-operation.

If $c = a+ib \in \mathbb{C}$, and we define $\widetilde{c} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, then

$$\widetilde{c_1 + c_2} = \tilde{c}_1 + \tilde{c}_2$$

and

$$\widetilde{c_1 c_2} = \widetilde{c}_1 \widetilde{c}_2.$$

Lemma *Let A, B be two complex matrices. Then*

$$\widetilde{A} + \widetilde{B} = \widetilde{A+B} \quad \textit{and} \quad \widetilde{A}\widetilde{B} = \widetilde{AB}$$

whenever the sum or product is defined.

Proof. The case where A and B are 1×1 matrices was the object of Exercise 1.2.20. For the general case, the formula for addition is obvious. For multiplication, note first that the i,jth entry of $\widetilde{AB}$ is

$$\widetilde{(AB)}_{ij} = \overbrace{\sum_k A_{ik}B_{kj}}^{\text{tilde}} = \sum \widetilde{A_{ik}B_{kj}} = \sum_k \widetilde{A_{ik}}\widetilde{B_{kj}}, \tag{2}$$

where we use an overbrace to denote an extra-wide tilde. The last term in (2) is the product of a "line matrix" whose entries are 2×2 matrices, and a "column matrix" whose entries are 2×2 matrices, all of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Thus we are reduced to showing the result when A is a line matrix and B a column matrix :

We use capital letters for the entries of A and B to distinguish them from the a and b in $c = a+ib$.

$$A = [A_1 \cdots A_m], \qquad B = \begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix}.$$

Then we argue by induction on m; the case $m = 1$ was treated in Exercise 1.2.20. If $m > 1$ we can write

$$\sum_{k=1}^{m} \widetilde{A_k}\widetilde{B_k} = \sum_{k=1}^{m-1} \widetilde{A_k}\widetilde{B_k} + \widetilde{A_m}\widetilde{B_m} \overbrace{=}^{\substack{\text{inductive} \\ \text{hypothesis}}} \sum_{k=1}^{m-1} \widetilde{A_k B_k} + \widetilde{A_m B_m} = \sum_{k=1}^{m} \widetilde{A_k B_k}. \qquad \square$$

Now suppose $C = E_1 \dots E_N$, where the E_k are elementary matrices. Then $\widetilde{C}$ can be written

$$\widetilde{C} = \widetilde{E}_1 \dots \widetilde{E}_n,$$

where (as in part b) each $\widetilde{E}_k$ is obtained from E_k by replacing each complex entry $c = a+ib$ by the 2×2 real matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

If E_k corresponds to adding a multiple of a row onto another, then $\det E_k = \det \widetilde{E_k} = 1$ (both E_k and $\widetilde{E_k}$ are triangular matrices with 1's on the diagonal). If E_k corresponds to multiplying a row through by $c = a+ib$, then

$$E_k = \begin{bmatrix} 1 & & & & \\ & \ddots & & \text{O} & \\ & & c & & \\ & \text{O} & & \ddots & \\ & & & & 1 \end{bmatrix}, \qquad \widetilde{E_k} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} & & & & \\ & \ddots & & \text{O} & \\ & & \begin{bmatrix} a & -b \\ b & a \end{bmatrix} & & \\ & \text{O} & & \ddots & \\ & & & & \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{bmatrix}$$

An elementary matrix that adds a multiple of a row onto another always has det 1; see equation 4.8.28. An elementary matrix that multiplies a row by c has $\det c$; see equation 4.8.16. An elementary matrix that exchanges two rows has $\det -1$; see equation 4.8.29.

so $\det E_k = c$, but $\det \widetilde{E_k} = \det \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = a^2 + b^2 = |c|^2$.

Finally, if E_k corresponds to exchanging two rows, then $\det E_k = -1$, but $\det \widetilde{E_k} = 1$. Thus in all cases, $\det \widetilde{E_k} = |\det E_k|^2$, and we see that

$$\begin{aligned} \det \widetilde{C} &= \det \widetilde{E_1} \cdots \det \widetilde{E_N} = |\det E_1|^2 \cdots |\det E_N|^2 \\ &= |\det E_1 \cdots \det E_N|^2 = |\det C|^2. \end{aligned}$$

d. The criterion for the basis $\mathbf{w}_1, i\mathbf{w}_1, \dots, \mathbf{w}_n, i\mathbf{w}_n$ to be a direct basis is that the change of basis matrix have positive determinant. The change of basis matrix is the matrix $\widetilde{C}$ above, with determinant $|\det C|^2$, which is certainly positive.

6.4.1 No, it does not preserve orientation. The derivative of the function $f(\mathbf{x}) = x^2 + y^2 - z^2$ is $[2x,\ 2y,\ -2z]$, so $\vec{\nabla} f = \begin{bmatrix} 2x \\ 2y \\ -2z \end{bmatrix}$; this is a normal vector field. We have

$$\det \left[\vec{\nabla}\gamma \begin{pmatrix} r \\ \theta \end{pmatrix}, \overrightarrow{D_1\gamma}, \overrightarrow{D_2\gamma} \right] = \det \begin{bmatrix} 2r\cos\theta & \cos\theta & -r\sin\theta \\ 2r\sin\theta & \sin\theta & r\cos\theta \\ -2r & 1 & 0 \end{bmatrix} = -4r^2.$$

Since $-4r^2 < 0$, the parametrization γ does not preserve the orientation given by $\vec{\nabla} f$.

6.4.3 If we take $0 < \theta < 2\pi$ and $-\pi/2 < \varphi < \pi/2$, then

$$\det \left[\vec{\mathbf{n}}, \overrightarrow{D_1\gamma}, \overrightarrow{D_2\gamma} \right] = -\cos\varphi$$

will always be negative; the mapping will be orientation reversing. However, if we take $0 < \theta < 2\pi$ and $\pi/2 < \varphi < 3\pi/2$, the parametrization γ will be orientation preserving.

6.4.5 Parametrize the domain of integration by $\begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ r^k\cos k\theta \\ r^k \sin k\theta \end{pmatrix}$, for $0 \le r \le 1$ and $0 \le \theta \le 2\pi$. This parametrization is orientation preserving, because

This parametrization uses de Moivre's formula:

$$z^k = r^k(\cos k\theta + i \sin k\theta).$$

$$\operatorname{sgn} dx_1 \wedge dy_1 \Bigg(\underbrace{\begin{bmatrix} \cos\theta \\ \sin\theta \\ kr^{k-1}\cos k\theta \\ kr^{k-1}\sin k\theta \end{bmatrix}}_{\overrightarrow{D_1}\gamma}, \underbrace{\begin{bmatrix} -r\sin\theta \\ r\cos\theta \\ -r^{k+1}\sin k\theta \\ r^{k+1}\cos k\theta \end{bmatrix}}_{\overrightarrow{D_2}\gamma} \Bigg) = \operatorname{sgn} r(\cos^2\theta + \sin^2\theta) = +1.$$

The integral becomes

$$\begin{aligned} &\int_0^{2\pi}\int_0^1 (dx_1 \wedge dy_1 + dy_1 \wedge dx_2)\left(\begin{bmatrix} \cos\theta \\ \sin\theta \\ kr^{k-1}\cos k\theta \\ kr^{k-1}\sin k\theta \end{bmatrix}, \begin{bmatrix} -r\sin\theta \\ r\cos\theta \\ -kr^k\sin k\theta \\ kr^k\cos k\theta \end{bmatrix} \right) dr\ d\theta \\ &= \int_0^{2\pi}\int_0^1 \left(\det \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} + \det \begin{bmatrix} \sin\theta & r\cos\theta \\ kr^{k-1}\cos k\theta & -kr^k\sin k\theta \end{bmatrix} \right) \\ &= \int_0^{2\pi}\int_0^1 r - kr^k\Big(\sin\theta\sin k\theta + \cos\theta\cos k\theta\Big) \\ &= \int_0^{2\pi}\int_0^1 \big(r - kr^k\cos(k-1)\theta\big)\, dr\, d\theta = \begin{cases} \pi & \text{if } k \ne 1 \\ 0 & \text{if } k = 1 \end{cases}. \end{aligned}$$

6.4.7 a. The space $M_1(2,2)$ contains all 2×2 matrices with determinant 0 except the zero matrix; i.e., it is the complement of the zero matrix in the locus defined by the equation $\det(A) = 0$. It separates 2×2 matrices with positive determinant from those with negative determinant. At all points A of $M_1(2,2)$, we have $[\mathbf{D}\det(A)] \ne 0$ (by Exercise 3.1.19), so $M_1(2,2)$ is a manifold and does have a tangent space. Therefore we can choose a vector field orthogonal to the tangent space, pointing towards matrices with positive determinant. This vector field orients the manifold. The normal vector field pointing towards matrices with negative determinant gives the opposite orientation.

Solution 6.4.7: You might think of $M_1(2,2)$ as a "double" cone like the one in Figure 5.2.1, the vertex being the zero matrix.

The derivative $[\mathbf{D}\det(A)]$ is a linear transformation from the space of 2×2 matrices to $\mathbb{R}$. Writing $[\mathbf{D}\det(A)] \ne 0$ suggests that it is a scalar, which it is not, but writing $[\mathbf{D}\det(A)] \ne [0]$ would suggest that it is a matrix, and it is not that either.

b. This is rather similar to Example 6.4.9, though the computation comes out differently. We will parametrize three open subsets

$$U_1, U_2, U_3 \subset M_1(3,3),$$

whose union is all of $M_1(3,3)$. We will see that the orientations induced by these parametrizations are all compatible. The subsets U_i are the subsets where the ith column is not $\mathbf{0}$, and the parametrizations are

$$\underbrace{\gamma_1\begin{pmatrix} a_1\\ b_1\\ c_1\\ u_1\\ v_1\end{pmatrix} = \begin{bmatrix} a_1 & u_1a_1 & v_1a_1\\ b_1 & u_1b_1 & v_1b_1\\ c_1 & u_1c_1 & v_1c_1\end{bmatrix}}_{U_1}, \qquad \underbrace{\gamma_2\begin{pmatrix} a_2\\ b_2\\ c_2\\ u_2\\ v_2\end{pmatrix} = \begin{bmatrix} v_2a_2 & a_2 & u_2a_2\\ v_2b_2 & b_2 & u_2b_2\\ v_2c_2 & c_2 & u_2c_2\end{bmatrix}}_{U_2},$$

$$\underbrace{\gamma_3\begin{pmatrix} a_3\\ b_3\\ c_3\\ u_3\\ v_3\end{pmatrix} = \begin{bmatrix} u_3a_3 & v_3a_3 & a_3\\ u_3b_3 & v_3b_3 & b_3\\ u_3c_3 & v_3c_3 & c_3\end{bmatrix}}_{U_3}$$

On $U_1 \cap U_2$, the parametrizing variables are related by

$$a_2 = u_1a_1, \quad b_2 = u_1b_1, \quad c_2 = u_1c_1, \quad u_2 = \frac{v_1}{u_1}, \quad v_2 = \frac{1}{u_1}. \tag{1}$$

Equations 1: For example, we get $a_2 = u_1a_1$ by comparing the first row, second entry in the matrices for U_1 and U_2. Then substituting u_1a_1 for a_2 in the first row, third entry of the second matrix and comparing it with the corresponding entry of the first matrix gives $u_2u_1 = v_1$.

The mapping $\gamma_2^{-1}\circ\gamma_1$ is given by

$$\gamma_2^{-1}\circ\gamma_1\begin{pmatrix} a_1\\ b_1\\ c_1\\ u_1\\ v_1\end{pmatrix} = \begin{pmatrix} u_1a_1\\ u_1b_1\\ u_1c_1\\ v_1/u_1\\ 1/u_1\end{pmatrix}, \text{ with derivative } \begin{bmatrix} u_1 & 0 & 0 & a_1 & 0\\ 0 & u_1 & 0 & b_1 & 0\\ 0 & 0 & u_1 & c_1 & 0\\ 0 & 0 & 0 & -\frac{v_1}{u_1^2} & \frac{1}{u_1}\\ 0 & 0 & 0 & -\frac{1}{u_1^2} & 0\end{bmatrix}.$$

The determinant of this matrix is $u_1^3/u_1^3 = 1$. Thus, by Proposition 6.4.8, γ_1 and γ_2 define orientations on U_1 and U_2 that are compatible on $U_1 \cap U_2$.

Exactly the same computation shows that γ_1 and γ_3 define compatible orientations on $U_1 \cap U_3$, so together they define an orientation of $M_1(3,3)$.

6.5.1 The work form (a) is identical to (j) and (l). The work form evaluated on $P_{\mathbf{x}}(\vec{\mathbf{v}})$ in (b) is identical to (i).

The flux form in (k) is identical to (d) and (h). The flux form evaluated on $P_{\mathbf{x}}(\vec{\mathbf{v}}, \vec{\mathbf{w}})$, given in (c), is identical to (e) and (f).

The mass form in (g) has no equivalents.

6.5.3 a. This is correct as stated, if $\vec{F}$ is a vector field on $\mathbb{R}^4$. If it is a vector field on $\mathbb{R}^3$, one can either change the flux form to a mass form, getting $M_f\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)\big)$, or change the 3-parallelogram to a 2-parallelogram, getting $\Phi_{\vec{F}}\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)\big)$.

b. The work form field is associated to a vector field, not a function; this should be $W_{\vec{F}}$.

c. If f is a function on $\mathbb{R}^2$, this expression makes sense, as a mass form on $\mathbb{R}^2$. If f is a function on $\mathbb{R}^3$, the mass form field is a function of a 3-parallelogram; this should be $M_f\big(P_{\mathbf{x}}(\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}})\big)$ (or $M_f\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)\big)$ or the equivalent).

d. The cross product takes two vectors and gives a vector, while the dot product takes two vectors and gives a number. This should be $\vec{\mathbf{v}}_1 \cdot (\vec{\mathbf{v}}_2 \times \vec{\mathbf{v}}_3)$.

e. This is correct; the flux form Φ is associated to a vector field $\vec{F}$.

f. This expression should be $\Phi_{\vec{F}} = F_1\, dy \wedge dz - F_2\, dx \wedge dz + F_3\, dx \wedge dy$.

g. $W_{\vec{F}}\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1)\Big)$ or $\Phi_{\vec{F}}\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)\Big)$.

h. The mass form is associated to a function, not a vector field; this should be M_f.

i. This is correct as is.

6.5.5 You do not want to evaluate $W_{\vec{F}} \wedge \Phi_{\vec{G}}$ on vectors, as we did in Example 6.1.14; that leads to horrendous computations. When carrying out the computation, remember to discard terms containing $dx_i \wedge dx_i$ (for example, $F_1 G_2\, dx \wedge dz \wedge dx$), since such terms equal 0.

We will first show that $W_{\vec{F}} \wedge \Phi_{\vec{G}} = M_{\vec{F}\cdot\vec{G}}$. By Proposition 6.1.15,

$$\begin{aligned}
&\underbrace{(F_1 dx + F_2 dy + F_3 dz)}_{W_{\vec{F}}} \wedge \underbrace{(G_1 dy \wedge dz + G_2 dz \wedge dx + G_x dz \wedge dy)}_{\Phi_{\vec{G}}}\\
&\qquad = F_1 G_1\, dx \wedge dy \wedge dz + F_2 G_2\, dy \wedge dz \wedge dx + F_3 G_3\, dz \wedge dx \wedge dy\\
&\qquad = (F_1 G_1 + F_2 G_2 + F_3 G_3)\, dx \wedge dy \wedge dz\\
&\qquad = \vec{F} \cdot \vec{G}\, dx \wedge dy \wedge z = M_{\vec{F}\cdot\vec{G}}.
\end{aligned}$$

It follows that $M_{\vec{F}\cdot\vec{G}} = M_{\vec{G}\cdot\vec{F}} = W_{\vec{G}} \wedge \Phi_{\vec{F}}$.

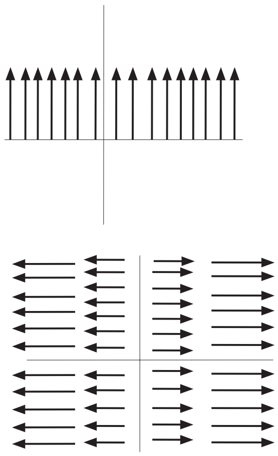


FIGURE FOR SOLUTION 6.5.7

TOP: The vector field $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ of part a.

BOTTOM: The vector field $\begin{bmatrix} x \\ 0 \end{bmatrix}$ of part b.

6.5.7 a. $dy = W_{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}$. The vector field $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is constant, pointing up; the work is large on a vertical path (positive if the path points up, negative if it points down); it is 0 on a horizontal path.

b. $x\, dx = W_{\begin{bmatrix} x \\ 0 \end{bmatrix}}$. The vector field $\begin{bmatrix} x \\ 0 \end{bmatrix}$ is horizontal, pointing to the right in the right halfplane, and to the left in the left halfplane. The work on a horizontal path from $\begin{pmatrix} a \\ 0 \end{pmatrix}$ to $\begin{pmatrix} b \\ 0 \end{pmatrix}$, with $0 < a < b$, is large. The work on a vertical path is 0, as is the work on a horizontal path from $\begin{pmatrix} -a \\ 0 \end{pmatrix}$ to $\begin{pmatrix} a \\ 0 \end{pmatrix}$.

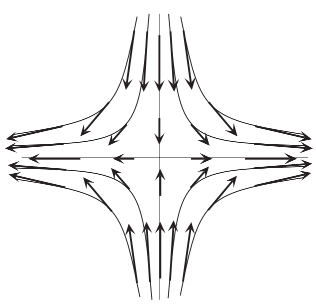


FIGURE FOR SOLUTION 6.5.7

The vector field $\begin{bmatrix} x \\ -y \end{bmatrix}$ of part c of Solution 6.5.7.

c. $x\, dx - y\, dy = W_{\begin{bmatrix} x \\ -y \end{bmatrix}}$. The vector field $\begin{bmatrix} x \\ -y \end{bmatrix}$ is tangent to the hyperbolas $xy = c$. The work of a path along one of the hyperbolas is large; the work along a segment of the lines $x = y$, $x = -y$ is 0.

6.5.9 a. One such parallelogram is anchored at $\begin{pmatrix}1\\1\\0\end{pmatrix}$ and spanned by $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}$, in that order:

$$\Big(y\,dy\wedge dz + x\,dx\wedge dz - z\,dx\wedge dy\Big)\left(P_{\begin{pmatrix}1\\1\\0\end{pmatrix}}\left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right)\right) = \Big(dy\wedge dz + dx\wedge dz\Big)\left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right)$$

$$= \det\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} + \det\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} = 2.$$

This means that the flux of the vector field $\vec{F}\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{bmatrix}y\\-x\\-z\end{bmatrix}$ through the parallelogram is 2.

One way to do this problem is to choose two vectors at random and do the computation. Almost any two vectors will give either positive or negative flux; if you get negative flux, just switch the order in which the two vectors are listed.

A more algebraic approach is to note that since the parallelogram is anchored at $\begin{pmatrix}1\\1\\0\end{pmatrix}$, the 2-form $y\,dy\wedge dz + x\,dx\wedge dz - z\,dx\wedge dy$ becomes $dy\wedge dz + dx\wedge dz$; construct your vectors $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$ so that

$$\det\begin{bmatrix}v_2 & w_2\\v_3 & w_3\end{bmatrix} + \det\begin{bmatrix}v_1 & w_1\\v_3 & w_3\end{bmatrix} > 0.$$

A more geometric approach is to note that Figure 6.5.4 suggests that the flux through any parallelogram lying in the vertical plane where $x = y$ will be nonzero. How can one guess, without doing the computation, whether the flux will be positive or negative? Remember (Proposition 1.4.21) that $\det[\vec{\mathbf{a}}, \vec{\mathbf{b}}, \vec{\mathbf{c}}]$ is positive if the vectors, in that order, satisfy the right-hand rule. So (Definition 6.5.2), the flux will be positive if $\vec{F}(\mathbf{x}), \vec{\mathbf{v}}, \vec{\mathbf{w}}$ satisfy the right-hand rule.

b. If we anchor the parallelogram at $\begin{pmatrix}0\\0\\-1\end{pmatrix}$, Φ gives a positive number (positive flux):

$$y\,dy\wedge dz + x\,dx\wedge dz - z\,dx\wedge dy\left(P_{\begin{pmatrix}0\\0\\-1\end{pmatrix}}\left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right)\right) = dx\wedge dy\left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right) = 1$$

If we anchor it at $\begin{pmatrix}0\\0\\1\end{pmatrix}$, Φ gives negative flux:

$$y\,dy\wedge dz + x\,dx\wedge dz - z\,dx\wedge dy\left(P_{\begin{pmatrix}0\\0\\1\end{pmatrix}}\left(\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix}\right)\right) = -dx\wedge dy\left(\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = -1.$$

Of course these aren't the only possibilities. How did we hit on these? The parallelogram is parallel to the (x, y)-plane, so Figure 6.5.4 and the right-hand rule say that anchoring it "below" that plane will give positive flux and anchoring it "above" that plane will give negative flux.

6.5.11 a. $W_{\begin{bmatrix}x\\y\end{bmatrix}}\left(P_{\begin{pmatrix}1\\1\end{pmatrix}}\begin{bmatrix}2\\3\end{bmatrix}\right) = \begin{bmatrix}1\\1\end{bmatrix}\cdot\begin{bmatrix}2\\3\end{bmatrix} = 5.$

b. $W_{\begin{bmatrix}x^2\\ \sin xy\end{bmatrix}}\left(P_{\begin{pmatrix}-1\\-\pi\end{pmatrix}}\begin{bmatrix}e\\ \pi\end{bmatrix}\right) = \begin{bmatrix}1\\0\end{bmatrix}\cdot\begin{bmatrix}e\\ \pi\end{bmatrix} = e.$

c. $\begin{bmatrix}0\\1\\1\end{bmatrix}\cdot\begin{bmatrix}2\\3\\-1\end{bmatrix} = 2.$ d. $\begin{bmatrix}\sin 1\\ \cos(-1)\\ e^0\end{bmatrix}\cdot\begin{bmatrix}0\\1\\0\end{bmatrix} = \cos 1.$

6.5.13 We need to show that the formula

$$\Phi_{\vec{F}(\mathbf{x})}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}) = \det[\vec{F}(\mathbf{x}),\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}]$$

is indeed multilinear and antisymmetric as a function of the $\vec{\mathbf{v}}_i$. This follows immediately from the corresponding statement about the determinant: it is linear as a function of each column, in particular as a function of the $\vec{\mathbf{v}}_i$, and it is antisymmetric as a function of its columns; in particular, if you exchange two of the $\vec{\mathbf{v}}_i$ you change the sign.

Another solution is to note that $\Phi_{\vec{F}(\mathbf{x})}$, written in coordinates, becomes equation 6.5.25 in the text: a linear combination of elementary $(n-1)$-forms on $\mathbb{R}^n$.

6.5.15 We have $\vec{F}(\mathbf{x}) = \begin{bmatrix}1\\0\\-1\end{bmatrix}$, and $f(\mathbf{x}) = -2$, so

$$W_{\vec{F}}\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1)\big) = \begin{bmatrix}1\\0\\-1\end{bmatrix}\cdot\begin{bmatrix}0\\1\\1\end{bmatrix} = -1$$

$$\Phi_{\vec{F}}\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2)\big) = \begin{bmatrix}1\\0\\-1\end{bmatrix}\cdot\left(\begin{bmatrix}0\\1\\1\end{bmatrix}\times\begin{bmatrix}1\\1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\0\\-1\end{bmatrix}\cdot\begin{bmatrix}-1\\1\\-1\end{bmatrix} = 0,$$

$$M_f\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2,\vec{\mathbf{v}}_3)\big) = -2\det\begin{bmatrix}0&1&-1\\1&1&1\\1&0&1\end{bmatrix} = -2\big(-1(1)+1(2)\big) = -2.$$

6.5.17 We will need to compute four integrals. The first side of the rectangle, denoted C_1, is parametrized by $\gamma_1(t) = \begin{pmatrix} 0 \\ t \end{pmatrix}$, for $0 \le t \le a$. This leads to

$$\int_{C_1} W_{\vec{F}} = \int_0^a \vec{F}\big(\gamma_1(t)\big) \cdot \gamma_1'(t)\, dt = \int_0^a \begin{bmatrix} 0 \\ t \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} dt = \frac{a^2}{2}.$$

Solution 6.5.17: These computations use equation 6.5.14.

Since $\vec{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} xy \\ ye^x \end{bmatrix}$, we have

$$\vec{F}\big(\gamma_1(t)\big) = \begin{bmatrix} 0 \\ t \end{bmatrix}.$$

The second side, C_2, is parametrized by $\gamma_2(t) = \begin{pmatrix} t \\ a \end{pmatrix}$, $0 \le t \le b$. This leads to

$$\int_0^b \begin{bmatrix} ta \\ ae^t \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} dt = \frac{ab^2}{2}.$$

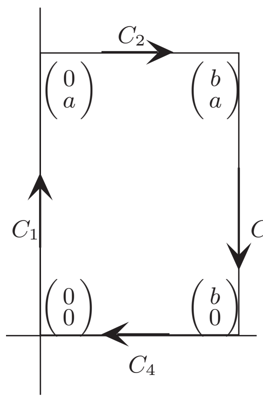


Figure for Solution 6.5.17.

The third side, C_3, is parametrized by $\gamma_3(t) = \begin{pmatrix} b \\ t \end{pmatrix}$, $0 \le t \le a$. But this parametrization reverses the orientation of C_3, so we find the integral

$$\int_a^0 \begin{bmatrix} bt \\ te^b \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} dt = -\frac{a^2 e^b}{2}.$$

Finally, the last side C_4 is parametrized by $\gamma_4(t) = \begin{pmatrix} t \\ 0 \end{pmatrix}$, $0 \le t \le b$. (Again, this parametrization reverses orientation.) This gives the integral

$$\int_b^0 \begin{bmatrix} 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} dt = 0.$$

So the total integral is

$$\frac{a^2}{2} + \frac{ab^2}{2} - \frac{a^2 e^b}{2}.$$

6.5.19 The obvious way to parametrize the sphere is with the spherical coordinates map

$$\gamma : \begin{pmatrix} \theta \\ \varphi \end{pmatrix} = \begin{pmatrix} R\cos\varphi\cos\theta \\ R\cos\varphi\sin\theta \\ R\sin\varphi \end{pmatrix}, \quad 0 \le \theta < 2\pi,\ -\frac{\pi}{2} < \varphi < \frac{\pi}{2}.$$

This is compatible with the orientation given by an outward-pointing vector field, since

$$\det\Bigg(\underbrace{\begin{bmatrix} R\cos\varphi\cos\theta \\ R\cos\varphi\sin\theta \\ R\sin\varphi \end{bmatrix}}_{\text{outward-pointing vector}}, \underbrace{\begin{bmatrix} -R\cos\varphi\sin\theta \\ R\cos\varphi\cos\theta \\ 0 \end{bmatrix}}_{\overrightarrow{D_\theta}\gamma\begin{pmatrix} \theta \\ \varphi \end{pmatrix}}, \underbrace{\begin{bmatrix} -R\sin\varphi\cos\theta \\ -R\sin\varphi\sin\theta \\ R\cos\varphi \end{bmatrix}}_{\overrightarrow{D_\varphi}\gamma\begin{pmatrix} \theta \\ \varphi \end{pmatrix}} \Bigg) = R^2\cos\varphi,$$

and $\cos\varphi > 0$ for $|\varphi| < \pi/2$. Thus (note that on the sphere, $r = R$) the flux becomes

$$\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} \underbrace{\begin{bmatrix} R^{a+1}\cos\varphi\cos\theta \\ R^{a+1}\cos\varphi\sin\theta \\ R^{a+1}\sin\varphi \end{bmatrix}}_{\vec{F}\gamma\begin{pmatrix}\theta\\ \varphi\end{pmatrix}} \cdot \left(\overrightarrow{D_\theta\gamma}\begin{pmatrix}\theta\\ \varphi\end{pmatrix} \times \overrightarrow{D_\varphi\gamma}\begin{pmatrix}\theta\\ \varphi\end{pmatrix}\right) d\varphi\, d\theta$$

$$= \int_0^{2\pi}\int_{-\pi/2}^{\pi/2} R^{a+1}\begin{bmatrix}\cos\varphi\cos\theta \\ \cos\varphi\sin\theta \\ \sin\varphi\end{bmatrix} \cdot \begin{bmatrix} R^2\cos^2\varphi\cos\theta \\ R^2\cos^2\varphi\sin\theta \\ R^2\cos\varphi\sin\varphi\end{bmatrix} d\varphi\, d\theta$$

$$= 2\pi R^{a+3}\int_{-\pi/2}^{\pi/2}\cos\varphi\, d\varphi = 4\pi R^{a+3}.$$

6.5.21 a. No, it does not preserve orientation, since

Here we use Definition 6.4.2.

$$\det\left[\mathbf{D}\gamma\begin{pmatrix}u\\ v\\ w\end{pmatrix}\right] = \det\begin{bmatrix}\cos v\cos w & -u\cos w\sin v & -R\sin w - u\cos v\sin w \\ \cos v\sin w & -u\sin w\sin v & R\cos w + u\cos v\cos w \\ \sin v & u\cos v & 0\end{bmatrix}$$
$$= -Ru - u^2\cos v.$$

This quantity is never positive (in fact, it is strictly negative except when $u = 0$, which happens on the core circle of the torus): by definition $R > 0$ and $u \geq 0$ and

$$Ru > |u^2\cos v|, \quad \text{since} \quad -u \leq u\cos v \leq u \text{ and } u \leq r < R,$$

so that even when v is between $\pi/2$ and $3\pi/2$ (so that $\cos v$ is negative), if $u > 0$, then $-Ru - u^2\cos v < 0$.

b. The integral becomes

$$\int_0^{2\pi}\int_0^{2\pi}\int_0^r f\left(\gamma\begin{pmatrix}u\\ v\\ w\end{pmatrix}\right)\det\left[\mathbf{D}\gamma\begin{pmatrix}u\\ v\\ w\end{pmatrix}\right] du\, dv\, dw$$

To compute $f(\gamma(\mathbf{u}))$, you don't need to multiply out in detail; just note that this becomes

$$(R + u\cos v)^2(\cos^2 w + \sin^2 w).$$

$$= \int_0^{2\pi}\int_0^{2\pi}\int_0^r -\overbrace{(R + u\cos v)^2}^{f(\gamma(\mathbf{u}))}\overbrace{u(R + u\cos v)}^{-\det[\mathbf{D}\gamma]}\, du\, dv\, dw$$

$$= 2\pi\int_0^{2\pi}\int_0^r (-R^3u - 3R^2u^2\cos v - 3Ru^3\cos^2 v - u^4\cos^3 v)\, du\, dv$$

$$= -2\pi\int_0^{2\pi}\left(\frac{R^3r^2}{2} + R^2r^3\cos v + \frac{3Rr^4\cos^2 v}{4} + \frac{r^5\cos^3 v}{5}\right) dv$$

$$= -\pi^2\left(2R^3r^2 + \frac{3Rr^4}{2}\right)$$

Since the parametrization reverses orientation, we should multiply that integral by -1.

6.6.1 a. Let X be the subset of $\mathbb{R}^3$ where $xyz \leq 1$ and $x^2 + y^2 + z^2 \leq 4$. This set is sort of like an apple with four shallow bites taken out. The sets $Z_1,\ Z_2$ of equation

$$g_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = xyz = 1 \quad \text{and} \quad g_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2 = 4$$

are both smooth 2-manifolds, since their derivatives do not vanish on Z_1 and Z_2 respectively. Therefore the points where both the equality $xyz = 1$ and the inequality $x^2 + y^2 + z^2 < 4$ are satisfied are smooth points of the boundary, as are the points where $xyz < 1$ and $x^2 + y^2 + z^2 - 4 = 0$.

Any nonsmooth points are necessarily points where both equalities

$$xyz = 1 \quad \text{and} \quad x^2 + y^2 + z^2 = 4$$

are satisfied. These points form a smooth curve C (actually, a disjoint union of four curves), since the derivative

$$\left[\mathbf{D} \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = \begin{bmatrix} 2x & 2y & 2z \\ yz & xz & xy \end{bmatrix}$$

has rank 2 at all points of C. So C is a manifold of dimension 1 and by Proposition 5.2.2 has 2-dimensional volume 0, satisfying part 1 of Definition 6.6.8.

Solution 6.6.1, part a: You may wonder how we know the derivative has rank 2. Certainly it must be either 1 or 2. It can't be 0, since x, y and z are nonzero; it can't be greater than 2, because the codomain is $\mathbb{R}^2$. Setting the determinant of the first two columns to 0 gives $2x^2z = 2y^2z$, so the first two columns can be linearly dependent only if $x = \pm y$ or if $z = 0$ (which is impossible, since $xyz = 1$).

Doing the same with the last two columns gives $2xy^2 = 2xz^2$, so those columns can be linearly dependent only if $y = \pm z$. Together these say that to satisfy $xyz = 1$, we must have $|x| = |y| = |z| = 1$, which contradicts $x^2+y^2+z^2 = 4$. Therefore, two of the columns are linearly independent.

Part 2 is clear: the smooth part of the boundary is the union of the part of the sphere of radius 2 where $xyz < 1$, which certainly has finite area, and the part of the smooth surface of equation $xyz = 1$ where $x^2+y^2+z^2 < 4$, which is a 2-dimensional piece-with-boundary, and has finite area by Theorem 6.6.16.

b. In this case we have sliced the bitten apple in two, and kept one of the pieces. (The pieces are not equal: one contains an entire bite and small pieces of the other three; the other has most of the other three. We kept the first.) Define Z_3 to be the plane of equation

$$g_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + y + z = 0.$$

We can repeat the above argument to show that any points where exactly one of the inequalities

$$g_1 \leq 1, \quad g_2 \leq 4, \quad g_3 \geq 0$$

is satisfied as an equality are smooth points of the boundary (either the peel of the apple, or exposed by the bite, or on the slice). Thus the nonsmooth points of the boundary are those where at least two of the inequalities are satisfied as equalities.

Any pair of equations defines a smooth curve, of two-dimensional volume 0, so the set of nonsmooth points, being a subset of a union of three such curves, has 2-dimensional volume 0. Moreover, as before, the projections of these curves onto the (x, y)-plane have area zero, so can be covered by dyadic squares of arbitrarily small area, and the parts of each of the surfaces Z_1, Z_2, Z_3 above these squares still have arbitrarily small area.

6.6.3 a. The only thing to check is that Ω is well defined, i.e., that at every $\mathbf{x} \in X$, we can find vectors $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1} \in T_{\mathbf{x}}X$ such that

$$\det\left[\vec{\nabla} f(\mathbf{x}), \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}\right] \neq 0.$$

That is the same as requiring that the vectors $\vec{\nabla} f(\mathbf{x}), \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}$ be linearly independent, and it is enough to take $\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}$ to be a basis of $T_{\mathbf{x}}X$, since $\vec{\nabla} f(\mathbf{x})$ is orthogonal to $T_{\mathbf{x}}X$.

b. Let X be the boundary of the region $Y = \{\mathbf{x} \mid f(\mathbf{x}) \leq 0\}$. We will give Y the standard orientation of $\mathbb{R}^n$, by det. Then Ω defines the boundary orientation of $\partial Y = X$. Indeed, $\vec{\nabla} f(\mathbf{x})$ is an outward-pointing vector. By Definition 6.6.21, the boundary orientation of an oriented piece $X \subset M$ at a point $\mathbf{x} \in \partial X$ is defined by

$$\Omega^{\partial}_M(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}) = \Omega_M(\vec{\mathbf{v}}_{\text{out}}, \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}),$$

where Ω_M is the orientation of the ambient manifold M. In our case, $\Omega_M = \Omega^{st}$, the standard orientation of $\mathbb{R}^n$, so the formula above reads

$$\Omega^{\partial}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}) = \operatorname{sgn}\det\left[\vec{\mathbf{v}}_{\text{out}}, \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}\right],$$

and the result follows since $\vec{\nabla} f(\mathbf{x})$ is an outward-pointing vector.

Solution 6.6.3: If we had considered X as the boundary of the set

$$\{\mathbf{x} \mid f(\mathbf{x}) \geq 0\},$$

then the orientation by Ω would be the opposite of the boundary orientation.

6.6.5 a. All vectors in $T_{\mathbf{x}}P$ are of the form $\begin{bmatrix} a \\ b \\ -a-b \end{bmatrix}$ for some a and b in $\mathbb{R}$. The 2-form defined by the normal $\vec{N} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ takes two vectors in $T_{\mathbf{x}}P$, $\vec{\mathbf{v}} = \begin{bmatrix} v_1 \\ v_2 \\ -v_1 - v_2 \end{bmatrix}$ and $\vec{\mathbf{w}} = \begin{bmatrix} w_1 \\ w_2 \\ -w_1 - w_2 \end{bmatrix}$, and returns the number

$$\det\left[\vec{N}, \vec{\mathbf{v}}, \vec{\mathbf{w}}\right] = \det \begin{bmatrix} 1 & v_1 & w_1 \\ 1 & v_2 & w_2 \\ 1 & -v_1 - v_2 & -w_1 - w_2 \end{bmatrix} = 3(v_1 w_2 - v_2 w_1).$$

Compare the action of the listed forms on $\vec{\mathbf{v}}$ and $\vec{\mathbf{w}}$:

$$\begin{aligned} dx \wedge dy(\vec{\mathbf{v}}, \vec{\mathbf{w}}) &= \quad v_1 w_2 - v_2 w_1 \\ dx \wedge dz(\vec{\mathbf{v}}, \vec{\mathbf{w}}) &= -(v_1 w_2 - v_2 w_1) \\ dy \wedge dz(\vec{\mathbf{v}}, \vec{\mathbf{w}}) &= \quad v_1 w_2 - v_2 w_1. \end{aligned}$$

The first and third forms, $dx \wedge dy$ and $dy \wedge dz$, equal the form defined by the normal, multiplied by the positive real number $1/3$, so they define the same orientation of P. The second is minus the form defined by the normal, so it defines the opposite orientation.

b. We will see that X satisfies the conditions of Definition 6.6.8, and that its nonsmooth boundary is empty, i.e., that all points in the boundary are smooth points (satisfy Definition 6.6.2).

Indeed, we can take

$$g_1\begin{pmatrix} x\\ y\\ z\end{pmatrix} = x+y+z \quad \text{and} \quad g_2\begin{pmatrix} x\\ y\\ z\end{pmatrix} = (x^2+y^2+z^2)-1.$$

Then

$$\left[\mathbf{D}\begin{pmatrix} g_1\\ g_2\end{pmatrix}\begin{pmatrix} x\\ y\\ z\end{pmatrix}\right] = \begin{bmatrix} 1 & 1 & 1\\ 2x & 2y & 2z\end{bmatrix},$$

which has rank two (hence is surjective) unless $x=y=z$. In P, this can only happen at the origin, which is not in $\partial_P X$.

It is easy to check that our map takes its values in $\partial_P X$. It is also easy to see that it has injective derivative, and that it is injective on $(0,2\pi)$. So the only problem is to see that it is surjective (with domain $[0,2\pi]$). Given two coordinates $\begin{pmatrix} a\\ b\end{pmatrix}$, we can solve

$$\begin{aligned} u-v&=a\\ -u-v&=b, \end{aligned} \qquad \text{to find} \qquad \begin{aligned} u&=\frac{a-b}{2}\\ v&=-\frac{a+b}{2} \end{aligned}.$$

If

$$2a^2+2b^2+2ab = \frac{3}{2}(a+b)^2+\frac{1}{2}(a-b)^2 = 1,$$

we find $|u|\le\sqrt{2}/2$ and $|v|\le\sqrt{6}/6$. Thus we can set

$$u=\frac{\cos t}{\sqrt{2}} \quad \text{and} \quad v=\frac{\sin s}{\sqrt{6}}$$

and since $2u^2+6v^2=1$, we can take $t=s$. This shows that the map is surjective.

c. The parametrization γ is consistent with the orientation of $\partial_P X$ if $\vec{\gamma}'(t)$ points in the same direction as the unit tangent vector field to $\partial_P X$ that defines the orientation of $\partial_P X$. Definition 6.6.21 and Example 6.6.24 tell us that this is equivalent to having

$$\det\left[\vec{N}, \vec{\mathbf{v}}_{\text{out}}(t), \vec{\gamma}'(t)\right] > 0$$

There is no right (or wrong) way to decide whether or not to put arrows on $\gamma(t)$; we are thinking of it both as a point and as a vector.

for $0\le t\le 2\pi$. Note that $\gamma(t)$ is in P and points out of X at $\mathbf{x}=\vec{\gamma}(t)$, so we can set $\vec{\mathbf{v}}_{\text{out}}(t)=\vec{\gamma}(t)$. From part a, $\det\left[\vec{N},\vec{\mathbf{v}},\vec{\mathbf{w}}\right] = 3\,dy\wedge dz(\vec{\mathbf{v}},\vec{\mathbf{w}})$ for two vectors $\vec{\mathbf{v}},\vec{\mathbf{w}}\in P$. Both $\vec{\gamma}(t)$ and $\vec{\gamma}'(t)$ are in P, so

$$\det\left[\vec{N},\vec{\gamma}(t),\vec{\gamma}'(t)\right] = 3\,dy\wedge dz\left[\vec{\gamma}(t),\vec{\gamma}'(t)\right].$$

It is easy to calculate the value of the latter expression:

$$3\,dy \wedge dz\,[\vec{\mathbf{v}}_{\text{out}}(t), \vec{\gamma}'(t)] = 3 \det \begin{bmatrix} -\frac{\cos t}{\sqrt{2}} - \frac{\sin t}{\sqrt{6}} & \frac{\sin t}{\sqrt{2}} - \frac{\cos t}{\sqrt{6}} \\ 2\,\frac{\sin t}{\sqrt{6}} & 2\,\frac{\cos t}{\sqrt{6}} \end{bmatrix} = -\sqrt{3}.$$

This is always negative, so the parametrization is not compatible with the boundary orientation of $\partial_P X$ and is, moreover, orientation reversing.

d. Since γ from part c is orientation reversing, a form compatible with the orientation of $\partial_P X$ will yield a negative real number when it acts on $\vec{\gamma}'(t)$ at $\mathbf{x} = \gamma(t)$. Consider the actions of the following 1-forms on $\vec{\gamma}'(t)$, at any position $\mathbf{x} = \gamma(t)$:

$$\begin{aligned} dx\,(\vec{\gamma}'(t)) &= -\frac{\sin t}{\sqrt{2}} - \frac{\cos t}{\sqrt{6}}, \\ dy\,(\vec{\gamma}'(t)) &= \ \ \frac{\sin t}{\sqrt{2}} - \frac{\cos t}{\sqrt{6}}, \\ dz\,(\vec{\gamma}'(t)) &= \ \ 2\,\frac{\cos t}{\sqrt{6}}. \end{aligned}$$

Since these are not always less than 0 for $0 \le t < 2\pi$, they cannot define the orientation at every point.

e. Use the same strategy as in part d. The actions of the listed forms on the vector $\vec{\gamma}'(t)$ anchored at $\mathbf{x} = \gamma(t)$ are as follows:

$$\begin{array}{lll} x\,dy - y\,dx & \text{yields} & -1/\sqrt{3}, \\ x\,dz - z\,dx & \text{yields} & \ \ 1/\sqrt{3}, \\ y\,dz - z\,dy & \text{yields} & -1/\sqrt{3}. \end{array}$$

The first and third forms, $x\,dy - y\,dx$ and $y\,dz - z\,dy$, are negative for all $0 \le t < 2\pi$, hence may be used to define the orientation of $\partial_P X$ at every point.

6.6.7 a. Remember that in $\mathbb{R}^n$ the length of the longest diagonal of a cube of sidelength s is $\sqrt{n}\,s$, so every point of the cube is within $s\sqrt{n}/2$ from its center. Thus any point of $\mathbb{R}^n$ is distance at most $\sqrt{n}/2^{N+1}$ from some point of $\frac{1}{2^N}\mathbb{Z}^n$ (the corners of the cubes of $\mathcal{D}_N(\mathbb{R}^n)$).

Let $\mathbf{x}$ be the center of C, and find such a point $\mathbf{y} \in \frac{1}{2^N}\mathbb{Z}^n$ distance at most $\sqrt{n}/2^{N+1}$ from $\mathbf{g}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$. Since A is orthogonal it preserves all lengths, so every point of $\mathbf{g}(C)$ is distance at most $\sqrt{n}/2^{N+1}$ from $\mathbf{g}(\mathbf{x})$, hence (by the triangle inequality) distance at most $\sqrt{n}/2^N$ from $\mathbf{y}$.

The ball of radius $\sqrt{n}/2^N$ around $\mathbf{y}$ is contained in the cube of sidelength $2(\sqrt{n}+1)/2^N$ which contains at most $(2(\sqrt{n}+1))^n$ cubes of $\mathcal{D}_N(\mathbb{R}^n)$.

b. This is almost the same: again, let $\mathbf{x}$ be the center of C, and find a point $\mathbf{y} \in \frac{1}{2^N}\mathbb{Z}^n$ distance at most $\sqrt{n}/2^{N+1}$ of $\mathbf{g}(\mathbf{x}) = A\mathbf{x} + \mathbf{c}$.

This time, every point of $\mathbf{g}(C)$ is within $|A|\sqrt{n}/2^{N+1}$ of $\mathbf{g}(\mathbf{x})$, so within $(|A|+1)\sqrt{n}/2^{N+1}$ from $|\mathbf{y}|$. Again to allow for fudging at the ends, $\mathbf{g}(C)$ is contained in the cube of sidelength $\big((|A|+1)\sqrt{n}+2\big)/2^N$, and as such contains at most $\big((|A|+1)\sqrt{n}+2\big)^n$ cubes of $\mathcal{D}_N(\mathbb{R}^n)$.

c. Note that we cannot write

$$\det(AB)^\top AB = \det\Big((B^\top A^\top A)B\Big) = \det B(B^\top A^\top A) = \det(BB^\top)\det(A^\top A).$$

and set $M = \det A^\top A$, because B is not square.

Use the SVD (Theorem 3.8.1) to write $B = PDQ^\top$, with P, Q orthogonal and D rectangular diagonal. Then (using the fact that the determinant of an orthogonal matrix is ± 1; see Exercise 4.8.23)

$$\det(AB)^\top AB = \det QD^\top PA^\top APDQ^\top = \det D^\top (AP)^\top (AP)D.$$

Write $\widetilde{D}$ for the diagonal $k \times k$ submatrix of D, and $\widetilde{AP}$ for the first k columns of AP. Then we have

$$D^\top (AP)^\top (AP)D = \widetilde{D}^\top (\widetilde{AP})^\top (\widetilde{AP})\widetilde{D},$$

so

$$\det\Big((AB)^\top AB\Big) = \det\Big(\widetilde{D}^\top (\widetilde{AP})^\top (\widetilde{AP})\widetilde{D}\Big) = (\det \widetilde{D})^2 \det\Big((\widetilde{AP})^\top (\widetilde{AP})\Big).$$

By Theorem 3.8.1 and Corollary 4.8.25, $\det \widetilde{D}$ is the product of the square roots of the eigenvalues of $B^\top B$, hence $(\det \widetilde{D})^2 = \det B^\top B$. It remains to show that $\det(\widetilde{AP})^\top(\widetilde{AP})$ is bounded by a constant that depends only on A; but P is an orthogonal matrix and the orthogonal group is compact (Exercise 3.2.11), so

$$M \overset{\text{def}}{=} \sup_{P \in O(n)} |\det\big((\widetilde{AP})^\top (\widetilde{AP})\big)|$$

is a number depending only on A.

6.6.9 The proof of Stokes theorem for manifolds with corners will essentially contain this exercise; most of the solution is in Proposition 6.10.7, which we repeat here.

Proposition 6.10.7. *Let $M \subset \mathbb{R}^n$ be a k-dimensional manifold, and let $Y \subset M$ be a piece-with-corners. Then every $\mathbf{x} \in Y$ is the center of a ball U in $\mathbb{R}^n$ such that there exists a diffeomorphism $\mathbf{F} : U \to \mathbb{R}^n$ satisfying*

$$\mathbf{F}(U \cap M) = \mathbf{F}(U) \cap \mathbb{R}^k \qquad 6.10.30$$

$$\mathbf{F}(U \cap Y) = \mathbf{F}(U) \cap \mathbb{R}^k \cap Z, \qquad 6.10.31$$

where Z is a region $x_1 \geq 0, \dots, x_j \geq 0$ for some $j \leq k$.

Proposition 6.10.7: The idea is to use the functions defining M and X within M as coordinate functions of a new space; this has the result of straightening out M and ∂X, since these become sets where appropriate coordinate functions vanish.

Proof. At any $\mathbf{x} \in Y$, Definition 6.6.5 of a corner point and Definition 3.1.10 defining a manifold known by equations give us a neighborhood $V \subset \mathbb{R}^n$ of $\mathbf{x}$, and a collection of C^1 functions $V \to \mathbb{R}$ with linearly independent derivatives: functions $f_1, \dots, f_{n-k}$ defining $M \cap V$ and functions $\widetilde{g}_1, \dots, \widetilde{g}_m$ (extensions of the coordinate functions $g_i, \dots, g_m$ of the map $\mathbf{g}$ of Definition 6.6.5). The number m may be 0; this will happen if $\mathbf{x}$ is in the interior of Y.

In each case we extend the collection of their derivatives to a maximal collection of n linearly independent linear functions, by adding linearly independent linear functions $\lambda_1, \dots, \lambda_{k-m} : \mathbb{R}^n \to \mathbb{R}$. These n functions define a map $\mathbf{F} : V \to \mathbb{R}^n$ whose derivative at $\mathbf{x}$ is invertible. We can apply the inverse function theorem to say that locally, $\mathbf{F}$ is invertible, so locally it is a diffeomorphism on some neighborhood $U \subset V$ of $\mathbf{x}$. Equations 6.10.30 and 6.10.31 then follow. $\square$

Using this, there are two things to show: the nonsmooth boundary has $(k-1)$- dimensional volume 0 and the smooth boundary has finite $(k-1)$-dimensonal volume. It is enough to show that this is true in U, or even in $U' \subset U$ containing $\mathbf{x}$ such that the closure $\overline{U}'$ is contained in U (and $\overline{U}'$ is compact of course since it is closed and bounded). Indeed, by the Heine-Borel theorem, since X is compact it is covered by finitely many such U', so both properties are true for all of X if they are true for each U'.

The nonsmooth boundary is a subset of a manifold of dimension $< k-1$ (see Proposition 5.2.2), since it is defined by at least two equations with surjective derivative. For the smooth boundary, consider the part where some g_i in U' it is a subset of a compact piece of a $k-1$-dimensional manifold, parametrized by F^{-1} restricted to a $(k-1)$-dimensional subspace of $\mathbb{R}^n$. So the $(k-1)$-dimensional volume of this part of the smooth boundary is an integral of a continuous function with compact support, hence finite.

6.6.11 Suppose $\left[\mathbf{D}\begin{pmatrix}\mathbf{f}\\ g\end{pmatrix}\left(\mathbf{x}\right)\right]$ is onto; then there exists $\vec{\mathbf{v}} \in \mathbb{R}^n$ such that $\left[\mathbf{D}\begin{pmatrix}\mathbf{f}\\ g\end{pmatrix}\left(\mathbf{x}\right)\right]\vec{\mathbf{v}} = \begin{pmatrix}\mathbf{0}\\ 1\end{pmatrix}$. The top entry says that $\vec{\mathbf{v}} \in T_{\mathbf{x}}M$, and the bottom entry says that $[\mathbf{D}g(\mathbf{x})]$ does not vanish on $\vec{\mathbf{v}}$, so $[\mathbf{D}g(\mathbf{x})] : T_{\mathbf{x}}M \to \mathbb{R}$ is onto.

For the converse, suppose that $[\mathbf{D}g(\mathbf{x})] : T_{\mathbf{x}}M \to \mathbb{R}$ is onto. Choose $\begin{bmatrix}\vec{\mathbf{w}}\\ w\end{bmatrix} \in \mathbb{R}^{n-k+1}$. We need to find $\vec{\mathbf{a}} \in \mathbb{R}^n$ such that

$$\left[\mathbf{D}\begin{pmatrix}\mathbf{f}\\ g\end{pmatrix}\ \left(\mathbf{x}\right)\right]\vec{\mathbf{a}} = \begin{bmatrix}\vec{\mathbf{w}}\\ w\end{bmatrix}.$$

Since $[\mathbf{Df}(\mathbf{x})]$ is onto we can find $\vec{\mathbf{v}}_1$ such that

$$[\mathbf{Df}(\mathbf{x})]\mathbf{v}_1 = \vec{\mathbf{w}},$$

and since $[\mathbf{D}g(\mathbf{x})] : T_{\mathbf{x}}M \to \mathbb{R}$ is onto we can find $\vec{\mathbf{v}}_2 \in T_{\mathbf{x}}M$ with $[\mathbf{D}g(\mathbf{x})]\vec{\mathbf{v}}_2 = w - [\mathbf{D}g(\mathbf{x})]\vec{\mathbf{v}}_1$. We will consider separately the top and bottom lines of $\left[\mathbf{D}\begin{pmatrix}\mathbf{f}\\ g\end{pmatrix}\ \left(\mathbf{x}\right)\right](\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2)$.

We have

$$[\mathbf{Df}(\mathbf{x})](\vec{\mathbf{v}}_2) = \mathbf{0}$$

since $\vec{\mathbf{v}}_2 \in T_{\mathbf{x}}M$ and, by Theorem 3.2.4,

$$T_{\mathbf{x}}M = \ker[\mathbf{Df}(\mathbf{x})].$$

The top line is

$$[\mathbf{Df}(\mathbf{x})](\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2) = [\mathbf{Df}(\mathbf{x})](\vec{\mathbf{v}}_1) + [\mathbf{Df}(\mathbf{x})](\vec{\mathbf{v}}_2) = \vec{\mathbf{w}} + \underbrace{[\mathbf{Df}(\mathbf{x})](\vec{\mathbf{v}}_2)}_{\mathbf{0}\text{ ; see margin}} = \vec{\mathbf{w}}$$

The bottom line is

$$[\mathbf{D}g(\mathbf{x})](\vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2) = [\mathbf{D}g(\mathbf{x})](\vec{\mathbf{v}}_1) + [\mathbf{D}g(\mathbf{x})](\vec{\mathbf{v}}_2) = [\mathbf{D}g(\mathbf{x})](\vec{\mathbf{v}}_1) + w - [\mathbf{D}g(\mathbf{x})](\vec{\mathbf{v}}_1) = w$$

Therefore if $\vec{\mathbf{a}} = \vec{\mathbf{v}}_1 + \vec{\mathbf{v}}_2$, we have

$$\begin{bmatrix} [\mathbf{D}\mathbf{f}(\mathbf{x})] \\ [\mathbf{D}g(\mathbf{x})] \end{bmatrix} \vec{\mathbf{a}} = \begin{bmatrix} \vec{\mathbf{w}} \\ w \end{bmatrix}, \quad \text{so} \quad \left[\mathbf{D}\begin{pmatrix} \mathbf{f} \\ g \end{pmatrix}(\mathbf{x})\right] : \mathbb{R}^n \to \mathbb{R}^{n-k+1} \quad \text{is onto.}$$

6.7.1 a.

$$\mathbf{d}(x_1x_3\, dx_3 \wedge dx_4) = \Big(D_1(x_1x_3)dx_1 + D_2(x_1x_3)dx_2 + D_3(x_1x_3)dx_3 + D_4(x_1x_3)dx_4\Big) \wedge dx_3 \wedge dx_4$$
$$= (x_3dx_1 + x_1dx_3) \wedge dx_3 \wedge dx_4 = x_3\, dx_1 \wedge dx_3 \wedge dx_4.$$

b. $\mathbf{d}(\cos xy\, dx \wedge dy) = (-y \sin xy\, dx - x \sin xy\, dy) \wedge dx \wedge dy = 0.$

6.7.3 When computing the exterior derivative, remember that any terms containing $dx_i \wedge dx_i$ are 0.

a.

$$\begin{aligned}
\mathbf{d}\Phi_{\vec{F}_2} &= \mathbf{d}\left(\frac{-y}{x^2+y^2}\, dx + \frac{x}{x^2+y^2} dy\right) = \mathbf{d}\left(\frac{-y}{x^2+y^2} dx\right) + \mathbf{d}\left(\frac{x}{x^2+y^2} dy\right) \\
&= \left(D_1 \frac{-y}{x^2+y^2} dx + D_2 \frac{-y}{x^2+y^2} dy\right) \wedge dx + \left(D_1 \frac{x}{x^2+y^2} dx + D_2 \frac{x}{x^2+y^2} dy\right) \wedge dy \\
&= \frac{y^2-x^2}{(x^2+y^2)^2} dy \wedge dx + \frac{(x^2+y^2)-2x^2}{(x^2+y^2)^2} dx \wedge dy = 0.
\end{aligned}$$

b. Set $\vec{F}_3 = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$. Then

$$\begin{aligned}
\mathbf{d}\Phi_{\vec{F}_3} &= \mathbf{d}\Big(F_1\, dy \wedge dz - F_2\, dx \wedge dz + F_3\, dx \wedge dy\Big) \\
&= \mathbf{d}\left(\frac{x}{(x^2+y^2+z^2)^{3/2}} dy \wedge dz - \frac{y}{(x^2+y^2+z^2)^{3/2}} dx \wedge dz + \frac{z}{(x^2+y^2+z^2)^{3/2}} dx \wedge dy\right) \\
&= \frac{(x^2+y^2+z^2)^{3/2} \cdot 1 - x\frac{3}{2}(x^2+y^2+z^2)^{1/2} \cdot 2x}{(x^2+y^2+z^2)^3}\, dx \wedge dy \wedge dz \\
&\quad - \frac{(x^2+y^2+z^2)^{3/2} - 3y^2(x^2+y^2+z^2)^{1/2}}{(x^2+y^2+z^2)^3}\, dy \wedge dx \wedge dz \\
&\quad + \frac{(x^2+y^2+z^2)^{3/2} - 3z^2(x^2+y^2+z^2)^{1/2}}{(x^2+y^2+z^2)^3}\, dz \wedge dz \wedge dy \\
&= \frac{3(x^2+y^2+z^2)^{3/2} - 3(x^2+y^2+z^2)^{1/2}(x^2+y^2+z^2)}{(x^2+y^2+z^2)^3} = 0.
\end{aligned}$$

6.7.5 a. Computing the exterior derivative from the definition means computing the integral on the right side of

$$\mathbf{d}(z^2\, dx \wedge dy)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)\big) = \lim_{h \to 0} \frac{1}{h^3} \int_{\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1, h\vec{\mathbf{v}}_2, h\vec{\mathbf{v}}_3)} z^2\, dx \wedge dy.$$

Thus we must integrate the 2-form field $z^2\, dx \wedge dy$ over the boundary of the 3-parallelogram spanned by $h\vec{\mathbf{v}}_1, h\vec{\mathbf{v}}_2, h\vec{\mathbf{v}}_3$, i.e., over the six faces of the 3-parallelogram shown in the margin on the next page.

Solution 6.7.5, part a: If you just integrate everything in sight, this is a long computation. To make it bearable you need to (1) not compute terms that don't need to be computed, since they will disappear in the limit, and (2) take advantage of cancellations early in the game. To interpret the answer, it also helps to know what one expects to find (more on that later).

We parametrize those six faces as follows, where $0 \le s,\ t \le h$, and $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$; we use Proposition 6.6.26 to determine which faces are taken with a plus sign and which are taken with a minus sign:

1. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + s\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_2$, minus. (This is the base of the box shown at left.)
2. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + h\vec{\mathbf{v}}_3 + s\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_2$, plus. (This is the top of the box.)
3. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + s\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_3$, plus. (This is the front of the box.)
4. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + h\vec{\mathbf{v}}_2 + s\vec{\mathbf{v}}_1 + t\vec{\mathbf{v}}_3$, minus. (This is the back of the box.)
5. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + s\vec{\mathbf{v}}_2 + t\vec{\mathbf{v}}_3$, minus. (This is the left side of the box.)
6. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + h\vec{\mathbf{v}}_1 + s\vec{\mathbf{v}}_2 + t\vec{\mathbf{v}}_3$, plus. (This is the right side of the box.)

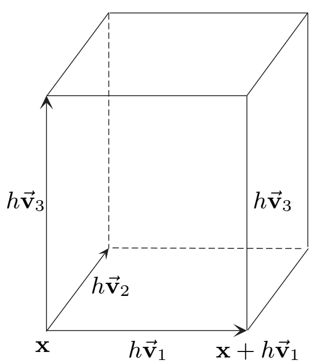


FIGURE FOR SOLUTION 6.7.5.

The 3-parallelogram spanned by $h\vec{\mathbf{v}}_1, h\vec{\mathbf{v}}_2, h\vec{\mathbf{v}}_3$.

Note that the two first faces are spanned by $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_2$ (or translates thereof), the base taken with $-$ and the top taken with $+$. The next two faces, with opposite signs, are spanned by $\vec{\mathbf{v}}_1$ and $\vec{\mathbf{v}}_3$, and the two sides, with opposite signs, are spanned by $\vec{\mathbf{v}}_2$ and $\vec{\mathbf{v}}_3$.

To integrate our form field over these parametrized domains we use Definition 6.2.1. The computations are tedious, but we do not have to integrate everything in sight. For one thing, common terms that cancel can be ignored; for another, anything that amounts to a term in h^4 can be ignored, since h^4/h^3 will vanish in the limit as $h \to 0$.

We will compute in detail the integrals over the first two faces. The integral over the first face (the base) is the following, where $v_{1,3}$ denotes the third entry of $\vec{\mathbf{v}}_1$ and $v_{2,3}$ denotes the third entry of $\vec{\mathbf{v}}_2$:

$$-\int_0^h \int_0^h \overbrace{(z + sv_{1,3} + tv_{2,3})^2 dx \wedge dy}^{z^2 dx\wedge dy \text{ eval. at } \gamma\begin{pmatrix} s \\ t \end{pmatrix}} \overbrace{(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)}^{\overrightarrow{D_s\gamma}, \overrightarrow{D_t\gamma}} \, ds\, dt$$

$$= -\int_0^h \int_0^h \Big(\underbrace{z^2}_{\text{term in } h^2} + \underbrace{s^2v_{1,3}^2 + t^2 v_{2,3}^2 + 2stv_{1,3}v_{2,3}}_{\text{terms in } h^4} + \underbrace{2szv_{1,3} + 2tzv_{2,3}}_{\text{terms in } h^3}\Big)(v_{1,1}v_{2,2} - v_{1,2}v_{2,1})\, ds\, dt$$

Note that the integral $\int_0^h \int_0^h z^2\, ds\, dt$ will give a term in h^2, and the next three terms will give a term in h^4; for example, $\int_0^h s^2 v_{1,3}^2\, ds$ gives an h^3 and $\int_0^h s^2 v_{1,3}^2\, dt$ gives an h, making h^4 in all. These higher degree terms can be disregarded; we will denote them below by $O(h^4)$. This gives the following integral over the first face:

The notation $O(h)$ is discussed in Appendix A.11.

$$-\int_0^h \int_0^h \big(z^2 + 2szv_{1,3} + 2tzv_{2,3} + O(h^4)\big)(v_{1,1}v_{2,2} - v_{1,2}v_{2,1})\, ds\, dt$$
$$= -\big(h^2 z^2 + h^3 zv_{1,3} + h^3 zv_{2,3} + O(h^4)\big)(v_{1,1}v_{2,2} - v_{1,2}v_{2,1}).$$

Before computing the integral over the second face (the top), notice that it is exactly like the first face except that it also has the term $h\vec{\mathbf{v}}_3$. This face comes with a plus sign, while the first face comes with a minus sign, so when we integrate over the second face, the identical terms cancel each other. Thus we didn't actually have to compute the integral over the first face at all! In computing the contribution of the first two faces to the integral, we need only concern ourselves with those terms in $(z + hv_{3,3} + sv_{1,3} + tv_{2,3})^2$ that contain $hv_{3,3}$: i.e., $h^2v_{3,3}^2$, $2hsv_{3,3}v_{1,3}$, $2htv_{3,3}v_{2,3}$, and $2zhv_{3,3}$. Integrating the first three would give terms in h^4, so the entire contribution to the integral of the first two faces is

$$\int_0^h \int_0^h \big(2zhv_{3,3} + O(h^4)\big)(v_{1,1}v_{2,2} - v_{1,2}v_{2,1})\, ds\, dt = (2zh^3v_{3,3} + O(h^4))(v_{1,1}v_{2,2} - v_{1,2}v_{2,1}).$$

Similarly, the entire contribution of the second pair of faces is

$$-\int_0^h \int_0^h \big(2zhv_{2,3} + O(h^4)\big)(v_{1,1}v_{3,2} - v_{3,1}v_{1,2})\, ds\, dt = -\Big(2zh^3v_{2,3} + O(h^4)\Big)\Big(v_{1,1}v_{3,2} - v_{3,1}v_{1,2}\Big).$$

(The partial derivatives are different, so we have $(v_{1,1}v_{3,2} - v_{3,1}v_{1,2})$, not $(v_{1,1}v_{2,2} - v_{1,2}v_{2,1})$ as before.) And the contribution of the last pair of faces is

$$\int_0^h \int_0^h \big(2zhv_{1,3} + O(h^4)\big)(v_{2,1}v_{3,2} - v_{2,2}v_{3,1})\, ds\, dt = (2zh^3v_{1,3} + O(h^4))(v_{2,1}v_{3,2} - v_{2,2}v_{3,1}).$$

Dividing by h^3 and taking the limit as $h \to 0$ gives

$$\begin{aligned}\mathbf{d}(z^2\, dx \wedge dy)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)\big) = (2zv_{3,3})&(v_{1,1}v_{2,2} - v_{1,2}v_{2,1}) - (2zv_{2,3})(v_{1,1}v_{3,2} - v_{3,1}v_{1,2}) \\ &+ (2zv_{1,3})(v_{2,1}v_{3,2} - v_{2,2}v_{3,1}).\end{aligned}$$

Note that in this solution we have reversed our usual rule for subscripts, where the row comes first and column second.

Now it helps to know what one is looking for. Since $\mathbf{d}(z^2\, dx \wedge dy)$ is a 3-form on $\mathbb{R}^3$, it is a multiple of the determinant. If you compute

$$\det[\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3] = \begin{bmatrix} v_{1,1} & v_{2,1} & v_{3,1} \\ v_{1,2} & v_{2,2} & v_{3,2} \\ v_{1,3} & v_{2,3} & v_{3,3} \end{bmatrix},$$

you will see that

$$\mathbf{d}(z^2\, dx \wedge dy)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)\big) = 2z \det[\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3].$$

b. Using Theorem 6.7.4, we have

$$\begin{aligned}\mathbf{d}(z^2\, dx \wedge dy) &\overbrace{=}^{\text{part 5}} \mathbf{d}\, z^2 \wedge dx \wedge dy \\ &\overbrace{=}^{\text{part 4}} (D_1 z^2\, dx + D_2 z^2\, dy + D_3 z^2\, dz) \wedge dx \wedge dy \\ &= 2z\, dz \wedge dx \wedge dy \\ &\overbrace{=}^{\text{Prop. 6.1.15}} 2z\, dx \wedge dy \wedge dz.\end{aligned}$$

6.7.7 a. The four edges of $P_{-\mathbf{e}_2}(h\vec{\mathbf{e}}_2, h\vec{\mathbf{e}}_3)$ are parametrized by

$$1. \quad t \mapsto \begin{pmatrix} 0 \\ -1+h \\ 0 \end{pmatrix} + t \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix} \qquad 2. \quad t \mapsto \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + t \begin{bmatrix} 0 \\ 0 \\ h \end{bmatrix}$$

$$3. \quad t \mapsto \begin{pmatrix} 0 \\ -1 \\ h \end{pmatrix} + t \begin{bmatrix} 0 \\ h \\ 0 \end{bmatrix} \qquad 4. \quad t \mapsto \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + t \begin{bmatrix} 0 \\ h \\ 0 \end{bmatrix}.$$

The second and third are taken with a minus sign, and the first and fourth with a plus sign.

Only the first two contribute to the integral around the boundary, since dx_3 returns 0 for vectors with no vertical component. They give

$$\int_0^1 (-1+h)^2 h\, dt - \int_0^1 (-1)^2 h\, dt = -2h^2 + h^3.$$

Divide by h^2 and take the limit as $h \to 0$, to find -2.

b. We find

$$\mathbf{d}(x_2^2 dx_3) = 2x_2\, dx_2 \wedge dx_3 \quad \text{and} \quad 2x_2\, dx_2 \wedge dx_3 P_{-\mathbf{e}_2}(\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3) = -2 \cdot 1 = -2.$$

6.7.9 First compute $\mathbf{d}\omega$:

$$\begin{aligned} \mathbf{d}\omega &= \mathbf{d}\Big(p(y,z)\, dx + q(x,z)\, dy\Big) \\ &= D_2 p(y,z)\, dy \wedge dx + D_3 p(y,z)\, dz \wedge dx + D_1 q(x,z)\, dx \wedge dy \\ &\qquad + D_3 q(x,z)\, dz \wedge dy; \end{aligned}$$

comparing this with the formula $\mathbf{d}\omega = x\, dy \wedge dz + y\, dx \wedge dz$, we see that

$$\begin{aligned} x\, dy \wedge dz &= D_3 q(x,z)\, dz \wedge dy; \quad \text{i.e.,} \quad D_3 q(x,z) = -x; \\ y\, dx \wedge dz &= D_3 p(y,z)\, dz \wedge dx; \quad \text{i.e.,} \quad D_3 p(y,z) = -y. \end{aligned}$$

In addition,

$$0 = D_1 q(x,z)\, dx \wedge dy - D_2 p(y,z)\, dx \wedge dy; \quad \text{i.e.,} \quad D_1 q(x,z) = D_2 p(y,z).$$

The functions $q(x,z) = -xz$ and $p(y,z) = -yz$ satisfy these constraints, which gives us the 1-forms of the form

$$\omega = -yz\, dx - xz\, dy.$$

But since $\mathbf{dd}f = 0$, the complete list of 1-forms satisfying the conditions is larger:

$$\omega = -yz\, dx - xz\, dy + \mathbf{d}f, \quad \text{where } f \text{ is any (at least) } C^2 \text{ function on } \mathbb{R}^3.$$

6.7.11 a. This is a restatement of Theorem 1.8.1, part 5.

b. Any k-form can be written as a sum of k-forms of the form

$$a(\mathbf{x})\, dx_{i_1} \wedge \cdots \wedge dx_{i_k},$$

and any l-form can be written as a sum of l-forms of the form

$$b(\mathbf{x})\, dx_{j_1} \wedge \cdots \wedge dx_{j_l}.$$

The result then follows from Theorem 6.7.4, part 2 (the exterior derivative of the sum equals the sum of the exterior derivatives) and from Proposition 6.1.15 (distributivity of the wedge product). We will work it out in the case of a k-form $\varphi = \varphi_1 + \varphi_2$ and an l-form ψ, where we assume that Theorem 6.7.9 is true for φ_1, φ_2, and ψ.

We have

$$\begin{aligned}
\mathbf{d}\big((\varphi_1+\varphi_2)\wedge\psi\big) &= \mathbf{d}(\varphi_1\wedge\psi + \varphi_2\wedge\psi) = \mathbf{d}(\varphi_1\wedge\psi) + \mathbf{d}(\varphi_2\wedge\psi)\\
&= \Big(\mathbf{d}\varphi_1\wedge\psi + (-1)^k\varphi_1\wedge\mathbf{d}\psi\Big) + \Big(\mathbf{d}\varphi_2\wedge\psi + (-1)^k\varphi_2\wedge\mathbf{d}\psi\Big)\\
&= \big((\mathbf{d}\varphi_1+\mathbf{d}\varphi_2)\wedge\psi\big) + (-1)^k\big(\varphi_1\wedge\mathbf{d}\psi + \varphi_2\wedge\mathbf{d}\psi\big)\\
&= \mathbf{d}(\varphi_1+\varphi_2)\wedge\psi + (-1)^k(\varphi_1+\varphi_2)\wedge d\psi.
\end{aligned}$$

c. To simplify notation, in the equation below we write $a(\mathbf{x})$ and $b(\mathbf{x})$ as a and b. Recall that we can write the wedge product of a 0-form f and a form α either as $f\alpha$ or as $f\wedge\alpha$.

We now treat the function ab as the function f of Theorem 6.7.4, part 5:

$$\begin{aligned}
\mathbf{d}\big(\varphi\wedge\psi\big) &= \mathbf{d}\big(a\, dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge b\, dx_{j_1}\wedge\cdots\wedge dx_{j_l}\big) && [1]\\
&= \mathbf{d}(ab)\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l} && [2]\\
&= (a\,\mathbf{d}b + b\,\mathbf{d}a)\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l} && [3]\\
&= a\,\mathbf{d}b\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l}\\
&\qquad + b\,\underbrace{\mathbf{d}a\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}}_{d\varphi}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l} && [4]\\
&= a\,\mathbf{d}b\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l}\\
&\qquad + \mathbf{d}\varphi\wedge\underbrace{b\wedge dx_{j_1}\wedge\cdots\wedge dx_{j_l}}_{\psi} && [5]\\
&= \underbrace{a\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}}_{\varphi}\wedge(-1)^k\,\underbrace{\mathbf{d}b\, dx_{j_1}\wedge\cdots\wedge dx_{j_l}}_{\mathbf{d}\psi} + \mathbf{d}\varphi\wedge\psi && [6]\\
&= \mathbf{d}\varphi\wedge\psi + (-1)^k\varphi\wedge\mathbf{d}\psi.
\end{aligned}$$

To move the b next to the a in line [2] we use Proposition 6.1.15 (skew commutativity): since b is a 0-form we have

$$\begin{aligned}
&dx_{i_1}\wedge\cdots\wedge dx_{i_k}\wedge b\\
&= (-1)^{0\cdot k} b\wedge dx_{i_1}\wedge\cdots\wedge dx_{i_k}\\
&= b\, dx_{i_1}\wedge\cdots\wedge dx_{i_k}.
\end{aligned}$$

To go to line [2] we also use equation 6.7.9.

To go to line [3] we apply the formula $\mathbf{d}(fg) = f\,\mathbf{d}g + g\,\mathbf{d}f$ (Theorem 6.7.9, with $k=0$).

To go to [4] we use the distributivity of the wedge product (Proposition 6.1.15).

To go from [4] to [5] we use the commutativity of 0-forms. To go from [5] to [6] we again use skew commutativity, this time moving the 1-form $\mathbf{d}b$.

6.8.1 a. Numbers: $dx \wedge dy(\vec{\mathbf{v}}, \vec{\mathbf{w}})$, $\vec{\mathbf{u}} \cdot (\vec{\mathbf{v}} \times \vec{\mathbf{w}})$, $\operatorname{grad} f(\mathbf{x}) \cdot \vec{\mathbf{v}}$. Function: $\operatorname{div} \vec{F}$. Vector fields: $\operatorname{grad} f$, $\operatorname{curl} \vec{F}$

b.
$$\begin{aligned}
\operatorname{grad} f &= \vec{\nabla} f\\
\operatorname{div} \vec{F} &= \vec{\nabla} \cdot \vec{F}\\
\operatorname{curl} \vec{F} &= \vec{\nabla} \times \vec{F}\\
\mathbf{d} f &= W_{\vec{\nabla} f} = W_{\operatorname{grad} f} = D_1 f\, dx_1 + D_2 f\, dx_2 + D_3 f\, dx_3\\
\mathbf{d} W_{\vec{F}} &= \Phi_{\operatorname{curl} \vec{F}} = \Phi_{\vec{\nabla} \times \vec{F}}\\
\mathbf{d} \Phi_{\vec{F}} &= M_{\operatorname{div} \vec{F}} = M_{\vec{\nabla} \cdot \vec{F}}
\end{aligned}$$

c. Of course more than one right answer is possible; for example, in (i), $\operatorname{grad} \vec{F}$ could be changed to $\operatorname{curl} \vec{F}$. In (iii), $\Phi_{\vec{F}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3)$ is meaningful if it is in $\mathbb{R}^4$.

i.	$\operatorname{grad} f$	ii.	$\operatorname{curl} \vec{F}$	iii.	$\Phi_{\vec{F}}(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2)$
iv.	$W_{\vec{F}}$	v.	unchanged	vi.	unchanged

6.8.3

$$\operatorname{grad} f = \begin{bmatrix} 2xy \\ x^2 \\ 1 \end{bmatrix}; \qquad \operatorname{curl} \vec{F} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \times \begin{bmatrix} -y \\ x \\ xz \end{bmatrix} = \begin{bmatrix} 0 \\ -z \\ 2 \end{bmatrix};$$
$$\operatorname{div} \vec{F} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \cdot \begin{bmatrix} -y \\ x \\ xz \end{bmatrix} = x.$$

6.8.5 a. $\vec{\nabla} f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 0 \\ 2y \end{bmatrix}$ b. $\vec{\nabla} f \begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} 2x \\ -2y \end{bmatrix}$

c. $\vec{\nabla} f \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{x^2+y^2} \begin{bmatrix} 2x \\ 2y \end{bmatrix}$ d. $\vec{\nabla} f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \dfrac{\operatorname{sgn}(x+y+z)}{|x+y+z|} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Note that $\frac{\operatorname{sgn}(x+y+z)}{|x+y+z|}$ can be simplified to $\frac{1}{x+y+z}$. Of course there is no solution to part d if $x+y+z=0$.

6.8.7 a. We can write the 1-form $\varphi = xy\, dx + z\, dy + yz\, dz$ as the work form $W_{\vec{F}}$, where $\vec{F} = \begin{bmatrix} xy \\ z \\ yz \end{bmatrix}$.

b. In the language of forms,

$$\begin{aligned}
\mathbf{d} W_{\vec{F}} &= \mathbf{d}(xy\, dx + z\, dy + yz\, dz) = \mathbf{d}(xy) \wedge dx + \mathbf{d}(z) \wedge dy + \mathbf{d}(yz) \wedge dz\\
&= (D_1 xy\, dx + D_2 xy\, dy + D_3 xy\, dz) \wedge dx + (D_1 z\, dx + D_2 z\, dy + D_3 z\, dz) \wedge dy\\
&\quad + (D_1 yz\, dx + D_2 yz\, dy + D_3 yz\, dz) \wedge dz \qquad (1)\\
&= -x(dx \wedge dy) + (z-1)(dy \wedge dz).
\end{aligned}$$

Since a 2-form in $\mathbb{R}^3$ can be written

$$\Phi_{\vec{G}} = G_1\,dy \wedge dz - G_2\,dx \wedge dz + G_3\,dx \wedge dy,$$

the last line of equation (1) can be written $\Phi_{\vec{G}}$ for $\vec{G} = \begin{bmatrix} z-1 \\ 0 \\ -x \end{bmatrix}$.

It can also be written as $\Phi_{\vec{\nabla}\times\vec{F}}$, since $\vec{G}$ is precisely the curl of $\vec{F}$:

$$\vec{\nabla} \times \vec{F} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \times \begin{bmatrix} xy \\ z \\ yz \end{bmatrix} = \begin{bmatrix} D_2 yz - D_3 z \\ -D_1 yz + D_3 xy \\ D_1 z - D_2 xy \end{bmatrix} = \begin{bmatrix} z-1 \\ 0 \\ -x \end{bmatrix}.$$

Solution 6.8.9, part a: A less direct way to show this is to say that the exterior derivative of

$$\mathbf{d}f = W_{\text{grad}\, f} = W_{\begin{bmatrix} F_1 \\ F_2 \end{bmatrix}} = (F_1\,dx + F_2\,dy)$$

must be 0:

$$\mathbf{d}(F_1\,dx + F_2\,dy) = D_2F_1\,dy \wedge dx + D_1F_2\,dx \wedge dy = 0,$$

i.e., $D_1F_2 - D_2F_1 = 0$.

6.8.9 a. If $\vec{F} = \text{grad}\, f$, then $D_1 f = F_1$ and $D_2 f = F_2$, so that $D_2F_1 = D_2D_1 f$ and $D_1F_2 = D_1D_2 f$; these crossed partials are equal since f is of class C^2 (Corollary 3.3.10).

b. We have changed the statement of the exercise to

"Show that this is not necessarily true if we only assume that all second partials of f exist everywhere. "

The solution is: It is possible for the first partial derivatives to have partial derivatives without being differentiable, and then their crossed partials are not necessarily equal. See Example 3.3.9.

6.8.11 a. $W_{\begin{bmatrix} x \\ y \\ z \end{bmatrix}} = x\,dx + y\,dy + z\,dz$ and

$$\mathbf{d}(x\,dx + y\,dy + z\,dz) = dx \wedge dx + dy \wedge dy + dz \wedge dz = 0.$$

Indeed, a function f exists such that $\text{grad}\, f = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$; it is the function $f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \dfrac{x^2+y^2+z^2}{2}$.

b. $\Phi_{\begin{bmatrix} x \\ y \\ z \end{bmatrix}} = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy$; the exterior derivative is

$$\begin{aligned} &\mathbf{d}(x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy) \\ &\qquad = dx \wedge dy \wedge dz + dy \wedge dz \wedge dx + dz \wedge dx \wedge dy \\ &\qquad = 3\,dx \wedge dy \wedge dz. \end{aligned}$$

6.8.13 a.

$$\text{div} \begin{bmatrix} x^2y \\ -2yz \\ x^3y^2 \end{bmatrix} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \cdot \begin{bmatrix} x^2y \\ -2yz \\ x^3y^2 \end{bmatrix} = 2xy - 2z$$

$$\text{curl} \begin{bmatrix} x^2y \\ -2yz \\ x^3y^2 \end{bmatrix} = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \times \begin{bmatrix} x^2y \\ -2yz \\ x^3y^2 \end{bmatrix} = \begin{bmatrix} 2x^3y + 2y \\ -3x^2y^2 \\ -x^2 \end{bmatrix}$$

b.

$$\operatorname{div}\begin{bmatrix}\sin xz\\ \cos yz\\ xyz\end{bmatrix} = \begin{bmatrix}D_1\\ D_2\\ D_3\end{bmatrix}\cdot\begin{bmatrix}\sin xz\\ \cos yz\\ xyz\end{bmatrix} = z\cos xz - z\sin yz + xy$$

$$\operatorname{curl}\begin{bmatrix}\sin xz\\ \cos yz\\ xyz\end{bmatrix} = \begin{bmatrix}D_1\\ D_2\\ D_3\end{bmatrix}\times\begin{bmatrix}\sin xz\\ \cos yz\\ xyz\end{bmatrix} = \begin{bmatrix}xz + y\sin yz\\ x\cos xz - yz\\ 0\end{bmatrix}.$$

c. **Curl of the vector field in part a**

We will compute $\mathbf{d}W_{\vec{F}} = \Phi_{\operatorname{curl}\vec{F}}$, the flux form field of $\operatorname{curl}\vec{F}$, and from that we will compute $\operatorname{curl}\vec{F}$. Since

$$W_{\vec{F}} = x^2y\,dx - 2yz\,dy + x^3y^2\,dz$$

is a 1-form, its exterior derivative is a 2-form. By Definition 6.7.1, we have

$$\mathbf{d}W_{\vec{F}} = \lim_{h\to 0}\frac{1}{h^2}\int_{\partial(P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,h\vec{\mathbf{v}}_2))} x^2y\,dx - 2yz\,dy + x^3y^2\,dz.$$

If you attempt to compute this using random vectors $\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2$, you will regret it.

Rather than computing this with vectors $\vec{\mathbf{v}}_1,\vec{\mathbf{v}}_2$, remember (Theorem 6.1.8) that any 2-form can be written in terms of the elementary 2-forms. So

$$\mathbf{d}W_{\vec{F}} = \Phi_{\operatorname{curl}\vec{F}} = a\,dx\wedge dy + b\,dy\wedge dz + c\,dx\wedge dz \tag{1}$$

for some coefficients a, b, c. Theorem 6.1.8 says that to determine the coefficients, we should evaluate $\mathbf{d}W_{\vec{F}}$ on the standard basis vectors.

Thus to determine the coefficient of $dx\wedge dy$ we will integrate $W_{\vec{F}}$ over the oriented boundary of the parallelogram spanned by $h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2$, computing

$$\lim_{h\to 0}\frac{1}{h^2}\int_{\partial P_{\mathbf{x}}(h\vec{\mathbf{e}}_1,h\vec{\mathbf{e}}_2)} x^2y\,dx - 2yz\,dy + x^3y^2\,dz.$$

To do this, we parametrize each edge of the parallelogram shown in the figure at left and give it a plus or minus depending on its orientation:

1.	$P_{\mathbf{x}}(h\vec{\mathbf{e}}_1)$	parametrized by	$\gamma_1(t) = \mathbf{x} + t\vec{\mathbf{e}}_1$
2.	$P_{\mathbf{x}+h\vec{\mathbf{e}}_1}(h\vec{\mathbf{e}}_2)$	parametrized by	$\gamma_2(t) = \mathbf{x} + h\vec{\mathbf{e}}_1 + t\vec{\mathbf{e}}_2$
3.	$P_{\mathbf{x}+h\vec{\mathbf{e}}_2}(h\vec{\mathbf{e}}_1)$	parametrized by	$\gamma_3(t) = \mathbf{x} + h\vec{\mathbf{e}}_2 + t\vec{\mathbf{e}}_1$
4.	$P_{\mathbf{x}}(h\vec{\mathbf{e}}_2)$	parametrized by	$\gamma_4(t) = \mathbf{x} + t\vec{\mathbf{e}}_2$

First, we will determine the orientation of each edge. Using Proposition 6.6.26, we have

$$\partial P_{\mathbf{x}}(h\vec{\mathbf{e}}_1,h\vec{\mathbf{e}}_2) = (-1)^0\big(P_{\mathbf{x}+h\vec{\mathbf{e}}_1}(h\vec{\mathbf{e}}_2) - P_{\mathbf{x}}(h\vec{\mathbf{e}}_2)\big) + (-1)^1\big(P_{\mathbf{x}+h\vec{\mathbf{e}}_2}(h\vec{\mathbf{e}}_1) - P_{\mathbf{x}}(h\vec{\mathbf{e}}_1)\big),$$

so edges 1 and 2 come with a plus sign, while 3 and 4 get a minus.

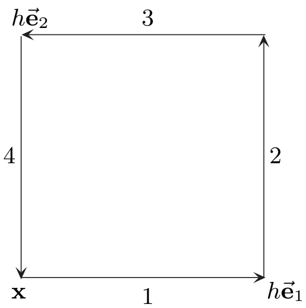


FIGURE FOR SOLUTION 6.8.13.

The parallelogram spanned by $h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2$. In our calculations, we denote by x, y, z the entries of $\mathbf{x}$: $\mathbf{x} = \begin{pmatrix} x \\ y \\ x \end{pmatrix}$.

We must integrate $x^2y\,dx - 2yz\,dy + x^3y^2\,dz$ over each parametrized edge, but note that for the first edge (and the third), we are only concerned with $x^2y\,dx$, since $dy(\vec{\mathbf{e}}_1) = 0$ and $dz(\vec{\mathbf{e}}_1) = 0$. If you don't see this, recall (equation 6.4.30) how we integrate forms. To integrate

$$x^2y\,dx - 2yz\,dy + x^3y^2\,dz$$

over the first edge, we integrate $x^2y\,dx - 2yz\,dy + x^3y^2\,dz(\overrightarrow{D_t\gamma_1}(t))$ over the first edge parametrized by γ_1. But $\overrightarrow{D_t\gamma_1}(t) = \vec{\mathbf{e}}_1$.

Thus for edge 1 we compute

$$\begin{aligned}\frac{1}{h^2}\int_{\partial(P_{\mathbf{x}}(h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2))} x^2y\,dx &= \frac{1}{h^2}\int_0^h \overbrace{(x+t)^2y}^{x^2y \text{ eval. at } \gamma_1(t)} \quad dx(\vec{\mathbf{e}}_1)\,dt \\ &= \frac{1}{h^2}\int_0^h x^2y + 2xyt + t^2y\,dt \\ &= \frac{1}{h^2}\left(x^2yh + xyh^2 + \cdots\right)\end{aligned}$$

The h in the denominator might be worrisome – what will happen as $h \to 0$? But we will see that it cancels with a term from another edge. That is the point of having an oriented boundary!

Notice that t^2y gives a term in h^3; once we divide by h^2 it will be a term in h, which will go to 0 as $h \to 0$. So we can ignore it. The terms that count for edge 1 are $\dfrac{x^2y}{h} + xy$, both taken with a +.

For edge 2 we are only concerned with $-2yz\,dy$, since $dx(\vec{\mathbf{e}}_2) = 0$ and $dz(\vec{\mathbf{e}}_2) = 0$. We compute

$$\frac{1}{h^2}\int_0^h -2yz - 2tz\,dt,$$

which gives $\dfrac{-2yz}{h} - z$.

A similar computation for edge 3 gives the terms $\dfrac{x^2y}{h} + xy + x^2$, each term taken with a minus sign; for edge 4 we get $\dfrac{-2yz}{h} - z$, also taken with a minus sign. Thus we have

$$\begin{aligned}\lim_{h\to 0}\frac{1}{h^2}\int_{\partial(P_{\mathbf{x}}(h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2))} & x^2y\,dx - 2yz\,dy + x^3y^2\,dz \\ &= \underbrace{\frac{x^2y}{h} + xy}_{\text{from edge 1}} + \underbrace{\frac{-2yz}{h} - z}_{\text{from edge 2}} \underbrace{- \frac{x^2y}{h} - xy - x^2}_{\text{from edge 3}} + \underbrace{\frac{2yz}{h} + z}_{\text{from edge 4}} \\ &= -x^2.\end{aligned}$$

We can substitute this for a in equation 1:

$$\Phi_{\operatorname{curl}\vec{F}} = -x^2\,dx \wedge dy + b\,dy \wedge dz + c\,dx \wedge dz.$$

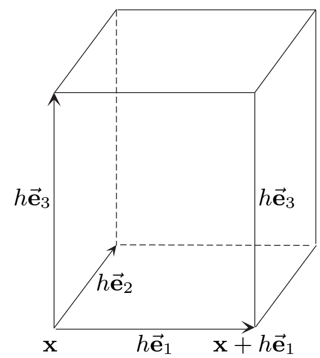
Recall that $\Phi_{\vec{F}} = F_1\,dy \wedge dz - F_2\,dx \wedge dz + F_3\,dx \wedge dy$, so $-x^2$ should be the third entry of $\operatorname{curl}\vec{F}$, which is indeed what we got in part a (with considerably less effort!).

For the coefficient of $dx \wedge dz$ we integrate over the boundary of the parallelogram spanned by $h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_3$. To save work, we note that for edges 1 and 3 we are interested only in $x^2y\,dx$, and for edges 2 and 4 we are only interested in $x^3y^2\,dz$. We get

$$\partial\big(P_{\mathbf{x}}(h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_3)\big) = \underbrace{\frac{x^2y}{h} + xy}_{\text{from edge 1}} + \underbrace{\frac{x^3y^2}{h} + 3x^2y^2}_{\text{from edge 2}} \underbrace{-\frac{x^2y}{h} - xy}_{\text{from edge 3}} - \frac{x^3y^2}{h} = 3x^2y^2.$$

Since in the formula for $\Phi_{\vec{F}}$ the coefficient of $dx \wedge dz$ is $-F_2$, the second entry of $\operatorname{curl}\vec{F}$ should be $-3x^2y^2$, which is what we got in part a.

A similar computation involving $\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3$ gives the coefficient for $dy \wedge dz$, i.e., the first entry of $\operatorname{curl}\vec{F}$.



The parallelogram spanned by $h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2, h\vec{\mathbf{e}}_3$.

Div of the vector field in part a

Now let's compute the divergence of the vector field in part a, by computing $\mathbf{d}\Phi_{\vec{F}} = M_{\operatorname{div}\vec{F}}$. Since

$$\Phi_{\vec{F}} = x^2y\,dy \wedge dz + 2yz\,dx \wedge dz + x^3y^2\,dx \wedge dy$$

is a 2-form, $\mathbf{d}\Phi_{\vec{F}}$ is a 3-form and can be written $\alpha\,dx \wedge dy \wedge dz$ for some coefficient $\alpha = \operatorname{div}\vec{F}$. We will compute $\mathbf{d}\Phi_{\vec{F}}$ by integrating

$$x^2y\,dy \wedge dz + 2yz\,dx \wedge dz + x^3y^2\,dx \wedge dy$$

over the oriented boundary of the parallelogram spanned by $h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2, h\vec{\mathbf{e}}_3$, shown at left.

The boundary consists of six faces, which we parametrize as follows, where $0 \le s, t \le h$, and $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$; we use Proposition 6.6.26 to determine which faces are taken with a plus sign and which are taken with a minus sign:

1. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + s\vec{\mathbf{e}}_1 + t\vec{\mathbf{e}}_2$, minus (the base of the box shown at left)
2. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + h\vec{\mathbf{e}}_3 + s\vec{\mathbf{e}}_1 + t\vec{\mathbf{e}}_2$, plus (the top of the box)
3. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + s\vec{\mathbf{e}}_1 + t\vec{\mathbf{e}}_3$, plus (the front of the box)
4. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + h\vec{\mathbf{e}}_2 + s\vec{\mathbf{e}}_1 + t\vec{\mathbf{e}}_3$, minus (the back of the box)
5. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + s\vec{\mathbf{e}}_2 + t\vec{\mathbf{e}}_3$, minus (the left side of the box)
6. $\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \mathbf{x} + h\vec{\mathbf{e}}_1 + s\vec{\mathbf{e}}_2 + t\vec{\mathbf{e}}_3$, plus (the right side of the box)

For the integral over the first face, we are concerned only with the term $x^3y^2\,dx \wedge dy$, since $dy \wedge dz(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2) = 0$ and $dx \wedge dz(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2) = 0$. So we have

$$\frac{1}{h^3}\int_{\partial P_{\mathbf{x}}(h\vec{\mathbf{e}}_1, h\vec{\mathbf{e}}_2)} x^3y^2\,dx \wedge dy = \frac{1}{h^3}\int_0^h\int_0^h \overbrace{(x+s)^3(y+t)^2 dx \wedge dy}^{x^3y^2dx\wedge dy \text{ eval. at } \gamma\begin{pmatrix} s \\ t \end{pmatrix}} \overbrace{(\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2)}^{\overrightarrow{D_s\gamma}, \overrightarrow{D_t\gamma}}\,ds\,dt$$

After discarding everything that will give terms in h^4, dividing by h^3, and giving the correct orientation, we are left with the following for the first face:

$$\frac{-x^3y^2}{h} - \frac{3x^2y^2}{2} - x^3y.$$

For the top of the box we get the same terms, with opposite sign.

For the third face (the front), we get

$$\frac{2yz}{h} + y.$$

For the fourth face (the back), we get

$$\frac{-2yz}{h} - y - 2z.$$

For the fifth and sixth faces, we are concerned only with $x^2y\,dy \wedge dz$. The fifth face gives

$$\frac{-x^2y}{h} - \frac{x^2}{2}.$$

For the sixth face, we get

$$\frac{x^2y}{h} + 2xy + \frac{x^2}{2}.$$

After cancellations, this leaves $2xy - 2z$. Thus

$$\mathbf{d}\Phi_{\vec{F}} = M_{\operatorname{div}\vec{F}} = (2xy - 2z)\,dx \wedge dy \wedge dz.$$

By equation 6.5.13, $\operatorname{div}\vec{F} = 2xy - 2z$, which is what we got in part a.

Curl and div of the vector field in part b

We do not work out curl and div of the second vector field, since the procedure is the same and you can check your results from part b.

6.9.1 a. To show that the pullback T^* is linear if T is a linear transformation, we need to show that for k-forms φ and ψ and scalar a,

$$\begin{aligned} T^*(\varphi + \psi) &= T^*(\varphi) + T^*(\psi); \\ T^*(a\varphi) &= aT^*\varphi. \end{aligned}$$

For the first:

We use Definition 6.1.5 of addition of k-forms to go from line 1 to line 2.

$$\begin{aligned} T^*(\varphi+\psi)(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k) &= (\varphi+\psi)\big(T(\vec{\mathbf{v}}_1),\dots,T(\vec{\mathbf{v}}_k)\big) \\ &= \varphi\big(T(\vec{\mathbf{v}}_1),\dots,T(\vec{\mathbf{v}}_k)\big) + \psi\big(T(\vec{\mathbf{v}}_1),\dots,T(\vec{\mathbf{v}}_k)\big) \\ &= T^*\varphi(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k) + T^*\psi(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k). \end{aligned}$$

For the second:

$$T^*\big(a\varphi(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big) = a\varphi\big(T(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big) = aT^*\varphi(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k).$$

b. To show that the pullback by a C^1 mapping $\mathbf{f}$ is linear, we have

$$\begin{aligned}\mathbf{f}^*(\varphi+\psi)\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big) &= (\varphi+\psi)\Big(P_{\mathbf{f}(\mathbf{x})}\big([\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_k\big)\Big)\\ &= \varphi\Big(P_{\mathbf{f}(\mathbf{x})}\big([\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_k\big)\Big)+\psi\Big(P_{\mathbf{f}(\mathbf{x})}\big([\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_k\big)\Big)\\ &= (\mathbf{f}^*\varphi)\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big)+(\mathbf{f}^*\psi)\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big)\end{aligned}$$

and

$$\begin{aligned}\mathbf{f}^*(a\varphi)\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big) &= (a\varphi)\Big(P_{\mathbf{f}(\mathbf{x})}\big([\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_k\big)\Big)\\ &= a\mathbf{f}^*\varphi\Big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\Big).\end{aligned}$$

6.9.3 a.

$$\begin{aligned}(\mathbf{f}^*W_\xi)P_{\mathbf{x}}(\vec{\mathbf{v}}) &= W_\xi P_{\mathbf{f}(\mathbf{x})}\big([\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}\big) = \xi\big(\mathbf{f}(\mathbf{x})\big)\cdot[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}} = \big(\xi\big(\mathbf{f}(\mathbf{x})\big)\big)^\top[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}\\ &= \Big([\mathbf{Df}(\mathbf{x})]^\top\xi\big(\mathbf{f}(\mathbf{x})\big)\Big)^\top\vec{\mathbf{v}} = \Big([\mathbf{Df}(\mathbf{x})]^\top\xi\big(\mathbf{f}(\mathbf{x})\big)\Big)\cdot\vec{\mathbf{v}}\\ &= W_{[\mathbf{Df}]^\top\xi\circ\mathbf{f}}P_{\mathbf{x}}(\vec{\mathbf{v}}),\end{aligned}$$

so $\mathbf{f}^*W_\xi = W_{[\mathbf{Df}]^\top\xi\circ\mathbf{f}}$.

b. We have $n=m$, so suppose $U\subset\mathbb{R}^n$ is open, $\mathbf{f}:U\to\mathbb{R}^n$ is C^1 and ξ is a vector field on $\mathbb{R}^n$. Then

$$\begin{aligned}\mathbf{f}^*\Phi_\xi\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1})\big) &= \Phi_\xi\Big(P_{\mathbf{f}(\mathbf{x})}\big([\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_{n-1}\big)\Big)\\ &= \det\Big[\xi\big(\mathbf{f}(\mathbf{x})\big),[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_{n-1}\Big]\\ &= \det\Big[[\mathbf{Df}(\mathbf{x})][\mathbf{Df}(\mathbf{x})]^{-1}\xi\big(\mathbf{f}(\mathbf{x})\big)[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_1,\dots,[\mathbf{Df}(\mathbf{x})]\vec{\mathbf{v}}_{n-1}\Big]\\ &= \det\Big([\mathbf{Df}(\mathbf{x})]\Big[[\mathbf{Df}(\mathbf{x})]^{-1}\xi\big((\mathbf{x})\big),\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}\Big]\Big)\\ &= \det[\mathbf{Df}(\mathbf{x})]\det\big[[\mathbf{Df}(\mathbf{x})]^{-1}\xi\big(\mathbf{f}(\mathbf{x})\big),\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1}\big]\\ &= \Phi_{(\det[\mathbf{Df}])[\mathbf{Df}]^{-1}(\xi\circ\mathbf{f})}\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_{n-1})\big)\end{aligned}$$

So $\mathbf{f}^*\Phi_\xi = \Phi_{(\det[\mathbf{Df}])[\mathbf{Df}]^{-1}(\xi\circ\mathbf{f})} = \det[\mathbf{Df}]\Phi_{[\mathbf{Df}]^{-1}(\xi\circ\mathbf{f})}$

c. Let $\widetilde{A}_{[i,j]}$ be the matrix obtained from A by replacing the (i,j)th entry by 1, and all other entries of the ith row and jth column by 0. For instance,

$$\text{if}\quad A=\begin{bmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{bmatrix}\qquad\text{then}\quad \widetilde{A}_{[3,1]}=\begin{bmatrix}0&b_1&c_1\\0&b_2&c_2\\1&0&0\end{bmatrix}$$

Note that

$$\det\widetilde{A}_{[i,j]} = (-1)^{i+j}\det A_{[i,j]}. \tag{1}$$

Recall (Exercise 4.8.18) that Cramer's rule says that if $A\vec{\mathbf{x}}=\vec{\mathbf{b}}$ and if $A_i(\vec{\mathbf{b}})$ denotes the matrix A where the ith column has been replaced by $\vec{\mathbf{b}}$, then $x_i \det A = \det A_i(\vec{\mathbf{b}})$. When A is invertible, so that $\det A \neq 0$, then Cramer's rule says that

$$A \underbrace{\begin{bmatrix} \frac{\det A_1(\vec{\mathbf{b}})}{\det A} \\ \vdots \\ \frac{\det A_n(\vec{\mathbf{b}})}{\det A} \end{bmatrix}}_{\vec{\mathbf{x}}} = \vec{\mathbf{b}}, \quad \text{i.e.,} \quad A \begin{bmatrix} \det A_1(\vec{\mathbf{b}}) \\ \vdots \\ \det A_n(\vec{\mathbf{b}}) \end{bmatrix} = (\det A)\vec{\mathbf{b}}. \tag{2}$$

In the second equation in (2), both sides are continuous functions of A, so the equation is true on the closure of invertible matrices, which is all matrices.

Apply this to $\vec{\mathbf{b}} = \vec{\mathbf{e}}_k$, noting that $\det \widetilde{A}_{[k,i]} = \det A_i(\vec{\mathbf{e}}_k)$. Then Cramer's rule says that

$$A \begin{bmatrix} \det \widetilde{A}_{[k,1]} \\ \vdots \\ \det \widetilde{A}_{[k,n]} \end{bmatrix} = (\det A)\vec{\mathbf{e}}_k,$$

so, using equation 1,

$$A \operatorname{adj}(A) = A \begin{bmatrix} \det \widetilde{A}_{[1,1]} & \cdots & \det \widetilde{A}_{[n,1]} \\ \vdots & \cdots & \vdots \\ \det \widetilde{A}_{[1,n]} & \cdots & \det \widetilde{A}_{[n,n]} \end{bmatrix} = (\det A)[\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_n] = (\det A)I. \tag{3}$$

Now we need to show that $\mathbf{f}^*\Phi_\xi = \Phi_{\operatorname{adj}[\mathbf{Df}](\xi\circ\mathbf{f})}$. For any vector $\xi \in \mathbb{R}^n$ and any invertible $n \times n$ matrix A we have

Equality 1 is the definition of the flux. Equality 2 introduces $AA^{-1} = I$. Equality 3 factors out A. Equality 4 is

$$\det(AB) = \det A \det B.$$

Equality 5 is multilinearity: multiplying one column of a matrix by a number multiplies the determinant by the same number. Equality 6 is $(\det A)A^{-1} = \operatorname{adj} A$, which follows from equation 3.

$$\begin{aligned} \Phi_\xi(P_{\mathbf{f}(\mathbf{x})}(A\vec{\mathbf{v}}_1, \dots, A\vec{\mathbf{v}}_{n-1}) &\underset{1}{=} \det[\xi(\mathbf{f}(\mathbf{x})), A\vec{\mathbf{v}}_1, \dots, A\vec{\mathbf{v}}_{n-1}] \\ &\underset{2}{=} \det[AA^{-1}\xi(\mathbf{f}(\mathbf{x})), A\vec{\mathbf{v}}_1, \dots, A\vec{\mathbf{v}}_{n-1}] \\ &\underset{3}{=} \det\Big(A[A^{-1}\xi(\mathbf{f}(\mathbf{x})), \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}]\Big) \\ &\underset{4}{=} \det A \det[A^{-1}\xi(\mathbf{f}(\mathbf{x})), \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}] \\ &\underset{5}{=} \det[(\det A)A^{-1}\xi(\mathbf{f}(\mathbf{x})), \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}] \\ &\underset{6}{=} \det[(\operatorname{adj}(A)\xi(\mathbf{f}(\mathbf{x})), \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}]. \end{aligned} \tag{4}$$

and again in both sides the first and last expressions are continuous functions of A, hence valid for all matrices. Applying this to $A = [\mathbf{D}f(\mathbf{x})]$ we find $\mathbf{f}^*\Phi_\xi = \Phi_{\operatorname{adj}[\mathbf{Df}](\xi\circ\mathbf{f})}$.

6.10.1 Give U its standard orientation as an open subset of $\mathbb{R}^3$, and ∂U the boundary orientation. Since

$$\begin{aligned} \mathbf{d}(z\,dx \wedge dy + y\,dz \wedge dx + x\,dy \wedge dz) &= dz \wedge dx \wedge dy + dy \wedge dz \wedge dx + dx \wedge dy \wedge dz \\ &= 3\,dx \wedge dy \wedge dz, \end{aligned}$$

we see that

$$\int_{\partial U} \frac{1}{3}(z\,dx \wedge dy + ydz \wedge dx + x\,dy \wedge dz) = \int_U dx \wedge dy \wedge dz,$$

which is the volume of U.

6.10.3 We will compute the integral in two ways. First we compute it directly. Let γ be the inverse of the projection mentioned in the text. The projection preserves orientation, so γ does too, and we may use γ to perform our integration. Let U be the region we wish to integrate over and set

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1, x_2, x_3 \geq 0, \text{ and } x_1 + x_2 + x_3 \leq a \right\}.$$

Since $\gamma(V) = U$, Definition 6.2.1 tells us that

$$\int_U x_1\, dx_2 \wedge dx_3 \wedge dx_4 = \int_V x_1\, dx_2 \wedge dx_3 \wedge dx_4 \left(\overrightarrow{D_1\gamma}, \overrightarrow{D_2\gamma}, \overrightarrow{D_3\gamma} \right) |dx_1\, dx_2\, dx_3|.$$

Since $\gamma \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ a - x_1 - x_2 - x_3 \end{pmatrix}$, with derivative $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}$, we have

$$x_1\, dx_2 \wedge dx_3 \wedge dx_4(\overrightarrow{D_1\gamma}, \overrightarrow{D_2\gamma}, \overrightarrow{D_3\gamma}), = x_1 \det \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix} = -x_1,$$

so we may write our integral as an iterated integral and evaluate:

$$\begin{aligned} \int_0^a \int_0^{a-x_1} \int_0^{a-x_1-x_2} -x_1\, dx_3\, dx_2\, dx_1 &= -\int_0^a \int_0^{a-x_1} x_1(a - x_1 - x_2)\, dx_2\, dx_1 \\ &= -\int_0^a \frac{x_1(a-x_1)^2}{2} dx_1 \\ &= -\frac{1}{2} \left[\frac{a^2 x_1^2}{2} - \frac{2a x_1^3}{3} + \frac{x_1^4}{4} \right]_0^a = -\frac{a^4}{24}. \end{aligned}$$

We can get the same result using the generalized Stokes's theorem. Let X be the region in $\mathbb{R}^4$ bounded by the three-dimensional manifolds given by equations $x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$, and $x_1 + x_2 + x_3 + x_4 = a$. Also let $\partial X_1, \dots, \partial X_5$ be the portions of these manifolds that bound X, respectively. We assume all are given orientations so that the projection of ∂X_5 onto (x_1, x_2, x_3)-coordinate space is orientation preserving. We may now apply the generalized Stokes's theorem (Theorem 6.10.2), which tells us that

$$\begin{aligned} \sum_{i=1}^{5} \int_{\partial X_i} x_1\, dx_2 \wedge dx_3 \wedge dx_4 &= \int_X \mathbf{d}(x_1\, dx_2 \wedge dx_3 \wedge dx_4) \\ &= \int_X dx_1 \wedge dx_2 \wedge x_3 \wedge x_4. \end{aligned}$$

The integrals over the first four portions, $\partial X_1, \dots, \partial X_4$, contribute nothing to the sum; the form $x_1\, dx_2 \wedge dx_3 \wedge dx_4$ is uniformly zero over these manifolds. We will spell it out for the first, ∂X_1.

Let $\varphi : \mathbb{R}^3 \to \mathbb{R}^4$ be the map

$$\varphi : \begin{pmatrix} x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix};$$

this mapping parametrizes the boundary of X_1. The function x_1 is 0 on this manifold, so the form evaluates to zero everywhere:

$$x_1\, dx_2 \wedge dx_3 \wedge dx_4(\overrightarrow{D_{x_2}}\varphi, \overrightarrow{D_{x_3}}\varphi, \overrightarrow{D_{x_4}}\varphi) = 0 \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0.$$

Similarly, the boundaries of X_2, X_3, and X_4 contribute nothing to the integral.

We now have

$$\int_{\partial X_5} x_1\, dx_2 \wedge dx_3 \wedge dx_4 = \int_X dx_1 \wedge dx_2 \wedge x_3 \wedge x_4,$$

and our problem has been reduced to calculating an integral over the region

$$X = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \middle|\, x_1, x_2, x_3, x_4 \geq 0,\ x_1 + x_2 + x_3 + x_4 \leq a \right\}.$$

To compute this integral, we first determine the orientation of X. All 4-forms in $\mathbb{R}^4$ are of the form $f(\mathbf{x})\, dx_1 \wedge dx_2 \wedge x_3 \wedge x_4 = f(\mathbf{x}) \det$. Assume X is oriented by $\operatorname{sgn} k$ det, for some constant $k \neq 0$. Since the vector $\vec{N} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ points out of X on ∂X_5, Definition 6.6.21 tells us that ∂X_5 is oriented by

$$\operatorname{sgn} k \det \left(\vec{N}, \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3 \right).$$

Since we wish the map γ to preserve orientation, we require that

$$k \det \left(\vec{N}, \overrightarrow{D_{x_1}}\gamma, \overrightarrow{D_{x_2}}\gamma, \overrightarrow{D_{x_3}}\gamma \right) = k \det \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} = -4k$$

be positive. This is satisfied if $k = -1$, so that $-\operatorname{sgn} \det$ orients X.

Now consider the identity parameterization of X. The partial derivatives with respect to $x_1, \dots, x_4$ are $\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_4$. Since $-\det$ evaluated on these vectors gives -1, the identity parameterization is orientation reversing. So

$$\int_{[I(X)]} dx_1 \wedge dx_2 \wedge x_3 \wedge x_4 = -\int_X |d^4\mathbf{x}|,$$

where I is the identity map on X. As an iterated integral this is

$$-\int_0^a \int_0^{a-x_1} \int_0^{a-x_1-x_2} \int_0^{a-x_1-x_2-x_3} dx_4\,dx_3\,dx_2\,dx_1 = -\int_0^a \int_0^{a-x_1} \int_0^{a-x_1-x_2} (a-x_1-x_2-x_3)\,dx_3\,dx_2\,dx_1$$
$$= -\int_0^a \int_0^{a-x_1} \frac{(a-x_1-x_2)^2}{2}\,dx_2\,dx_1$$
$$= -\int_0^a \frac{(a-x_1)^3}{6}\,dx_1 = -\frac{a^4}{24},$$

the same result as before.

6.10.5 First we will do this using Stokes's theorem. Set

$$\omega = x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy,$$

and call A the octant of solid ellipsoid given by

$$0 \le x,\ y,\ z \quad \text{and} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \le 1.$$

Observe that

$$\mathbf{d}\omega = 3\,dx \wedge dy \wedge dz.$$

Next, observe that the integral of $x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy$ over any subset of the coordinate planes vanishes. For instance, on the (x,y)-plane, only the term in $dx \wedge dy$ could contribute, and it doesn't, since $z = 0$ there. So $\int_S \omega = \int_A dx \wedge dy \wedge dz$.

The octant A is the image of the first octant in the unit sphere under the linear transformation $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, with determinant abc. So

$$\mathrm{vol}_3(A) = abc\left(\frac{1}{8}\right)\left(\frac{4\pi}{3}\right) = \frac{abc\,\pi}{6}.$$

Finally, $\displaystyle\int_S \omega = 3\,\mathrm{vol}_3(A) = \frac{abc\,\pi}{2}.$

Now we repeat the exercise using the parametrization

$$\gamma\begin{pmatrix} \theta \\ \varphi \end{pmatrix} = \begin{pmatrix} a\cos\varphi\cos\theta \\ b\cos\varphi\sin\theta \\ c\sin\varphi \end{pmatrix} \quad \text{with} \quad 0 \le \theta \le \frac{\pi}{2}, 0 \le \varphi \le \frac{\pi}{2}.$$

This parametrization is orientation preserving, and we have

$$\left[\mathbf{D}\gamma\begin{pmatrix} \theta \\ \varphi \end{pmatrix}\right] = \begin{bmatrix} -a\cos\varphi\sin\theta & -a\sin\varphi\cos\theta \\ b\cos\varphi\cos\theta & -b\sin\varphi\sin\theta \\ 0 & c\cos\varphi \end{bmatrix}.$$

Thus

$$\int_S x\,dy\wedge dz + y\,dz\wedge dx + z\,dx\wedge dy =$$

$$= \int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\Bigg(a\cos\varphi\cos\theta\det\begin{bmatrix} b\cos\varphi\cos\theta & -b\sin\varphi\sin\theta\\ 0 & c\cos\varphi\end{bmatrix}$$

$$+\, b\cos\varphi\sin\theta\det\begin{bmatrix} 0 & c\cos\varphi\\ -a\cos\varphi\sin\theta & -a\sin\varphi\cos\theta\end{bmatrix}$$

$$+\, c\sin\varphi\det\begin{bmatrix} -a\cos\varphi\sin\theta & -a\sin\varphi\cos\theta\\ b\cos\varphi\cos\theta & -b\sin\varphi\sin\theta\end{bmatrix}\Bigg)d\theta\,d\varphi.$$

$$= \int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\Big(abc\cos^3\varphi\cos^2\theta + abc\cos^3\varphi\sin^2\theta + abc\sin^2\varphi\cos\varphi\sin^2\theta$$

$$+\, abc\sin^2\varphi\cos\varphi\cos^2\theta\Big)d\varphi\,d\theta$$

$$= abc\int_0^{\frac{\pi}{2}}\int_0^{\frac{\pi}{2}}\big(\cos^3\varphi + \sin^2\varphi\cos\varphi\big)\,d\varphi\,d\theta$$

$$= abc\frac{\pi}{2}\int_0^{\frac{\pi}{2}}\cos\varphi\,d\varphi = abc\frac{\pi}{2}\Big[\sin\varphi\Big]_0^{\frac{\pi}{2}} = abc\frac{\pi}{2}.$$

Solution 6.10.7: Recall that the "bump" function $\beta_R : \mathbb{R}^n \to \mathbb{R}$ is given by

$$\begin{cases} e^{-1/(R^2-\mathbf{x}|^2)} & \text{if } |\mathbf{x}|^2 \le R^2\\ 0 & \text{if } |\mathbf{x}|^2 > R^2.\end{cases}$$

6.10.7 This is an application of the principle that "exponentials win over polynomials": suppose that p is a polynomial of some degree m, then

$$\lim_{x\to\infty}\frac{e^x}{|p(x)|} = \infty \quad \text{or, taking inverses,} \quad \lim_{x\to\infty} p(x)e^{-x} = 0.$$

To see this, note that the power series for e^x contains a term $x^{m+1}/(m+1)!$ and all the other terms are positive, and $|p(x)| \le Cx^m$ for an appropriate C and x sufficiently large, so

$$\lim_{x\to\infty}\frac{e^x}{|p(x)|} \ge \lim_{x\to\infty}\frac{x^{m+1}}{C(m+1)!x^m} = \infty.$$

Restating this one more time, for any poynomials p and q we have

$$\lim_{u\searrow 0}\frac{p(u)}{q(u)}e^{-1/u} = 0.$$

The function

$$e^{-1/(R^2-|\mathbf{x}|^2)}, \quad \text{for} \quad |\mathbf{x}| < R$$

is the composition of $g : \mathbf{x} \mapsto R^2 - |\mathbf{x}|^2$ and the function $u \mapsto e^{-1/u}$. The map g is C^∞, so the issue is to show that the map

$$F : u \mapsto \begin{cases} e^{-1/u} & \text{if } u > 0\\ 0 & \text{if } u \le 0\end{cases}$$

is C^∞, or equivalently if for all k the right-derivatives satisfy

$$\lim_{u\searrow 0} F^{(k)}(u) = 0,$$

to fit with the left-derivatives, which are obviously all 0. We have

$$F'(u) = \frac{1}{u^2}e^{-1/u},\ F''(u) = \left(-\frac{2}{u^3} + \frac{1}{u^4}\right)e^{-1/u}, \dots,$$

and it is not hard to see that all derivatives are of this form, hence indeed $\lim_{u \searrow 0} f^{(k)}(u) = 0$.

6.10.9 a. Let A be an $n \times k$ matrix, and consider the map $\operatorname{Mat}(n,k) \to \mathbb{R}$ given by

$$B \mapsto \det A^\top B.$$

Viewed as a function of the columns of B, this mapping is k-linear and alternating. It is alternating because exchanging two columns of B exchanges the same two columns of $A^\top B$, and hence changes the sign of $\det A^\top B$. To see that it is multilinear, write

$$B' = \left[\vec{\mathbf{b}}_1, \dots, \vec{\mathbf{b}}_{i-1}, \vec{\mathbf{b}}_i', \vec{\mathbf{b}}_{i+1}, \dots, \vec{\mathbf{b}}_k\right], \quad B'' = \left[\vec{\mathbf{b}}_1, \dots, \vec{\mathbf{b}}_{i-1}, \vec{\mathbf{b}}_i'', \vec{\mathbf{b}}_{i+1}, \dots, \vec{\mathbf{b}}_k\right],$$

and

$$B = \left[\vec{\mathbf{b}}_1, \dots, \vec{\mathbf{b}}_{i-1}, \alpha\vec{\mathbf{b}}_i' + \beta\vec{\mathbf{b}}_i', \vec{\mathbf{b}}_{i+1}, \dots, \vec{\mathbf{b}}_k\right].$$

Then $A^\top B$ is obtained from $A^\top B'$ and $A^\top B''$ by the same process of keeping all the columns except the ith, and replacing its ith column by

$$\alpha \cdot \text{the } i\text{th column of } A^\top B' + \beta \cdot \text{the } i\text{th column of } A^\top B'',$$

so that $\det A^\top B = \alpha \det A^\top B' + \beta \det A^\top B''$.

b. By part a, $B \mapsto \det A^\top B$ defines an element of $\varphi \in A^k(\mathbb{R}^n)$. By Theorem 6.1.8, it is of the form

$$\varphi = \sum_{1 \le i_1 < \dots < i_k \le n} c_{i_1,\dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

for coefficients $c_{i_1,\dots,i_k}$, which can be evaluated as follows:

$$\begin{aligned} c_{i_1,\dots i_k} &= \det\left(A^\top \left[\vec{\mathbf{e}}_{i_1}, \dots, \vec{\mathbf{e}}_{i_k}\right]\right) \\ &= \det \begin{bmatrix} a_{i_1,1} & \dots & a_{i_1,k} \\ \vdots & \ddots & \vdots \\ a_{i_k,1} & \dots & a_{i_k,k} \end{bmatrix}^\top . \end{aligned}$$

c. Thus

$$\varphi(\vec{\mathbf{a}}_1,\dots,\vec{\mathbf{a}}_k) = \det A^\top A = \sum_{1\le i_1<\dots<i_k\le n} \det \begin{bmatrix} a_{i_1,1} & \dots & a_{i_1,k} \\ \vdots & \ddots & \vdots \\ a_{i_k,1} & \dots & a_{i_k,k}\end{bmatrix}^\top dx_{i_1}\wedge\dots\wedge dx_{i_k}\Big[\vec{\mathbf{a}}_1,\dots,\vec{\mathbf{a}}_k\Big]$$

$$= \sum_{1\le i_1<\dots<i_k\le n} \det\left(\begin{bmatrix} a_{i_1,1} & \dots & a_{i_1,k} \\ \vdots & \ddots & \vdots \\ a_{i_k,1} & \dots & a_{i_k,k}\end{bmatrix}^\top \begin{bmatrix} a_{i_1,1} & \dots & a_{i_1,k} \\ \vdots & \ddots & \vdots \\ a_{i_k,1} & \dots & a_{i_k,k}\end{bmatrix}\right)$$

$$= \sum_{1\le i_1<\dots<i_k\le n} \left(\det \begin{bmatrix} a_{i_1,1} & \dots & a_{i_1,k} \\ \vdots & \ddots & \vdots \\ a_{i_k,1} & \dots & a_{i_k,k}\end{bmatrix}\right)^2.$$

Solution 6.10.9, part c: Since $A = [\vec{\mathbf{a}}_1,\dots,\vec{\mathbf{a}}_k]$, we can write $A \mapsto \det A^\top A$ as

$$\varphi(\vec{\mathbf{a}}_1,\dots,\vec{\mathbf{a}}_k) = \det A^\top A.$$

To go to line 3, remember (Theorem 4.8.8) that $\det A = \det A^\top$ and (Theorem 4.8.4) that if A, B are $n\times n$ matrices,

$$\det AB = \det A \det B.$$

6.10.11 a. Since P is compact, it is enough to prove the result for a single point $\mathbf{p}\in P$. Clearly we can rotate and translate M so that $\mathbf{p}$ is the origin, and the tangent plane $T_{\mathbf{p}}M$ is $\mathbb{R}^k$. Then in a neighborhood of $\mathbf{p}$, the manifold M is the graph of a C^2 map $\mathbf{f}: U \to \mathbb{R}^{n-k}$, where U is a neighborhood of the origin in $\mathbb{R}^k$, with $[\mathbf{Df}(\mathbf{0})] = [0]$. By the compactness of P we may assume that U is the ball of radius $2R$ for some fixed R, and that both $\mathbf{f}$ and its derivative depend continuously on $\mathbf{p}$.

Write $F = \begin{pmatrix} F_1 \\ F_2\end{pmatrix}$ and note that

$$F_2\begin{pmatrix}\mathbf{u}\\ \mathbf{y}\end{pmatrix} = \mathbf{y} - \mathbf{f}(\mathbf{u})$$

so that F_2 is by definition of class C^2, with

$$\Big[\mathbf{D}F_2\begin{pmatrix}\mathbf{0}\\ \mathbf{y}\end{pmatrix}\Big]\begin{bmatrix}\mathbf{h}\\ \mathbf{k}\end{bmatrix} = \mathbf{k}.$$

All the trouble is with F_1. Even it is not so bad: it can be written

$$F_1(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|}\sqrt{|\mathbf{u}|^2 + |\mathbf{f}(\mathbf{u})|^2} = \mathbf{u}\sqrt{1+\frac{|\mathbf{f}(\mathbf{u})|^2}{|\mathbf{u}|^2}}$$

$$= \mathbf{u}\left(1 + \frac{1}{2}\frac{|\mathbf{f}(\mathbf{u})|^2}{|\mathbf{u}|^2} + o\left(\frac{|\mathbf{f}(\mathbf{u})|^2}{|\mathbf{u}|^2}\right)\right) = \mathbf{u} + o(|\mathbf{u}|)$$

It is not obvious that F_1 is of class C^2 or even of class C^1. It looks quite a bit like the function in Example 1.9.5, the typical function with directional derivatives, but whose derivative is discontinuous: in both cases, there is an $|\mathbf{u}|$ in the denominator.

This proves that F is differentiable at the origin and that its derivative is the identity.

By the Taylor rules, F will be differentiable with Lipschitz derivative if and only if $\dfrac{\mathbf{u}|\mathbf{f}(\mathbf{u})|^2}{|\mathbf{u}|^2}$ is.

We saw many such examples in Chapter 1; see Examples 1.9.3 and 1.9.5 and Exercises 1.9.1 and 1.9.2. The numerator starts with terms of degree 5, and they are divided by $|\mathbf{u}|^2$. This is differentiable, and the various partial derivatives have terms of degree 4 in the numerator, and still $|\mathbf{u}|^2$ in the denominator. So they are also differentiable, with second partials having terms of degree 3 in the numerator, and still $|\mathbf{u}|^2$ in the denominator.

This proves part a.

b. The point $\begin{pmatrix}\mathbf{u}\\ \mathbf{y}\end{pmatrix}$ is in $(M \cap \partial B_r(\mathbf{0}))$ if and only if $\mathbf{y} = \mathbf{f}(\mathbf{u})$ and $|\mathbf{u}|^2 + |\mathbf{f}(\mathbf{u})|^2 = r^2$, and then

$$F\begin{pmatrix}\mathbf{u}\\ \mathbf{y}\end{pmatrix} = \begin{pmatrix} r\frac{\mathbf{u}}{|\mathbf{u}|}\\ \mathbf{0}\end{pmatrix} \in T_{\mathbf{0}}M \cap \partial B_r(\mathbf{0}).$$

c. Let $Y_r = M \cap \partial B_r(\mathbf{0})$, and let $\gamma : V \to Y_r$ be a parametrization, where V is an appropriate subset of $\mathbb{R}^{k-1}$. Then

$$\operatorname{vol}_{k-1} Y_r = \int_V \sqrt{\det\Big([\mathbf{D}\gamma(\mathbf{v})]^\top [\mathbf{D}\gamma(\mathbf{v})]\Big)}\; |d^{k-1}\mathbf{v}|,$$

and $F \circ \gamma$ is a parametrization of $F(Y_r)$, so

$$\begin{aligned}\operatorname{vol}_{k-1} F(Y_r) &= \int_V \sqrt{\det\Big([\mathbf{D}(F\circ\gamma)(\mathbf{v})]^\top [\mathbf{D}(F\circ\gamma)(\mathbf{v})]\Big)}\; |d^{k-1}\mathbf{v}|\\ &= \int_V \sqrt{\det\Big([\mathbf{D}\gamma(\mathbf{v})]^\top\Big) \underbrace{[\mathbf{D}F(\gamma(\mathbf{v}))]^\top [\mathbf{D}F(\gamma(\mathbf{v}))]}_{\approx\ \mathrm{id}}[\mathbf{D}\gamma(\mathbf{v})]\Big)}\; |d^{k-1}\mathbf{v}|.\end{aligned}$$

For r sufficiently small, the matrix $[\mathbf{D}F(\gamma(\mathbf{v}))]^\top[\mathbf{D}F(\gamma(\mathbf{v}))]$ is arbitrarily close to the identity: for any $\epsilon > 0$, there exists $\rho > 0$ such that if $0 < r < \rho$ we have

$$\det([\mathbf{D}\gamma(\mathbf{v})]^\top[\mathbf{D}\gamma(\mathbf{v})]) \le (1+\epsilon)^2 \det\Big([\mathbf{D}\gamma(\mathbf{v})]^\top[\mathbf{D}F(\gamma(\mathbf{v}))]^\top[\mathbf{D}F(\gamma(\mathbf{v}))][\mathbf{D}\gamma(\mathbf{v})]\Big).$$

Thus

$$\operatorname{vol}_{k-1} Y_r \le (1+\epsilon)\operatorname{vol}_{k-1} F(Y_r) \le (1+\epsilon) r^{k-1} \operatorname{vol}_{k-1} S^{k-1}.$$

6.11.1 Since $\operatorname{div}\vec{F} = 3$ and (by Theorem 6.8.3)

$$\mathbf{d}\Phi_{\vec{F}} = M_{\operatorname{div}\vec{F}} = \operatorname{div}\vec{F}\, dx \wedge dy \wedge dz,$$

Stokes's theorem says that

$$\int_S \Phi_{\vec{F}} = \int_X 3\, dx\wedge dy\wedge dz = 3\int_X |dx\, dy\, dz|,$$

where X is the solid torus bounded by S. This last integral can be evaluated in many ways: one is to integrate first with respect to $|dx\, dy|$, to find

$$\pi\Big((2+\sqrt{1-z^2})^2 - (2-\sqrt{1-z^2})^2\Big) = 8\pi\sqrt{1-z^2},$$

so the final result is

$$24\pi\int_{-1}^{1}\sqrt{1-z^2}\, dz = 12\pi^2.$$

6.11.3 a. In cylindrical coordinates, the integral becomes

$$\operatorname{vol}_3(Z_\alpha) = \int_0^\alpha \left(\int_X r\,|dr\, dz|\right) d\theta = \alpha\int_X x\,|dx\, dz|.$$

One way of saying this is that the volume is the area of X times the distance that the center of gravity travels. Recall (Definition 4.2.1) that the center of gravity of X is the point $\overline{\mathbf{x}}$ whose x-coordinate is given by

$$\overline{x} = \frac{\int_X x|d^2\mathbf{x}|}{\int_X |d^2\mathbf{x}|},$$

so that

$$\mathrm{vol}_3(Z_\alpha) = \underbrace{\alpha\overline{x}}_{\substack{\text{distance center of}\\ \text{gravity has traveled}}} \underbrace{\int_X |d^2\mathbf{x}|}_{\text{area of } X} = \alpha \int_X x\,|dx\,dz|.$$

b. The formula above says that it is $\pi(4\pi) = 4\pi^2$.

Solution 6.11.3, part c: The doughnut is cut in half vertically, not sliced horizontally like a bagel to be toasted.

c. Consider the full half-doughnut, bounded by the part of the torus where $y \geq 0$, and two circles. Since the vector field is radial, its flow through these two endpoints is 0. Notice that the divergence of our vector field is 3, so the flux of the vector field through the surface is equal to 3 times the volume of the half-doughnut, i.e., $6\pi^2$.

6.11.5 Denote by P the boundary of the polygon and by S a circle centered at the origin and with radius small enough so that S is inside the polygon (any circle will do, but a small one is convenient). Then the region U between S and the 11-sided polygon is a region bounded by S and the polygon. The curve P carries the boundary orientation of U, i.e., it is oriented counterclockwise (see Figure 6.6.11). If S is oriented as part of the boundary of U, it would be oriented clockwise. Give it the opposite orientation. Green's theorem (Stokes's theorem applied to curves) now says that for any 1-form ω,

$$\int_U \mathbf{d}\omega = \int_P \omega - \int_S \omega.$$

Solution 6.11.5: We have

$$\begin{aligned}\mathbf{d}W_{\begin{bmatrix} -y/(x^2+y^2) \\ x/(x^2+y^2)\end{bmatrix}} &= D_2\frac{-y}{x^2+y^2}\,dy\wedge dx \\ &\quad + D_1\frac{x}{x^2+y^2}dx\wedge dy \\ &= \frac{-(x^2+y^2)-2y(-y)}{(x^2+y^2)^2}dy\wedge dx \\ &\quad + \frac{(x^2+y^2)-2x(x)}{(x^2+y^2)^2}dx\wedge dy \\ &= \frac{x^2-y^2+y^2-x^2}{(x^2+y^2)^2}\,dx\wedge dy = 0\end{aligned}$$

Now observe the convenient fact that $\mathbf{d}W_{\begin{bmatrix} -y/(x^2+y^2) \\ x/(x^2+y^2)\end{bmatrix}} = 0$ (see the margin note). So we have $\int_P \omega = \int_S \omega$:

$$\begin{aligned}\int_P W_{\begin{bmatrix} -y/(x^2+y^2) \\ x/(x^2+y^2)\end{bmatrix}} &= \int_S W_{\begin{bmatrix} -y/(x^2+y^2) \\ x/(x^2+y^2)\end{bmatrix}} \\ &= \int_S -\frac{y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = 2\pi.\end{aligned}$$

6.11.7 Let U be the region bounded by C. Then

$$\int_C W_{\begin{bmatrix} 0 \\ x\end{bmatrix}} = \int_C x\,dy = \int_U dx\wedge dy,$$

whereas

$$\int_C W_{\begin{bmatrix} y \\ 0 \end{bmatrix}} = \int_C y\,dx = \int_U dy \wedge dx;$$

the first returns the area of U, and the second returns the negative of the area.

6.11.9 By Stokes's theorem,

$$\int_{\partial U_R} W_{\vec{F}} = \int_{U_R} \Phi_{\operatorname{curl} \vec{F}}.$$

The curl of $\vec{F}$ is almost constant on a small disc, so

$$\int_{U_R} \Phi_{\operatorname{curl} \vec{F}} = \underbrace{\pi R^2}_{\text{area of disc}} \big(\operatorname{curl} \vec{F}(\mathbf{a})\big) \cdot \vec{\mathbf{v}} + \underbrace{o(R^2)}_{\substack{\text{higher-degree terms from} \\ \text{Taylor expansion of } \vec{F} \text{ at } \mathbf{a}.}}$$

Thus

$$\lim_{R \to 0} \frac{1}{R^2} \int_{\partial U_R} W_{\vec{F}} \;=\; \pi \big(\operatorname{curl} \vec{F}(\mathbf{a})\big) \cdot \vec{\mathbf{v}}.$$

Solution 6.11.9: To express a vector field $\vec{F}$ as its Taylor series, express each component as its Taylor series:

$$\vec{F}(\mathbf{a}+\vec{\mathbf{h}}) = \begin{bmatrix} F_1(\mathbf{a}+\vec{\mathbf{h}}) \\ F_2(\mathbf{a}+\vec{\mathbf{h}}) \\ F_3(\mathbf{a}+\vec{\mathbf{h}}) \end{bmatrix} = \begin{bmatrix} F_1(\mathbf{a}) + [\mathbf{D}F_1(\mathbf{a})]\vec{\mathbf{h}} + \cdots \\ F_2(\mathbf{a}) + [\mathbf{D}F_2(\mathbf{a})]\vec{\mathbf{h}} + \cdots \\ F_3(\mathbf{a}) + [\mathbf{D}F_3(\mathbf{a})]\vec{\mathbf{h}} + \cdots \end{bmatrix}$$

6.11.11 You could simply argue that the flux must be 0 because the sphere is symmetrical about the origin, so that the flux through one part is canceled by the flux through another. Being more formal, you can use the divergence theorem to say

$$\int_{\partial \text{ball}} \Phi_{\begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}} = \int_{\text{ball}} \mathbf{d}\Phi_{\begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}} = \int_{\text{ball}} (2x+2y+2z)\,dx \wedge dy \wedge dz.$$

Again, at this point you could stop and point out that this integral is 0 by symmetry. But for practice setting up a multiple integrals we can go further:

$$\int_{\text{ball}} (2x+2y+2z)\,dx \wedge dy \wedge dz = \int_{-1}^{1} \left(\int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} \left(\int_{-\sqrt{1-y^2-z^2}}^{\sqrt{1-y^2-z^2}} (2x+2y+2z)\,dx \right) dy \right) dz.$$

(Actually computing this integral is unpleasant; take advantage of symmetry whenever you can.)

Where did we use the fact that the surface is oriented by the outward-pointing normal? The surface is the boundary of the ball, which has the standard orientation, so with the boundary orientation, the surface is oriented by the outward-pointing vector (see equation 6.6.22). If the surface had been oriented by the inward-pointing normal, the divergence theorem would say

$$\int_{\partial \text{ball}} \Phi_{\begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}} = -\int_{\text{ball}} \mathbf{d}\Phi_{\begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}}.$$

6.11.13 This is easiest using Stokes's theorem. Let S be the disc $x^2+y^2 \le 1,\ z=3$, oriented by the downward normal (downward because the circle is oriented clockwise). Then

$$\int_{\partial S} W_{\begin{bmatrix} -3y \\ 3x \\ 1 \end{bmatrix}} = \int_S \mathbf{d}W_{\begin{bmatrix} -3y \\ 3x \\ 1 \end{bmatrix}} = \int_S \Phi_{\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \times \begin{bmatrix} -3y \\ 3x \\ 1 \end{bmatrix}} = \int_S \Phi_{\begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}} = \int_S 6\, dx \wedge dy.$$

This integral is -6π: 6 times the area of the circle, with a minus sign because the vector field is pointing upwards, against the downwards orientation of S.

6.11.15 a. Suppose $\vec{\mathbf{w}} \in L$ is another vector. Then we can write

$$\vec{\mathbf{w}} = (a+bi)(\vec{\mathbf{v}}) = a\vec{\mathbf{v}} + bi\vec{\mathbf{v}} \quad \text{and} \quad i\vec{\mathbf{w}} = i(a+bi)(\vec{\mathbf{v}}) = -b\vec{\mathbf{v}} + a(i\vec{\mathbf{v}}).$$

Solution 6.11.15: Similar results were sketched in greater generality in Exercise 6.3.15.

Thus the change of basis matrix from the basis $\vec{\mathbf{w}}, i\vec{\mathbf{w}}$ to the basis $\vec{\mathbf{v}}, i\vec{\mathbf{v}}$ is

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

with determinant $a^2+b^2>0$, so these bases define the same orientation.

b. We will parametrize our domain by the map

$$z \mapsto \begin{pmatrix} z^q \\ -z^p \end{pmatrix}, \quad |z| \le R.$$

Recall that q and p are relatively prime, so at least one is odd. Assume that q is odd, and note that if we set $z_1 = z^q$ and $z_2 = -z^p$, then points in the image of the mapping satisfy the equation $(z^q)^p + (-z^p)^q = 0$. (If q and p were both even, this would not be true.)

To show that the mapping parametrizes our domain we also have to show (Definition 5.2.3) that it is one to one, of class C^1, with locally Lipschitz derivative, and that its derivative is also one to one. Since the coordinate functions z^q and $-z^p$ are polynomials, the mapping is not just C^1, it is C^∞; by Proposition 2.8.9, its derivative is Lipschitz.

Showing that the mapping is one to one means showing that a given pair $\begin{pmatrix} z^q \\ -z^p \end{pmatrix}$ corresponds to only one z. The number z^q has q qth roots (Proposition 0.7.7). If ζ is one of them, then the numbers

$$\zeta e^{2\pi i k/q}, \quad k = 0, 1, \dots, q-1$$

is a list of all of them. If we raise these to the pth power, we get

$$\zeta^p e^{2\pi i pk/q},\ k = 0, 1, \dots, q-1,$$

and these numbers are all different; indeed if $e^{2\pi i pk/q} = e^{2\pi i pl/q}$, then $e^{2\pi i p(k-l)/q} = 1$, which occurs only if $p(k-l)/q$ is an integer; but $p(k-l)$ is one of the integers $-p(q-1)\dots, -p, 0, p, \dots, p(q-1)$, and the only one of these which is divisible by q is 0. Thus our mapping is one to one. It should be clear that the derivative is one to one except at the origin.

Now we will use polar coordinates, writing (as in equation 0.7.10) $z = r(\cos\theta + i\sin\theta)$. In these terms our surface is parametrized by

$$\begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r^q \cos q\theta \\ r^q \sin q\theta \\ -r^p \cos p\theta \\ -r^p \sin p\theta \end{pmatrix}.$$

We wish to integrate over the part of $X_{p,q}$ where $|z_1| \le R_1, |z_2| \le R_2$. Thus $|z^q| \le R_1$ and $|-z^p| \le R_2$, which gives $|z| \le R_1^{1/q}, |z| \le R_2^{1/p}$. Set $R = \min(R_1^{1/q}, R_2^{1/p})$. Then the disc of radius R corresponds to the part of $X_{p,q}$ where $|z_1| \le R_1, |z_2| \le R_2$.

This leads to the integral

$$\begin{aligned}
\int_0^{2\pi}\int_0^R (dx_1 \wedge dy_1 + dx_2 \wedge dy_2)&\left(\begin{bmatrix} qr^{q-1}\cos q\theta \\ qr^{q-1}\sin q\theta \\ -pr^{p-1}\cos p\theta \\ -pr^{p-1}\sin p\theta \end{bmatrix}, \begin{bmatrix} -qr^q \sin q\theta \\ qr^q\cos q\theta \\ pr^p \sin p\theta \\ -pr^p\cos p\theta \end{bmatrix}\right) dr\, d\theta \\
&= \int_0^{2\pi}\int_0^R (q^2 r^{2q-1} + p^2 r^{2p-1})\, dr\, d\theta \\
&= 2\pi \left[\frac{q^2 r^{2q}}{2q} + \frac{p^2 r^{2p}}{2p}\right]_0^R \\
&= \pi(qR^{2q} + pR^{2p}).
\end{aligned}$$

6.12.1 Theorem 6.8.3 (or Table 6.8.1) tells you how to compute exterior derivatives of forms on $\mathbb{R}^3$ using grad, curl, and div. Here we have forms on $\mathbb{R}^4$, the vector fields $\vec{\mathbf{E}}$ and $\vec{\mathbf{B}}$ are time dependent, and the derivatives with respect to time have to be taken into account when relevant.

For the term $W_{\vec{\mathbf{E}}} \wedge c\,dt$, the derivatives of $\vec{\mathbf{E}}$ with respect to time do not contribute to the exterior derivative, since they would lead to terms with $dt \wedge dt$. We find

$$\begin{aligned}
\mathbf{d}W_{\vec{\mathbf{E}}} \wedge c\,dt &= \mathbf{d}(E_x\,dx + E_y\,dy + E_z\,dz) \wedge c\,dt \\
&= (D_y E_x\,dy + D_z E_x\,dz) \wedge dx \wedge c\,dt + (D_x E_y\,dx + D_z E_y\,dz) \wedge dy \wedge c\,dt \\
&\quad + (D_x E_z\,dx + D_y E_z\,dy) \wedge dz \wedge c\,dt.
\end{aligned}$$

Collecting terms, this leads to

$$\begin{aligned}
\mathbf{d}W_{\vec{\mathbf{E}}} \wedge c\,dt &= (D_y E_z - D_z E_y)\,dy \wedge dz \wedge c\,dt + (D_z E_x - D_x E_z)\,dz \wedge dx \wedge c\,dt \\
&\quad + (D_x E_y - D_y E_x)\,dx \wedge dy \wedge c\,dt \\
&= \Phi_{\vec{\nabla} \times \vec{\mathbf{E}}} \wedge c\,dt.
\end{aligned}$$

Note that we have simply repeated the computation in part 2 of Theorem 6.8.3, with the $c\,dt$ tagging along for the ride.

When we compute $\mathbf{d}\Phi_{\vec{\mathbf{B}}}$, the time derivatives of $\vec{\mathbf{B}}$ do contribute to the exterior derivative. We find

$$\begin{aligned}
\mathbf{d}\Phi_{\vec{\mathbf{B}}} &= \mathbf{d}(B_x dy \wedge dz + B_y dz \wedge dx + B_z dx \wedge dy) \\
&= (D_x B_x dx + D_t B_x dt) \wedge dy \wedge dz + (D_y B_y dy + D_t B_y dt) \wedge dz \wedge dx \\
&\quad + (D_z B_z dz + D_t B_z dt) \wedge dx \wedge dy).
\end{aligned}$$

Moving the dt's from the beginning to the end does not change the sign.

Separate the terms involving dt from the purely spatial terms, to find

$$\begin{aligned}\mathbf{d}\Phi_{\vec{\mathbf{B}}} &= (D_xB_x + D_yB_y + D_zB_z)\,dx\wedge dy\wedge dz\\ &\quad + D_tB_x\,dy\wedge dz\wedge dt + D_tB_y\,dz\wedge dx\wedge dt + D_tB_z\,dx\wedge dy\wedge dt\\ &= M_{\vec{\nabla}\cdot\vec{\mathbf{B}}} + \Phi_{\frac{1}{c}D_t\vec{\mathbf{B}}}\wedge c\,dt.\end{aligned}$$

Altogether, this does lead to

$$\mathbf{d}\mathbb{F} = \Phi_{\vec{\nabla}\times\vec{\mathbf{E}}}\wedge c\,dt + M_{\vec{\nabla}\cdot\vec{\mathbf{B}}} + \Phi_{\frac{1}{c}D_t\vec{\mathbf{B}}}\wedge c\,dt.$$

When we integrate a vector-valued function in $\mathbb{R}^n$ (or in $\mathbb{R}^3$, as in the case of the Biot-Savart formula), the answer is a vector in $\mathbb{R}^n$ (or in $\mathbb{R}^3$); we compute separately the integral of each entry:

$$\int\begin{bmatrix}a_1\\a_2\\a_2\end{bmatrix} = \begin{bmatrix}\int a_1\\ \int a_2\\ \int a_2\end{bmatrix}.$$

6.12.3 This is one of those unpleasant integrals requiring a couple of changes of variables. First,

$$\begin{bmatrix}I\\0\\0\end{bmatrix}\times\begin{bmatrix}x-s\\y\\z\end{bmatrix} = I\begin{bmatrix}0\\-z\\y\end{bmatrix}.$$

so we are left with showing that

$$\int_{-\infty}^{\infty}\frac{ds}{((x-s)^2+y^2+z^2)^{3/2}} = \frac{2}{y^2+z^2}.$$

To lighten notation, set $y^2+z^2=a^2$, and make a first change of variables $x-s=-au,\ ds=a\,du$, leading to

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{ds}{((x-s)^2+y^2+z^2)^{3/2}} &= \int_{\infty}^{\infty}\frac{a\,du}{(a^2u^2+a^2)^{3/2}}\\ &= \frac{1}{a^2}\int_{\infty}^{\infty}\frac{du}{(u^2+1)^{3/2}}.\end{aligned}$$

Next, set $u=\tan\theta,\ du=d\theta/\cos^2\theta$. This time the region of integration $u\in(-\infty,\infty)$ corresponds to $\theta\in(-\pi/2,\pi/2)$, and the integral becomes

$$\begin{aligned}\frac{1}{a^2}\int_{\infty}^{\infty}\frac{du}{(u^2+1)^{3/2}} &= \frac{1}{a^2}\int_{-\pi/2}^{\pi/2}\frac{\cos^3\theta\,d\theta}{\cos^2\theta} = \frac{1}{a^2}\int_{-\pi/2}^{\pi/2}\cos\theta\,d\theta\\ &= \frac{2}{a^2} = \frac{2}{y^2+z^2}.\end{aligned}$$

6.12.5 First let us check the integrability of the right side. Passing to spherical coordinates, we find

$$\int_{\mathbb{R}^3}\frac{1}{|\mathbf{u}|^2}\mathbf{1}_{B_R(\mathbf{0})}(\mathbf{u})|d^3\mathbf{u}| = \int_0^R\int_{-\pi/2}^{\pi/2}\int_0^{2\pi}\frac{1}{r^2}r^2\cos\varphi\,d\theta\,d\varphi\,dr = 4\pi R.$$

Now let us check that the inequality is correct. Suppose the support of $\vec{\mathbf{j}}$ is contained in the $B_\rho(\mathbf{0})$ and that $R>\rho+|\mathbf{x}_1|+1$; then we have $\vec{\mathbf{j}}(\mathbf{x}_1-\mathbf{u})=\vec{\mathbf{j}}(\mathbf{x}_2-\mathbf{u})=0$, so the inequality is true on the ball of radius R around $\mathbf{0}$.

Solution 6.12.5: It seems that inequality 1 should be an immediate consequence of Corollary 1.9.2. But it isn't quite, because that result concerns scalar-valued functions. Moreover, the proof doesn't go through for vector-valued functions, since it is based on Theorem 1.9.1, which is false for vector-valued functions. You can apply the proof for the various components of a vector-valued function, but then the points **c** you obtain for the components may be different, and conceivably the sum of the squares of the sup's at different points might be bigger than the sup of the sum of squares at one point. The argument given here shows that Corollary 1.9.2 is true anyway. The same argument is used in the proof of Proposition A5.1.

For any two vectors $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathbb{R}^3$ forming an angle θ,

$$|\vec{\mathbf{a}} \times \vec{\mathbf{b}}| = |\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}| \sin\theta \le |\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}|;$$

the length $|\vec{\mathbf{a}} \times \vec{\mathbf{b}}|$ is the area of the parallelogram spanned by $\vec{\mathbf{a}}, \vec{\mathbf{b}}$, i.e., $|\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}| \sin\theta$. Since

$$\left|\frac{\mathbf{u}}{|\mathbf{u}|^3}\right| = \frac{1}{|\mathbf{u}|^2},$$

we are left with checking that

$$|\vec{\mathbf{j}}(\mathbf{x}_1 - \mathbf{u}) - \vec{\mathbf{j}}(\mathbf{x}_2 - \mathbf{u})| \le |\mathbf{x}_1 - \mathbf{x}_2| \sup_{\mathbb{R}^3} |[\mathbf{D}\vec{\mathbf{j}}]|. \tag{1}$$

Set $\mathbf{g}(t) = \vec{\mathbf{j}}(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1))$, so that

$$|\vec{\mathbf{j}}(\mathbf{x}_2) - \vec{\mathbf{j}}(\mathbf{x}_1)| = |\mathbf{g}(1) - \mathbf{g}(0)|.$$

Now

$$\begin{aligned}
|\mathbf{g}(1) - \mathbf{g}(0)| &= \left|\int_0^1 \mathbf{g}'(t)dt\right| \\
&= \left|\int_0^1 [\mathbf{D}\vec{\mathbf{j}}(\mathbf{g}(t))](\mathbf{x}_2 - \mathbf{x}_1)\,dt\right| \\
&\le \int_0^1 \Big|[\mathbf{D}\vec{\mathbf{j}}(\mathbf{g}(t))]\Big| |\mathbf{x}_2 - \mathbf{x}_1|\,dt \le \sup_{\mathbb{R}^3} |[\mathbf{D}\vec{\mathbf{j}}]|\,|\mathbf{x}_2 - \mathbf{x}_1|.
\end{aligned}$$

Solution 6.12.7: A transformation satisfying

$$|\mathbf{x}'|^2 - c^2(t')^2 = |\mathbf{x}|^2 - c^2t^2$$

is called a *Lorentz transformation* because it preserves the *Lorentz pseudolength*, analogous to ordinary Euclidean length.

6.12.7 Important as this result is, it is a straightforward algebraic calculation. Since $y' = y$ and $z' = z$, we can ignore those terms. Then

$$\begin{aligned}
|x'|^2 - c^2(t')^2 &= \frac{(x - vt)^2}{1 - v^2/c^2} - \frac{(t - vx/c^2)^2}{1 - v^2/c^2} \\
&= \frac{c^2x^2 - 2c^2cvt + c^2v^2t^2 - t^2c^4 + 2c^2tvx - v^2x^2}{c^2 - v^2} \\
&= \frac{c^2(x^2 - t^2c^2) - v^2(x^2 - c^2t^2)}{c^2 - v^2} = x^2 - c^2t^2.
\end{aligned}$$

6.12.9 a. In the absence of charges and currents, one of Maxwell's equations reads

$$\operatorname{div}\vec{\mathbf{E}} = g'(x - at) = 0.$$

b. No, unless g is constant. The sign of $\vec{\mathbf{B}}$ (or $\vec{\mathbf{E}}$, take your pick) is wrong. The Faraday 2-form is

$$\mathbb{F} = g(y - ct)\,dx \wedge c\,dt + g(y - ct)\,dx \wedge dy,$$

so

$$\begin{aligned}
\mathbf{d}\mathbb{F} &= g'(y - ct)\,dy \wedge dx \wedge c\,dt - g'(y - ct)\,c\,dt \wedge dx \wedge dy \\
&= -2g'(y - ct)\,dx \wedge dy \wedge c\,dt,
\end{aligned}$$

and this is 0 only if g is constant.

6.12.11 a. The vector field $\vec{F}_1$ is $\vec{x}/|x|$, i.e., the unit vector field that points right when $x > 0$ and left when $x < 0$. Like all the $\vec{F}_i$, it isn't defined at the origin. The flux of a vector field on $\mathbb{R}^n$ is always an $n-1$-form; in particular the flux $\Phi_{\vec{F}_1}$ of F_1 is a 0-form, i.e., a function, in fact $\Phi_{\vec{F}_1}(x) = \operatorname{sgn}(x)$.

Here f plays the role of ρ in equation 6.12.27, and g plays the role of $\vec{\mathbf{E}}$, so g' plays the role of $\operatorname{div}\vec{\mathbf{E}}$.

Thus the 1-dimensional analogue of the computation 6.12.27 says that if f is a C^1 function with compact support on $\mathbb{R}$, and

$$g(x) = \frac{1}{2}\int_{\mathbb{R}} f(y)\operatorname{sgn}(x-y)|dy|, \quad \text{then} \quad g'(x) = f(x).$$

b. Let us go through the computation step by step.

Step 1.

$$f(x) = -\lim_{r\to 0}\frac{1}{2}\int_{\partial B_r(x)} f(y)\Phi_{F_1(x-y)} = -\lim_{r\to 0}\frac{1}{2}\int_{\partial B_r(x)} f(y)\operatorname{sgn}(x-y).$$

To see this, remember how to integrate a 0-form, i.e., a function, over a union of oriented points:

At one boundary point we have $y = x + r$, so $\operatorname{sgn}(-r)$ is really $\operatorname{sgn}(x-(x+r))$; at the other boundary point $y = x - r$, so $\operatorname{sgn}(r)$ is really $\operatorname{sgn}(x-(x-r)$.

$$\begin{aligned}\frac{1}{2}\int_{\partial[x-r,x+r]} f(y)\operatorname{sgn}(x-y) &= \frac{1}{2}\left(f(x+r)\operatorname{sgn}(-r) - f(x-r)\operatorname{sgn}(r)\right)\\ &= -\frac{1}{2}\left(f(x+r)+f(x-r)\right)\end{aligned}$$

whose limit clearly tends to $-f(x)$ when $r \to 0$.

In steps 2–4 we omit the $\lim_{r\to 0}\frac{1}{2}$ that proceeds each integral.

Step 2.

$$\int_{\partial B_r(x)} f(y)\Phi_{\vec{F}_1(x-y)} = \int_{\partial B_r(0)} f(x+u)\Phi_{\vec{F}_1(u)}.$$

This is just a change of variables $y = x + u$.

Step 3.

$$\int_{\partial B_r(0)} f(x+u)\Phi_{\vec{F}_1(u)} = -\int_{A_{R,r}} D_u\Big(f(x+u)\operatorname{sgn}(u)\Big)du.$$

This is the fundamental theorem of calculus: integrating the derivative is the difference of the values at the endpoints. The region $A_{R,r}$ consists of the two intervals $[-R,-r]\cup[r,R]$; the endpoints $\pm R$ contribute nothing, since f vanishes there, and the endpoints $\pm r$ do contribute, but with $+r$ with a minus sign, since it is a lower limit, and $-r$ contributing with a plus sign, since it is an upper limit.

Step 4.

In step 4 we go from the derivative with respect to u to the derivative with respect to x (like going from $\operatorname{div}_{\mathbf{u}}$ to $\operatorname{div}_{\mathbf{x}}$ in equation 6.12.27).

$$-\int_{A_{R,r}} D_u\Big(f(x+u)\operatorname{sgn}(u)\Big)\,du = -\int_{A_{R,r}} D_x(f(x+u))\operatorname{sgn}(u)\,du.$$

A priori, $D_u\Big(f(x+u)\operatorname{sgn}(u)\Big)$ is the derivative of a product, so by the Leibnitz rule, $D(f_1f_2) = f_1D(f_2) + (Df_1)f_2$; it should be the sum of two terms. But one term vanishes: clearly the derivative of $\operatorname{sgn}(u)$ is 0 (except at the origin where it isn't defined). So we have two functions:

$$f_1(u) = f(x+u) \quad \text{and} \quad f_2(x) = f(x+u).$$

and both have derivative $f'(x+u)$, by the chain rule, so

$$D_u\Big(f(x+u)\operatorname{sgn}(u)\Big) = D_x\Big(f(x+u)\operatorname{sgn}(u)\Big) = f'(x+u)\operatorname{sgn}(u).$$

Step 5. Two things happen in step 5. For one thing, we can get rid of the $\lim_{r\to 0}$; now that we are differentiating with respect to x, not u, the integrand is a perfectly well-defined function, so

$$\lim_{r\to 0}\frac{1}{2}\int_{A_{R,r}} D_x(f(x+u))\operatorname{sgn}(u)\,du = \frac{1}{2}\int_{\mathbb{R}} D_x(f(x+u))\operatorname{sgn}(u)\,du.$$

For another thing, we can differentiate under the integral sign: the condition (from Theorem 4.11.22) is that

$$\frac{f(x_1+u)\operatorname{sgn}(u) - f(x_2+u)\operatorname{sgn}(u)}{x_1 - x_2} \le h(u)$$

for some integrable function h, but by the mean value theorem we may take

$$h(u) = \sup_{x\in\mathbb{R}} |f'(x)|\mathbf{1}_{[-R,R]}(u).$$

Step 6.

$$-D_x\frac{1}{2}\int_{\mathbb{R}} f(x+u)\operatorname{sgn}(u)\,du = D_x\frac{1}{2}\int_{\mathbb{R}} f(y)\operatorname{sgn}(x-y)\,dy = g'(x).$$

Again, this is just the change of variables $y = x+u$, this time as $u = y-x$. The change of signs comes from $\operatorname{sgn}(u) = \operatorname{sgn}(y-x) = -\operatorname{sgn}(x-y)$.

6.13.1 We will assume that the functions are all defined on either $\mathbb{R}^2$ or $\mathbb{R}^3$, which have no holes. Asking whether a vector field is the gradient of a function is the same as asking whether a function exists whose partial derivatives are the entries of the vector field. We know that $\mathbf{d}f = W_{\operatorname{grad} f}$, and that $\mathbf{d}\,\mathbf{d}f = 0$, so it is simply a matter of writing $\mathbf{d}\,\mathbf{d}f = \mathbf{d}(W_{\vec{F}})$ and seeing whether this is 0. If not, no such function f exists. If so (assuming the domain has no holes) then such a function does exist.

The vector fields (a)-(e), and (g) are gradients of functions. The vector fields (f), (h), (i), (j), (k), and (l) are not:

a. $W_{\begin{bmatrix}0\\1\end{bmatrix}} = dy, \quad \text{and} \quad \mathbf{d}(dy) = 0.$

b. $W_{\begin{bmatrix}x\\0\end{bmatrix}} = x\,dx, \quad \text{and} \quad \mathbf{d}(x\,dx) = 0.$

c. $W_{\begin{bmatrix}x\\y\end{bmatrix}} = x\,dx + y\,dy, \quad \text{and} \quad \mathbf{d}(x\,dx + y\,dy) = 0.$

d. $W_{\begin{bmatrix}x\\-y\end{bmatrix}} = x\,dx - y\,dy, \quad \text{and} \quad \mathbf{d}(x\,dx - y\,dy) = 0.$

e. $W_{\begin{bmatrix} y \\ x \end{bmatrix}} = y\,dx + x\,dy$ and $\mathbf{d}(y\,dx + x\,dy) = dy \wedge dx + dx \wedge dy = 0.$

f. $W_{\begin{bmatrix} -y \\ x \end{bmatrix}} = -y\,dx + x\,dy$ and $\mathbf{d}(-y\,dx + x\,dy) = 2\,dx \wedge dy$

g. $W_{\begin{bmatrix} y \\ x-y \end{bmatrix}} = y\,dx + (x-y)\,dy$ and

$$\mathbf{d}\big(y\,dx + (x-y)\,dy\big) = dy \wedge dx + dx \wedge dy = 0$$

h. $W_{\begin{bmatrix} x-y \\ x+y \end{bmatrix}} = (x-y)\,dx + (x+y)\,dy$

$$\mathbf{d}\big((x-y)\,dx + (x+y)\,dy\big) = -dy \wedge dx + dx \wedge dy = 2\,dx \wedge dy$$

i. $W_{\begin{bmatrix} x^2-y-1 \\ x-y \end{bmatrix}} = (x^2 - y - 1)\,dx + (x-y)\,dy$

$$\mathbf{d}\big((x^2 - y - 1)\,dx + (x-y)\,dy\big) = -dy \wedge dx + dx \wedge dy = 2\,dx \wedge dy$$

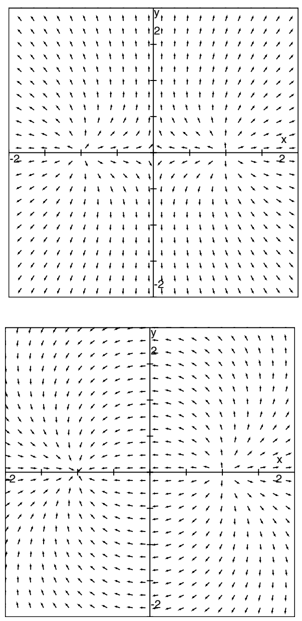


FIGURE FOR SOLUTION 6.13.3.

TOP: The vector field of part b. BOTTOM: The vector field of part (c). Since the z-coordinate of these vector fields is 0, and since z does not appear in the formulas, it is enough to draw the vector field in the (x, y)-plane.

6.13.3 a. The electric field is the gradient of the potential, i.e.

$$\vec{\nabla} V \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{x^2+y^2} \begin{bmatrix} 2x \\ 2y \\ 0 \end{bmatrix}.$$

Parts b and c: The figures at left show the vector fields

$$\text{b.} \quad \frac{1}{(x-1)^2+y^2} \begin{bmatrix} 2(-1) \\ 2y \\ 0 \end{bmatrix} + \frac{1}{(x+1)^2+y^2} \begin{bmatrix} 2(+1) \\ 2y \\ 0 \end{bmatrix}$$

and

$$\text{c.} \quad \frac{1}{(x-1)^2+y^2} \begin{bmatrix} 2(-1) \\ 2y \\ 0 \end{bmatrix} - \frac{1}{(x+1)^2+y^2} \begin{bmatrix} 2(+1) \\ 2y \\ 0 \end{bmatrix}.$$

6.13.5 a.

$$\begin{aligned} \mathbf{c}\big(x\,dx + y\,dy\big)\left(P_{\begin{pmatrix} u \\ v \end{pmatrix}}\right) &= \int_0^1 t^0 \big(x\,dx + y\,dy\big)\left(P_{t\begin{pmatrix} u \\ v \end{pmatrix}} \begin{bmatrix} u \\ v \end{bmatrix}\right) dt \\ &= \int_0^1 \Big((tu)u + (tv)v\Big)\,dt = \left[\frac{t^2}{2}\right]_0^1 u^2 + \left[\frac{t^2}{2}\right]_0^1 v^2 \\ &= \frac{u^2}{2} + \frac{v^2}{2}. \end{aligned}$$

Solution 6.13.5: Since only t_0 appears, we omit the index and write just t.

Indeed, $\mathbf{d}\left(\frac{u^2}{2} + \frac{v^2}{2}\right) = x\,dx + y\,dy.$

b.

$$\mathbf{c}\big(x\,dy - y\,dx\big)\Big(P_{\begin{pmatrix} u \\ v \end{pmatrix}}\Big) = \int_0^1 t^0\big(x\,dy - y\,dx\big)\Big((P_{t\begin{pmatrix} u \\ v \end{pmatrix}}\begin{bmatrix} u \\ v \end{bmatrix}\Big)dt$$

$$= \int_0^1 \Big((tu)v - (tv)u\Big)\,dt = 0$$

We want to check that

$$\overbrace{\mathbf{dc}(x\,dy - y\,dx)}^{0} + \mathbf{c}\,\overbrace{\mathbf{d}(x\,dy - y\,dx)}^{2\,dx\wedge dy} = x\,dy - y\,dx.$$

Indeed,

$$\mathbf{c}(2\,dx \wedge dy) = \int_0^1 2t\,dx \wedge dy\Big(P_{t\begin{pmatrix} a \\ b \end{pmatrix}}\Big(\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} u \\ v \end{bmatrix}\Big)\Big)dt$$

$$= \int_0^1 2t(av - bu)\,dt = av - bu,$$

so $\mathbf{c}(2\,dx \wedge dy) = x\,dy - y\,dx$, which confirms that

$$\mathbf{dc}(x\,dy - y\,dx) + \mathbf{cd}(x\,dy - y\,dx) = x\,dy - y\,dx.$$

6.13.7

Recall that

$$\mathbb{F} = W_{\vec{\mathbf{E}}} \wedge c\,dt + \Phi_{\vec{\mathbf{B}}}.$$

a. For $\vec{\mathbf{E}} = \begin{bmatrix} 0 \\ g(x-ct) \\ 0 \end{bmatrix}, \qquad \vec{\mathbf{B}} = \begin{bmatrix} 0 \\ 0 \\ g(x-ct) \end{bmatrix}$, we have

$$\mathbb{F} = W_{\vec{\mathbf{E}}} \wedge c\,dt + \Phi_{\vec{\mathbf{B}}} = g(x-ct)dy \wedge c\,dt + g(x-ct)\,dx \wedge dy$$

b. Define functions G and H by

$$G(u) \stackrel{\text{def}}{=} \int_0^u g(v)\,dv \quad \text{so that} \quad G' = g,\ G(0) = 0,\ \text{and}$$

$$H(u) \stackrel{\text{def}}{=} \int_0^u G(v)\,dv \quad \text{so that} \quad H' = G,\ H(0) = 0.$$

To compute $\mathbf{c}\mathbb{F}$, we use equation 6.13.24 in the case $k = 2$ to compute separately

$$\mathbf{c}\Big(g(x-ct)(dy \wedge c\,dt)\Big) \quad \text{and} \quad \mathbf{c}\Big(g(x-ct)\,dx \wedge dy\Big).$$

We could have evaluated the cone $\mathbf{c}\varphi$ first on

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \text{ then on } \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

In equations 1 and 3 below, s plays the role of t_0 in equation 6.13.24, and $\begin{bmatrix} 0 \\ \alpha \\ 0 \\ \beta \end{bmatrix} = \alpha\vec{\mathbf{e}}_2 + \beta\vec{\mathbf{e}}_4$ plays the role of parallelogram $P_{\mathbf{x}}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{k-1})$. We set the first and third entries to be 0 since there are no terms in dx and dz.

To compute the coefficients of dy and dt we compute the coefficients of α and β:

$$\begin{aligned}
\mathbf{c}\Big(\underbrace{g(x-ct)(dy \wedge c\,dt)}_{\varphi \text{ in eq. } 6.13.24}\Big) \begin{bmatrix} 0 \\ \alpha \\ 0 \\ \beta \end{bmatrix} &= \int_0^1 g\big(s(x-ct)\big)s \left(dy \wedge cdt \left(\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}, \begin{bmatrix} 0 \\ \alpha \\ 0 \\ \beta \end{bmatrix} \right) \right) ds \\
&= c(y\beta - t\alpha) \int_0^1 sg(s(x-ct))\,ds \\
&= c(y\beta - t\alpha) \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right). \qquad (1)
\end{aligned}$$

The coefficient of dy is the coefficient of α and the coefficient of dt is the coefficient of β, so

$$\begin{aligned}
\mathbf{c}\Big(g(x-ct)(dy \wedge c\,dt)\Big) = &- \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right) ct\,dy \\
&+ \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right) cy\,dt.
\end{aligned} \qquad (2)$$

Similarly,

$$\begin{aligned}
\mathbf{c}\Big(g(x-ct)\,dx \wedge dy\Big) \begin{bmatrix} \alpha \\ \beta \\ 0 \\ 0 \end{bmatrix} &= \int_0^1 sg\big(s(x-ct)\big) \left(dx \wedge dy \left(\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \\ 0 \\ 0 \end{bmatrix} \right) \right) ds \\
&= (\beta x - \alpha y) \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right) \qquad (3)
\end{aligned}$$

This time the coefficient of α is the coefficient of dx and the coefficient of β is the coefficient of dy, so:

$$\begin{aligned}
\mathbf{c}\Big(g(x-ct)\,dx \wedge dy\Big) = &-y \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right) dx \\
&+ x \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right) dy.
\end{aligned} \qquad (4)$$

Equation 5: Note that equations 2 and 4 both have the common factor
$$\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2}.$$

We have now computed the potential $\mathbb{A}$:

$$\mathbb{A} \overset{\text{def}}{=} \mathbf{c}\mathbb{F} = \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2} \right) \Big(-y\,dx + (x-ct)\,dy + y\,cdt \Big). \qquad (5)$$

c. Now we will check that $\mathbf{d}\mathbb{A} = \mathbb{F}$, using Theorem 6.7.9 for computing the exterior derivative of a wedge product:

$$\mathbf{d}\mathbb{A} = \mathbf{d}c\mathbb{F} = \mathbf{d}\Bigg(\overbrace{\Big(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2}\Big)}^{\varphi} \wedge \overbrace{\Big(-y\,dx + (x-ct)dy + y\,cdt\Big)}^{\psi}\Bigg)$$

$$= \mathbf{d}\Big(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2}\Big) \wedge \Big(-y\,dx + (x-ct)dy + y\,cdt\Big)$$
$$+ \Big(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2}\Big) \wedge \mathbf{d}\Big(-y\,dx + (x-ct)dy + y\,cdt\Big).$$

Theorem 6.7.9:
$\mathbf{d}(\varphi \wedge \psi)$
$= \mathbf{d}\varphi \wedge \psi + (-1)^{\deg \varphi} \varphi \wedge \mathbf{d}\psi$;
in this case, $\deg \varphi = 0$

First we compute

$$\mathbf{d}\Big(\frac{G(x-ct)}{x-ct}\Big) = \frac{(x-ct)\mathbf{d}G - G(x-ct)(dx - c\,dt)}{(x-ct)^2}.$$

Since $\mathbf{d}G = g(x-ct)\,dx - g(x-ct)\,c\,dt$, we have

$$\mathbf{d}\Big(\frac{G(x-ct)}{x-ct}\Big) = \frac{(x-ct)\Big(g(x-ct)\,dx - g(x-ct)\,c\,dt\Big) - G(x-ct)(dx - cdt)}{(x-ct)^2}.$$

Similarly,

$$\mathbf{d}\left(\frac{H(x-ct)}{(x-ct)^2}\right) = \frac{(x-ct)\Big(G(x-ct)\,dx - cG(x-ct)\,dt\Big) - 2H(x-ct)\,(dx - c\,dt)}{(x-ct)^3}.$$

(The function H is really the map $\begin{pmatrix} x \\ t \end{pmatrix} \mapsto H(x-ct)$; the exterior derivative of that function is $H'\,dx - cH'\,dt$; recall that $H' = G$.)

This gives

$$\mathbf{d}c\mathbb{F} = \left(\frac{g(x-ct)(dx-cdt)}{x-ct} - \frac{G(x-ct)(dx-cdt)}{(x-ct)^2} - \frac{G(x-ct)(dx-c\,dt)}{(x-ct)^2}\right.$$
$$\left. + \frac{2H(x-ct)(dx-c\,dt)}{(x-ct)^3}\right) \wedge \Big(-y\,dx + (x-ct)\,dy + y\,c\,dt\Big)$$
$$+ \left(\frac{G(x-ct)}{x-ct} - \frac{H(x-ct)}{(x-ct)^2}\right)\Big(2\,dx \wedge dy + 2\,dy \wedge c\,dt\Big).$$

Let us see that the terms in H cancel:

$$\frac{2H(x-ct)}{(x-ct)^2}\left(\frac{dx - c\,dt}{(x-ct)} \wedge \Big(-y\,dx + (x-ct)\,dy + y\,c\,dt\Big)\right) - \frac{2H(x-ct)}{(x-ct)^2}\Big(dx \wedge dy + dy \wedge c\,dt\Big)$$
$$= \frac{2H(x-ct)}{(x-ct)^2}\Big(dx \wedge dy - cdt \wedge dy - dx \wedge dy - dy \wedge c\,dt\Big) = 0.$$

Next let us get the terms in G to cancel:

$$\frac{2G(x-ct)}{x-ct}\left(\left(-\frac{dx-cdt}{x-ct}\right)\wedge\Big(-y\,dx+(x-ct)\,dy+y\,c\,dt\Big)+dx\wedge dy+dy\ \wedge c\,dt\right)$$

$$=\frac{2G(x-ct)}{x-ct}\Big(-dx\wedge dy-\underbrace{\frac{y}{x-ct}\,dx\wedge c\,dt-\frac{y\,c\,dt\wedge dx}{x-ct}}_{0}+c\,dt\wedge dy+dx\wedge dy+dy\wedge c\,dt\Big)=0.$$

Finally, we check that the surviving terms in g give us $\mathbb{F}$:

$$\frac{g(x-ct)}{x-ct}\left((dx-c\,dt)\wedge\Big(-y\,dx+(x-ct)\,dy+y\,c\,dt\Big)\right)$$

$$=\frac{g(x-ct)}{x-ct}\left((x-ct)\,dx\wedge dy+\overbrace{y\,dx\wedge c\,dt+y\,c\,dt\wedge dx}^{0}-(x-ct)\,c\,dt\wedge dy\right)$$

$$=g(x-ct)(dx\wedge dy+dy\wedge c\,dt)$$

6.13.9 a. The sequence of equalities reads

$$\begin{aligned}
\mathbf{d}(\mathbf{c}\varphi)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big) &\overset{(1)}{=} \lim_{h\to 0}\frac{1}{h^k}\int_{\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\mathbf{c}\varphi\\
&\overset{(2)}{=} \lim_{h\to 0}\frac{1}{h^k}\int_{C\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi\\
&\overset{(3)}{=} \lim_{h\to 0}\frac{1}{h^k}\Big(\int_{P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi-\int_{\partial CP_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi\Big)\\
&\overset{(4)}{=} \lim_{h\to 0}\frac{1}{h^k}\Big(\int_{P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi-\int_{CP_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\mathbf{d}\varphi\Big)\\
&\overset{(5)}{=} \varphi\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big)-\mathbf{c}(\mathbf{d}\varphi)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big).
\end{aligned}$$

Solution 6.13.9, part b: For equation (A), see equation 6.2.5 and Exercise 6.2.7.

b. From the definition of the integral of a k-form we get

$$\int_{P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi=h^k\varphi\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big)+o(h^k). \qquad (A)$$

Similarly, for the exterior derivative of a $k-1$-form we find

$$\int_{\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi=h^k(\mathbf{d}\varphi)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big)+o(h^k). \qquad (B)$$

and for the cone operator of a $k+1$-form we get

$$\int_{CP_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)}\varphi=h^k(\mathbf{c}\varphi)\big(P_{\mathbf{x}}(\vec{\mathbf{v}}_1,\dots,\vec{\mathbf{v}}_k)\big)+o(h^k). \qquad (C)$$

- Equality 1 in the equation in part a is the definition of the exterior derivative.

- Equality 2 follows from

$$\frac{1}{h^k}\int_{\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)} \mathbf{c}\varphi = \frac{1}{h}\mathbf{c}\varphi\left(\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)\right) + o\left(\frac{1}{h}\right)$$
$$= \frac{1}{h}\left(\frac{1}{h^{k-1}}\int_{C\partial P_{\mathbf{x}}(h\vec{\mathbf{v}}_1,\dots,h\vec{\mathbf{v}}_k)} \varphi + o\left(\frac{1}{h^{k-1}}\right)\right).$$

  The first equality is an instance of equation (A), the second an instance of equation (C).
- Equality 3 is equation 6.13.13: $\partial CP = P - C(\partial P)$.
- Equality 4 is Stokes's theorem.
- Equality 5 is an instance of equation (A) for the first term, and an instance of equation (C) for the second.

Solutions for review exercises, chapter 6

6.1 All are numbers except for (c) and (i). The cross product given in (c) is a vector in $\mathbb{R}^3$. In (i), $A^k(\mathbb{R}^k)$ is a 1-dimensional vector space; it is the set of k-forms on $\mathbb{R}^k$.

6.3 If φ is a 1-form and ψ is a 3-form, then

$$\varphi \wedge \psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4) = \varphi(\vec{\mathbf{v}}_1)\psi(\vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4) - \varphi(\vec{\mathbf{v}}_2)\psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_3, \vec{\mathbf{v}}_4) + \varphi(\vec{\mathbf{v}}_3)\psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_4) - \varphi(\vec{\mathbf{v}}_4)\psi(\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3).$$

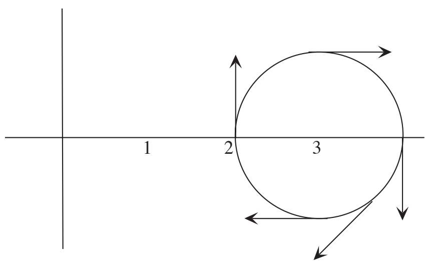


Figure for Solution 6.5.

At $\begin{pmatrix} 4 \\ 0 \end{pmatrix}$ we attach the vector $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$; at $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$ we attach $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$; at $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ we attach $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and so on. These all correspond to

$$\vec{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{bmatrix} y \\ 3-x \end{bmatrix}.$$

To confirm that this works at all points of the circle, we construct the translated radial vector field $\begin{bmatrix} x-3 \\ y \end{bmatrix}$; we have

$$\begin{bmatrix} x-3 \\ y \end{bmatrix} \cdot \begin{bmatrix} y \\ 3-x \end{bmatrix} = 0,$$

so $\begin{bmatrix} y \\ 3-x \end{bmatrix}$ is indeed tangent to the circle at every point.

6.5 The tangent vector field $\begin{bmatrix} y \\ 3-x \end{bmatrix}$ orients the circle clockwise. This corresponds to orientation by $y\,dx + (3-x)\,dy$, since

$$y\,dx + (3-x)\,dy\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = ay + (3-x)b = \begin{bmatrix} y \\ 3-x \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix}.$$

How did we find this tangent vector field? The sketch at left suggests one approach.

6.7 a. By Theorem 3.2.4, the tangent space $T_{\mathbf{x}}S^3$ is the kernel of the derivative at $\mathbf{x}$ of the equation $x_1^2 + x_2^2 + x_3^2 + x_4^2 - 1 = 0$. At the point $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ in S^3, the derivative is $[2\ 0\ 0\ 0]$, so the vectors $\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3, \vec{\mathbf{e}}_4 \in \mathbb{R}^4$ are in the tangent space $T_{\mathbf{x}}S^3$. But $dx_1 \wedge dx_2 \wedge dx_3(\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3, \vec{\mathbf{e}}_4) = 0$. So $dx_1 \wedge dx_2 \wedge dx_3$ is not a nonzero element of $A^3(T_{\mathbf{x}}S^3)$, and it cannot orient S^3.

b. The 3-form

$$\Omega_{\mathbf{x}} : (\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3) \mapsto \det[\mathbf{x}, \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3]$$

is certainly multilinear and alternating as a function of the $\vec{\mathbf{v}}_i$, since the determinant is, so $\Omega_{\mathbf{x}} \in A^3(T_{\mathbf{x}}S^3)$. The only problem is showing that at no point $\mathbf{x} \in S^3$ is $\Omega_{\mathbf{x}}$ the zero element of $A^3(T_{\mathbf{x}}S^3)$.

But if $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly independent elements of $T_{\mathbf{x}}S^3$, then $\mathbf{x}$ is orthogonal to $\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2$, and $\vec{\mathbf{v}}_3$, so $\mathbf{x}, \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3$ are linearly independent elements of $\mathbb{R}^4$, and hence

$$\det[\mathbf{x}, \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \vec{\mathbf{v}}_3] \neq 0.$$

6.9 We cannot use the chain rule to compute

$$[\mathbf{D}(\gamma_2^{-1} \circ \gamma_1)(\theta)] = [\mathbf{D}\gamma_2^{-1}(\gamma_1(\theta))][\mathbf{D}\gamma_1(\theta)];$$

the problem is that γ_2^{-1} is not defined on an open subset of $\mathbb{R}^2$, so it does not have a derivative $[\mathbf{D}\gamma_2^{-1}]$.

Instead we compute it directly. First we find θ_2 in terms of θ_1 such that

$$\begin{pmatrix}\sin\theta_2\\ \cos\theta_2\end{pmatrix} = \begin{pmatrix}\cos\theta_1\\ \sin\theta_1\end{pmatrix};$$

this gives $\theta_2 = \pi/2 - \theta_1$. So

$$\theta_1 \underbrace{\mapsto}_{\gamma_1} \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = \begin{pmatrix}\sin(\pi/2-\theta_1)\\ \cos(\pi/2-\theta_1)\end{pmatrix} \underbrace{\mapsto}_{\gamma_2^{-1}} \pi/2-\theta_1,$$

i.e. $(\gamma_2^{-1}\circ\gamma_1)(\theta_1) = \pi/2-\theta_1$, which gives

$$\det[\mathbf{D}(\gamma_2^{-1}\circ\gamma_1)(\theta_1)] = \det(-1) = -1.$$

6.11 a. The 1-form $(x^2+y^2)\,dz$ corresponds to the vector field

$$\vec{F}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{bmatrix}0\\ 0\\ x^2+y^2\end{bmatrix},$$

sketched at left. This vector field points straight up everywhere. It vanishes on the z-axis.

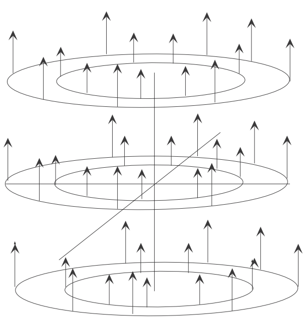


FIGURE FOR SOLUTION 6.11. The vector field

$$\vec{F}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{bmatrix}0\\ 0\\ x^2+y^2\end{bmatrix},$$

which we saw already in the solution to Exercise 1.1.7.

The work of this 1-form is 0 over any path contained entirely in the (x,y)-plane or in any plane parallel to the (x,y)-plane. It is large (and positive) over any path pointing straight up (in the direction of the positive z-axis), except that it is 0 over a path contained in the z-axis. The work is large but negative over any path pointing straight down.

If two paths both point straight up and begin at a point in the (x,y)-plane, then the work will be larger over the path that starts further from the origin.

A few simple computations confirm these conclusions for sample paths. For instance, take a straight path going from $x=0$ to $x=10$ in the horizontal plane at $z=1$. This path is parametrized by $\gamma : t \mapsto \begin{pmatrix}t\\ 0\\ 1\end{pmatrix}$, for $0<x<10$, so that $\vec{\gamma}'(t) = \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}$ and $\vec{F}_{\gamma(t)} = \begin{bmatrix}0\\ 0\\ t^2\end{bmatrix}$. Thus the work of $(x^2+y^2)\,dx$ over this path is

$$\int_0^{10}\begin{bmatrix}0\\ 0\\ t^2\end{bmatrix}\cdot\begin{bmatrix}1\\ 0\\ 0\end{bmatrix}\,dt = 0.$$

Now parametrize the straight vertical path going from $\begin{pmatrix}2\\ 2\\ 0\end{pmatrix}$ to $\begin{pmatrix}2\\ 2\\ 1\end{pmatrix}$ by $\gamma : t\mapsto \begin{pmatrix}2\\ 2\\ t\end{pmatrix}$, for $0<t<1$.

The work is 8, since

$$\int_0^1 \vec{F}\big(\gamma(t)\big) \cdot \vec{\gamma}'(t) = \int_0^1 \begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} dt = 8.$$

But the work over the vertical path from $\begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} 1/2 \\ 1/2 \\ 1 \end{pmatrix}$ is only 1/2.

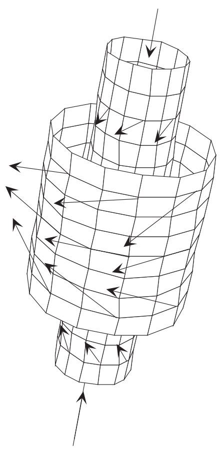


FIGURE FOR SOLUTION 6.11, part b: The vector field

$$\vec{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} y \\ -x \\ -z \end{bmatrix},$$

b. The 1-form $y\,dx - x\,dy - z\,dz$ corresponds to the vector field

$$\vec{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} y \\ -x \\ -z \end{bmatrix},$$

sketched at left and in the solution to Exercise 1.1.7. The work of this 1-form will be large over a path spiraling clockwise around the z-axis. It will be large on the path going straight down from $\begin{pmatrix} 0 \\ 0 \\ 10 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. It will be 0 on any horizonal radial path going straight out from the z-axis.

Again, we can confirm these conclusions with some simple computations. The vertical path from $\begin{pmatrix} 0 \\ 0 \\ 10 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ can be parametrized by the map $\gamma : t \mapsto \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix}$ for t from 10 to 0; the work of $y\,dx - x\,dy - z\,dz$ over this path is

$$\int_{10}^0 \overbrace{\begin{bmatrix} 0 \\ 0 \\ -t \end{bmatrix}}^{\vec{F}(\gamma(t))} \cdot \overbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}^{\vec{\gamma}'(t)} dt = \int_{10}^0 -t\,dt = \left[\frac{-t^2}{2}\right]_{10}^0 = 50.$$

The horizontal radial path from $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ to $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ can be parametrized by $\gamma : t \mapsto \begin{pmatrix} t \\ t \\ 1 \end{pmatrix}$, for $0 < t < 1$; the work of $y\,dx - x\,dy - z\,dz$ over this path is $\int_0^1 \begin{bmatrix} t \\ -t \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} dt = 0.$

6.13 The mapping

$$\gamma : \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ r \end{pmatrix}, 0 \le r \le \sqrt{R},\ 0 \le \theta \le \frac{\pi}{2}$$

(cylindrical change of coordinates, Definition 4.10.9) parametrizes the sector of cone, and

$$\overrightarrow{D_1\gamma} \times \overrightarrow{D_2\gamma} = \begin{bmatrix} -r\cos\theta \\ -r\sin\theta \\ r \end{bmatrix}$$

points inwards, so the parametrization preserves orientation.

Now we can compute the flux; since we have already computed $\overrightarrow{D_1\gamma} \times \overrightarrow{D_2\gamma}$ it is probably easier to use the equivalence $\det[\vec{\mathbf{a}}, \vec{\mathbf{b}}, \vec{\mathbf{c}}] = \vec{\mathbf{a}} \cdot (\vec{\mathbf{b}} \times \vec{\mathbf{c}})$:

$$\begin{aligned}\int_C \Phi_{\vec{F}} &= \int_0^{\sqrt{R}} \int_0^{\pi/2} \det\left[\vec{F}\left(\gamma\begin{pmatrix} r \\ \theta\end{pmatrix}\right), \overrightarrow{D_1\gamma}, \overrightarrow{D_2\gamma}\right] |d\theta||dr| \\ &= \int_0^{\sqrt{R}} \int_0^{\pi/2} \vec{F}\left(\gamma\begin{pmatrix} r \\ \theta\end{pmatrix}\right) \cdot \left(\overrightarrow{D_1\gamma} \times \overrightarrow{D_2\gamma}\right) d\theta||dr| \\ &= \int_0^{\sqrt{R}} \int_0^{\pi/2} \begin{bmatrix} r\sin\theta \\ -r \\ r^2\sin\theta \end{bmatrix} \cdot \begin{bmatrix} -r\cos\theta \\ -r\sin\theta \\ r \end{bmatrix} |d\theta||dr| \\ &= \int_0^{\sqrt{R}} \int_0^{\pi/2} (-r^2\cos\theta\sin\theta + r^2\sin\theta + r^3\sin\theta)|d\theta|\,|dr| \\ &= \frac{R\sqrt{R}}{6} + \frac{R^2}{4}.\end{aligned}$$

Solution 6.13: If you wish to check orientation using the determinant rather than the cross product, note that the cross product $\overrightarrow{D_1\gamma} \times \overrightarrow{D_2\gamma}$ is normal to the surface. In this case, it points inwards, so we can choose it as our inward-pointing normal. This gives

$$\det\left[\left(\overrightarrow{D_1\gamma} \times \overrightarrow{D_2\gamma}\right), \overrightarrow{D_1\gamma}, \overrightarrow{D_2\gamma}\right] = 2r^2 > 0.$$

In higher dimensions, where the cross product is not available, we could use the gradient. The cone is given by

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 - z^2 = 0,$$

so one vector normal to the cone is the gradient of f:

$$\vec{\nabla} f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} 2x \\ 2y \\ -2z \end{bmatrix}.$$

The gradient $\vec{\nabla}$ points in the direction f is increasing fastest, i.e., outwards, so $-\vec{\nabla}$ is an inward-pointing normal.

6.15 a. Denote by X the part of S^3 where $x_4 \le 0$. To show that X is a piece-with-boundary of S^3 according to Definition 6.6.8, we first need to define the smooth boundary (which is the entire boundary), using Definition 6.6.2. Let the U in that definition be $\mathbb{R}^4$, and define $f : U \to \mathbb{R}$ by

$$f(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + x_4^2 - 1,$$

so that $S^3 \cap \mathbb{R}^4$ is defined by the equation $f = 0$. Clearly the derivative $[\mathbf{D}f(\mathbf{x})] = [2x_i\ 2x_2\ 2x_3\ 2x_4]$ is surjective, since it vanishes only at the origin, which is not in S^3.

For the C^1 function g of Definition 6.6.2, we choose $g(\mathbf{x}) = -x_4$, with derivative $[0\ 0\ 0\ -1]$. At a point $\mathbf{x}$ of the smooth boundary, $g(\mathbf{x}) = 0$, so the derivative

$$\left[\mathbf{D}\begin{pmatrix} f \\ g \end{pmatrix}(\mathbf{x})\right] = \begin{bmatrix} 2x_1 & 2x_2 & 2x_3 & 2x_4 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

is surjective: since $x_4 = 0$, at least one of x_1, x_2, x_3 is nonzero. Thus by Proposition 6.6.3, $g : T_{\mathbf{x}}S^3 \to \mathbb{R}$ is onto.

It follows that the boundary of X consists of the sphere $x_1^2 + x_2^2 + x_3^2 = 1$, which certainly has finite 2-dimensional volume, so X satisfies condition 2 of Definition 6.6.8. There is no nonsmooth boundary, so it also satisfies condition 1.

b. At the point $\mathbf{x} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$, the tangent space $T_{\mathbf{x}}S^3$ is spanned by $\vec{\mathbf{e}}_2,\ \vec{\mathbf{e}}_3,\ \vec{\mathbf{e}}_4$, and the tangent space $T_{\mathbf{x}}\partial X$ is spanned by $\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3$. Moreover, $\vec{\mathbf{e}}_4$ is tangent to S^3 at $\mathbf{x}$ and outward-pointing from X. So $\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3$ is a direct basis of $T_{\mathbf{x}}\partial X$ if and only if $\vec{\mathbf{e}}_4, \vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3$ is a direct basis of $T_{\mathbf{x}}S^3$, which it is if and only if

$$\omega_{\mathbf{x}}(\vec{\mathbf{e}}_4, \vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3) = \det[\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_4,\ \vec{\mathbf{e}}_2,\ \vec{\mathbf{e}}_3]$$

is positive. This is indeed the case, so $\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3$ is a direct basis of $T_{\mathbf{x}}\partial X$.

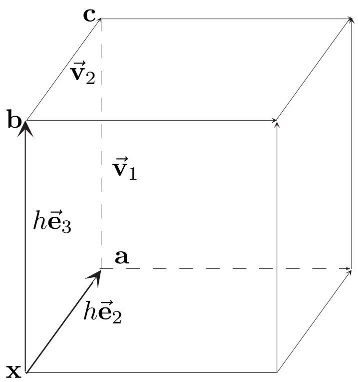


FIGURE FOR SOLUTION 6.17.

Above, $\mathbf{x}$ is the point $\begin{pmatrix}1\\2\\3\end{pmatrix}$, and the parallelogram spanned by $h\vec{\mathbf{e}}_2, h\vec{\mathbf{e}}_3$ is the left face of the cube.

The boundary orientation of $P_x(h\vec{\mathbf{e}}_2, h\vec{\mathbf{e}}_3)$ is given by Proposition 6.6.26. We interpret our sides as parallelograms; then the boundary consists of

$$\begin{aligned}+h\vec{\mathbf{e}}_2 &= P_{\mathbf{x}}(h\vec{\mathbf{e}}_2)\\ +\vec{\mathbf{v}}_1 &= P_{\mathbf{x}+h\mathbf{e}_2}(h\vec{\mathbf{e}}_3)\\ -\vec{\mathbf{v}}_2 &= -P_{\mathbf{x}+h\vec{\mathbf{e}}_3}(h\vec{\mathbf{e}}_2)\\ -h\vec{\mathbf{e}}_3 &= -P_{\mathbf{x}}(h\vec{\mathbf{e}}_3),\end{aligned}$$

where we use equation 6.6.23 to determine the signs. This corresponds to walking from $\mathbf{x}$ to $\mathbf{a}$ to $\mathbf{c}$ to $\mathbf{b}$ and back to $\mathbf{x}$.

Our 1-form is $\varphi = xyz\,dy$, which for the first parametrization corresponds to $1(2+th)(3)\,dy$, for the second to $1(2)(3+th)\,dy$, and so on.

6.17 a. The parallelogram $P_{\begin{pmatrix}1\\2\\3\end{pmatrix}}(h\vec{\mathbf{e}}_2, h\vec{\mathbf{e}}_3)$ is the left face of the cube in the margin; its boundary consists of $h\vec{\mathbf{e}}_2$, $h\vec{\mathbf{e}}_3$, $\vec{\mathbf{v}}_1$, and $\vec{\mathbf{v}}_2$. We have

$$\mathbf{a} = \begin{pmatrix}1\\2+h\\3\end{pmatrix},\quad \mathbf{b} = \begin{pmatrix}1\\2\\3+h\end{pmatrix},\quad \mathbf{c} = \begin{pmatrix}1\\2+h\\3+h\end{pmatrix}$$

so $h\vec{\mathbf{e}}_2$ is parametrized by γ_1 below; $h\vec{\mathbf{e}}_3$ is parametrized by γ_2; the side $\vec{\mathbf{v}}_1$ is parametrized by γ_4, and $\vec{\mathbf{v}}_2$ by γ_3:

$$\gamma_1(t) = \begin{pmatrix}1\\2+th\\3\end{pmatrix},\qquad \gamma_2(t) = \begin{pmatrix}1\\2\\3+th\end{pmatrix},$$

$$\gamma_3(t) = \begin{pmatrix}1\\2+th\\3+h\end{pmatrix},\qquad \gamma_4(t) = \begin{pmatrix}1\\2+h\\3+th\end{pmatrix},$$

all for $0 \le t \le 1$, and taken with the signs $+,\ -,\ -,\ +$. Using Definition 6.2.1, we need to compute, for each parametrizationγ_1, the integral

$$\int_0^1 xyz\,dy\Big(P_{\gamma_i(\mathbf{u})}\big(\gamma_i'(\mathbf{u})\big)\Big)|d^k\mathbf{u}|.$$

These integrals are

$$\int_0^1 1(2+th)(3)\,dy\begin{bmatrix}0\\h\\0\end{bmatrix}dt = \int_0^1 h(6+3th)\,dt = 6h + \frac{3}{2}h^2$$

$$-\int_0^1 1(2)(3+th)\,\left(dy\begin{bmatrix}0\\0\\h\end{bmatrix}\right)dt = 0$$

$$-\int_0^1 1(2+th)(3+h)\,dy\begin{bmatrix}0\\h\\0\end{bmatrix}dt = -6h - \frac{3}{2}h^2 - 2h^2 - \frac{h^3}{2}$$

$$+\int_0^1 1(2+h)(3+th)\,dy\begin{bmatrix}0\\0\\h\end{bmatrix}dt = 0$$

Add, divide by h^2, and take the limit as $h \to 0$, to find -2.

b. We have

$$\begin{aligned}\mathbf{d}(xyz\,dy) &= D_1 xyz\,dx \wedge dy + D_2 xyz\,dy \wedge dy + D_3 xyz\,dz \wedge dy\\ &= yz\,dx \wedge dy + xy\,dz \wedge dy.\end{aligned}$$

Evaluated on $\vec{\mathbf{e}}_2,\ \vec{\mathbf{e}}_3$, only the second term contributes, to give

$$\begin{aligned}xy\,dz \wedge dy\left(P_{\begin{pmatrix}1\\2\\3\end{pmatrix}}(\vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3)\right) &= 2\,dz \wedge dy\left(\begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right)\\ &= 2\det\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = -2.\end{aligned}$$

6.19 We have $\operatorname{curl}(\operatorname{grad} f) = \vec{\mathbf{0}}$ since

$$0 = \mathbf{dd}f = \mathbf{d}W_{\vec{\nabla} f} = \Phi_{\nabla \times \vec{\nabla} f} = \Phi_{\operatorname{curl}(\operatorname{grad} f)}.$$

We have $\operatorname{div}\operatorname{curl}\vec{F} = 0$ since

$$0 = \mathbf{dd}W_{\vec{F}} = \mathbf{d}\Phi_{\operatorname{curl}\vec{F}} = M_{\operatorname{div}\operatorname{curl}\vec{F}} = (\operatorname{div}\operatorname{curl}\vec{F})\,dx \wedge dy \wedge dz.$$

6.21 a. Write

$$\begin{aligned}&D_1\Big(x(x^2+y^2+z^2)^m\Big) + D_2\Big(y(x^2+y^2+z^2)^m\Big) + D_3\Big(z(x^2+y^2+z^2)^m\Big)\\ &= m(x^2+y^2+z^2)^{m-1}(2x^2) + (x^2+y^2+z^2)^m + m(x^2+y^2+z^2)^{m-1}(2y^2)\\ &\qquad + (x^2+y^2+z^2)^m + m(x^2+y^2+z^2)^{m-1}(2z^2) + (x^2+y^2+z^2)^m\\ &= (x^2+y^2+z^2)^{m-1}\Big(2m(x^2+y^2+z^2) + 3(x^2+y^2+z^2)\Big)\\ &= (2m+3)(x^2+y^2+z^2)^m.\end{aligned}$$

Solution 6.21: In part a, the exponent $m = -3/2$ corresponds to the vector field $\vec{F} = \vec{r}/r^3$. This gives an inverse square law (not a cube law because the $\vec{r}$ in the numerator has length r). This really explains the inverse square laws seen in gravitation and in Gauss's law in electricity, which says that the flux of the electric field through a surface measures the charge in the body bounded by the surface; the flux is not affected by the shape and area of the surface.

The only way for this to vanish is if $m = -3/2$.

b. Here are two possible approaches:

i. To simplify notation, set $\vec{F} = r^{2m}\vec{r}$. Note that $\vec{F}$ is a radial vector field. You should picture the vectors radiating out from the origin and passing through concentric $(n-1)$-dimensional spheres. The length of each vector depends only on the distance from the origin:

$$|\vec{F}| = r^{2m}|\vec{r}| = r^{2m+1},$$

with r being the distance from the origin. If $S_a(\mathbf{0})$ is the $(n-1)$-dimensional sphere of radius a centered at the origin, the integral of $\Phi_{\vec{F}}$ over $S_a(\mathbf{0})$ measures the flow of $\vec{F}$ through that sphere; i.e., the volume of the sphere times the length of the vector field on the sphere:

$$\int_{S_a(\mathbf{0})} \Phi_{\vec{F}} = \underbrace{a^{2m+1}}_{|\vec{F}|}\ \mathrm{vol}_{n-1}\, S_a(\mathbf{0}). \qquad (1)$$

For the sphere of radius 1 this gives

$$\int_{S_1(\mathbf{0})} \Phi_{\vec{F}} = \mathrm{vol}_{n-1} S_1(\mathbf{0}). \tag{2}$$

Further, since $\mathrm{vol}_{n-1} S_a(\mathbf{0}) = a^{n-1} \mathrm{vol}_{n-1} S_1(\mathbf{0})$, we have

$$\int_{S_a(\mathbf{0})} \Phi_{\vec{F}} = a^{n-1} a^{2m+1} \ \mathrm{vol}_{n-1} S_1(\mathbf{0}) \tag{3}$$

We will now use Stokes's theorem to see that if $\mathbf{d}\Phi_{\vec{F}} = 0$, the flow of $\vec{F}$ through any sphere centered at the origin is the same. Note that we cannot use Stokes to say that $\int_B \mathbf{d}\Phi_{\vec{F}} = \int_S \Phi_{\vec{F}}$, where B is the ball of which S is the boundary; since $\mathbf{d}\Phi_{\vec{F}} = 0$, that would give integrals of 0, which contradicts equations 1 and 2. Evidently, $\vec{F}$ is not defined at the origin. But we can set A to be the part of the ball between $S_1(\mathbf{0})$ and $S_a(\mathbf{0})$, which gives

Equation 4, far left: The integral over $S_1(\mathbf{0})$ is taken with a minus sign because the boundary orientation of $S_1(\mathbf{0})$ is the opposite of the boundary orientation of $S_a(\mathbf{0})$.

$$\int_{S_a(\mathbf{0})} \Phi_{\vec{F}} - \int_{S_1(\mathbf{0})} \Phi_{\vec{F}} = \int_A \mathbf{d}\Phi_{\vec{F}} = 0, \quad \text{i.e.,} \quad \int_{S_a(\mathbf{0})} \Phi_{\vec{F}} = \int_{S_1(\mathbf{0})} \Phi_{\vec{F}}. \tag{4}$$

Substituting equations 1 and 2 in equation 4 gives

$$a^{n-1}\, a^{2m+1} = a^{2m+n} = 1, \quad \text{i.e.,} \quad m = \frac{-n}{2}.$$

ii. We can do the computation as in part a: everything is as above, except that the 3 becomes an n. So the condition for the exterior derivative to vanish is $(2m) + n = 0$, i.e., $m = -n/2$.

6.23 a. The exterior derivative is

$$\left(D_1 \frac{x}{(x^2+y^2+z^2)^{3/2}} + D_2 \frac{y}{(x^2+y^2+z^2)^{3/2}} + D_3 \frac{z}{(x^2+y^2+z^2)^{3/2}}\right) dx \wedge dy \wedge dz$$
$$= \left(\frac{x^2+y^2+z^2-3x^2}{(x^2+y^2+z^2)^{5/2}} + \frac{x^2+y^2+z^2-3y^2}{(x^2+y^2+z^2)^{5/2}} + \frac{x^2+y^2+z^2-3z^2}{(x^2+y^2+z^2)^{5/2}}\right) dx \wedge dy \wedge dz = 0.$$

b. We cannot apply Stokes's theorem to the sphere as the boundary of a ball, since φ is not defined at the origin. Instead, parametrize the sphere, for instance by spherical coordinates, with $r = 1$:

$$\gamma \begin{pmatrix} \theta \\ \varphi \end{pmatrix} = \begin{pmatrix} \cos\theta\cos\varphi \\ \sin\theta\cos\varphi \\ \sin\varphi \end{pmatrix}, \quad \text{for} \quad 0 \le \theta < 2\pi, \ \frac{-\pi}{2} \le \varphi \le \frac{\pi}{2},$$

with

$$\overrightarrow{D_1}\gamma = \begin{bmatrix} -\sin\theta\cos\varphi \\ \cos\theta\cos\varphi \\ 0 \end{bmatrix} \quad \text{and} \quad \overrightarrow{D_1}2\gamma = \begin{bmatrix} -\cos\theta\sin\varphi \\ -\sin\theta\sin\varphi \\ \cos\varphi \end{bmatrix}.$$

This parametrization preserves orientation: the vector

$$\vec{\mathbf{n}} = \begin{bmatrix} x = \cos\theta\cos\varphi \\ y = \sin\theta\cos\varphi \\ z = \sin\varphi \end{bmatrix}$$

is an outward-pointing normal, and $\det\left[\vec{\mathbf{n}}, \overrightarrow{D_1}\gamma, \overrightarrow{D_2}\gamma\right] = \cos\varphi$, which is positive for φ between $-\pi/2$ and $\pi/2$.

Thus we find (note that the denominator $(x^2+y^2+z^2)^{3/2}$ is 1 on the unit sphere)

$$\int_0^{2\pi}\int_{-\pi/2}^{\pi/2}(\cos\theta\cos\varphi\, dy\wedge dz + \sin\theta\cos\varphi\, dz\wedge dx + \sin\varphi\, dx\wedge dy)\Bigg(\overbrace{\begin{bmatrix}-\sin\theta\cos\varphi\\ \cos\theta\cos\varphi\\ 0\end{bmatrix}}^{\overrightarrow{D_1}\gamma}, \overbrace{\begin{bmatrix}-\cos\theta\sin\varphi\\ -\sin\theta\sin\varphi\\ \cos\varphi\end{bmatrix}}^{\overrightarrow{D_2}\gamma}\Bigg)\, d\varphi\, d\theta$$

$$= \int_0^{2\pi}\int_{-\pi/2}^{\pi/2}\cos\varphi\; d\varphi\, d\theta = 4\pi.$$

c. The unit ball and the cube of side 4 together bound the region between the two, in which φ is well defined, so we can apply the divergence theorem, to show that the integral over the boundary of the cube is also 4π.

d. If φ is a 2-form and $\varphi = d\psi$, then the integral of φ over any oriented closed surface S is 0, since $\partial S = \phi$:

A compact manifold has empty boundary.

$$\int_S \varphi = \int_S \mathbf{d}\psi = \int_{\partial S}\psi = 0.$$

In our case, the integral over the unit sphere is 4π, so this 2-form cannot be written as the exterior derivative of a 1-form.

6.25 The divergence theorem says that if U is given the standard orientation,

$$\int_S \Phi_{\begin{bmatrix}x\\y\\z\end{bmatrix}} = \int_U \operatorname{div}\begin{bmatrix}x\\y\\z\end{bmatrix} dx\wedge dy\wedge dz.$$

Since the divergence above is 3, we find $\int_S \Phi_{\begin{bmatrix}x\\y\\z\end{bmatrix}} = 3\,\mathrm{vol}_3(U)$.

6.27 We will use primes to denote the coordinates of the system with the charge at rest. By equation 6.12.24, the electric and magnetic fields of the charge q at rest at the origin are

$$\vec{\mathbf{E}}(\mathbf{x}') = \frac{q}{4\pi|\mathbf{x}'|^3}\,\vec{\mathbf{x}}', \qquad \vec{\mathbf{B}}(\mathbf{x}') = \vec{\mathbf{0}}$$

i.e. (see equation 6.5.29),

$$\mathbb{F} = \frac{q}{4\pi({x'}^2+{y'}^2+{z'}^2)^{3/2}}\Big(x'\,dx' + y'\,dy' + z'\,dz'\Big)\wedge c\,dt', \qquad (1)$$

Equation 2: See equation 6.12.75, and recall that γ is the the Lorentz factor
$$\gamma = \frac{1}{\sqrt{1-v^2/c^2}}.$$

We need to write $f^*\mathbb{F}$, where denote by x, y, z, t the "moving" coordinates:

$$f\begin{pmatrix}x\\y\\z\\t\end{pmatrix} = \begin{pmatrix}\gamma(x-vt)\\ y\\ z\\ \gamma(t-vx/c^2)\end{pmatrix} = \begin{pmatrix}x'\\y'\\z'\\t'\end{pmatrix} \qquad (2)$$

Thus we can substitute into equation 1

$$\begin{aligned} x' &= \gamma(x - vt) & dx' &= \gamma(dx - v\,dt) \\ y' &= y & dy' &= dy \\ z' &= z & dz' &= dz \\ t' &= \gamma\left(t - \frac{v}{c^2}x\right) & dt' &= \gamma\left(dt - \frac{v}{c^2}\,dx\right) \end{aligned}$$

to find

$$f^*\mathbb{F} = \frac{q}{4\pi\Big(\gamma^2(x-vt)^2+y^2+z^2\Big)^{3/2}}\left(\Big(\gamma(x-vt)\gamma(dx-v\,dt)+y\,dy+z\,dz\Big) \wedge c\gamma\left(dt - \frac{v}{c^2}dx\right)\right).$$

On the right we have a 1-form wedge a 1-form, i.e., a 2-form.

Multiplying out gives

$$f^*\mathbb{F} = \frac{q}{4\pi\Big(\gamma^2(x-vt)^2+y^2+z^2\Big)^{3/2}}\left(\gamma^3(x-vt)\left(1-\frac{v^2}{c^2}\right)dx\wedge c\,dt + \gamma y\,dy\wedge c\,dt\right.$$
$$\left. + \gamma z\,dz\wedge c\,dt + 0\,dy\wedge dz + \frac{-vz\gamma}{c}\,dz\wedge dx + \frac{vy\gamma}{c}\,dx\wedge dy.\right)$$

Noticing that $\gamma^3\left(1-\frac{v^2}{c^2}\right) = \gamma$ gives

$$f^*\mathbb{F} = \frac{q}{4\pi\Big(\gamma^2(x-vt)^2+y^2+z^2\Big)^{3/2}}\left(\gamma(x-vt)\,dx\wedge c\,dt + \gamma y\,dy\wedge c\,dt\right.$$
$$\left. + \gamma z\,dz\wedge c\,dt - \frac{vz\gamma}{c}\,dz\wedge dx + \frac{vy\gamma}{c}\,dx\wedge dy\right),$$

leading (see again equation 6.5.29) to the electric and magnetic fields

$$\vec{\mathbf{E}} = \frac{q\gamma}{4\pi\Big(\gamma^2(x-vt)^2+y^2+z^2\Big)^{3/2}}\begin{bmatrix} x-vt \\ y \\ z \end{bmatrix}$$
$$\vec{\mathbf{B}} = \frac{q\gamma}{4\pi\Big(\gamma^2(x-vt)^2+y^2+z^2\Big)^{3/2}}\quad\frac{v}{c}\begin{bmatrix} 0 \\ -z \\ y \end{bmatrix}.$$

6.29 a. Since $\mathbf{d}\mathbb{F} = 0$ (see equation 6.12.11), we have

$$0 = \int_U \mathbf{d}\mathbb{F} = \int_{\partial U} \mathbb{F}.$$

In this case ∂U is a surface bounding a region in $\mathbb{R}^3$, and since it is in a region at fixed time,

$$\int_{\partial U} W_{\vec{\mathbf{E}}} \wedge c\,dt = 0.$$

The other term is

$$\int_{\partial U} \Phi_{\vec{\mathbf{B}}} = 0;$$

this says that exactly "as much" of $\vec{\mathbf{B}}$ flows in as flows out of U, i.e., that $\vec{\mathbf{B}}$ is "incompressible." It is often expressed by saying that there are no *magnetic monopoles*, no sources of the magnetic field.

b. This time, we have

$$4\pi \int_U J = \int_U \mathbf{d}\mathbb{M} = \int_{\partial U} \mathbb{M}.$$

Again, obviously

$$\int_{\partial U} W_{\vec{\mathbf{B}}} \wedge c\,dt = 0.$$

The other term is

$$\int_{\partial U} \Phi_{\vec{\mathbf{E}}} = 4\pi \int_U M_\rho.$$

The 4π in this expression depends on the units; in SI, the constant is $1/\epsilon_0$.

The flux of the vector field through a surface bounding a region U is 4π times the total charge in the region. The electric field does have sources (electrons and protons, for instance) and the flux of the electric field through a surface enclosing charge is proportional to this charge.

6.31 Suppose first that $R > 1$. Then S_R^{n-1} oriented by the outward-pointing normal and the unit sphere S_1^{n-1} with the orientation given by the inward-pointing normal form the oriented boundary of the region

$$B_{1,R} \stackrel{\text{def}}{=} \{\, \vec{\mathbf{x}} \mid 1 \le |\vec{\mathbf{x}}| \le R \,\}.$$

So by Stokes's theorem (giving S_1^{n-1} the orientation of the outward-pointing normal, hence the minus sign), we have

$$\int_{S_R^{n-1}} \Phi_{\vec{F}_n} - \int_{S_1^{n-1}} \Phi_{\vec{F}_n} = \int_{B_{1,R}} \mathbf{d}\Phi_{\vec{F}_n} = 0,$$

by equation 6.7.21.

If the word flux means what it supposed to mean, the fact that the flux of $\vec{F}_n$ through the unit sphere is the volume of the unit sphere should be obvious.

When $R < 1$, the same proof works: S_1^{n-1} oriented by the outward-pointing normal and S_R^{n-1} oriented by the inward-pointing normal form the oriented boundary of

$$B_{R,1} \stackrel{\text{def}}{=} \{\, \vec{\mathbf{x}} \mid R \le |\vec{\mathbf{x}}| \le 1 \,\}.$$

In all cases $\int_{S_R^{n-1}} \Phi_{\vec{F}_n} = \int_{S_1^{n-1}} \Phi_{\vec{F}_n}$, and the latter gives $\operatorname{vol}_{n-1} S_1^{n-1}$, since for any $\mathbf{x} \in S_1^{n-1}$ and any

$$\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1} \in T_{\vec{\mathbf{x}}} S_1^{n-1},$$

the vector $\vec{\mathbf{x}}$ is orthogonal to all the $\vec{\mathbf{v}}_i$, so

$$\begin{aligned}\Phi_{\vec{F}_n} P_{\vec{\mathbf{x}}}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}) &= \det[\vec{\mathbf{x}}, \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}] \\ &= \sqrt{\det[\vec{\mathbf{x}}, \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}]^\top [\vec{\mathbf{x}}, \vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}]} \\ &= \sqrt{\det[\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}]^\top [\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}]} \\ &= \text{vol}_{n-1} P_{\vec{\mathbf{x}}}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{n-1}).\end{aligned} \tag{1}$$

To see how we got from line 2 to line 3 in equation 1, look at equation 5.1.9 giving the product $T^\top T$. The entry in the first row, first column is $|\vec{\mathbf{x}}| = 1$, and all other entries of the first row and first column are 0.

6.33 Choosing one basis $\{\mathbf{u}\}$ of a real vector space and declaring it direct orients the vector space, according to the rule

$$\Omega(\{\mathbf{u}'\}) = \operatorname{sgn} \det P \quad \text{if} \quad \{\mathbf{u}'\} = \{\mathbf{u}\}P.$$

Solution 6.33: For the notation $\{\mathbf{u}\}P$ see the margin note next to Proposition and Definition 2.6.20.

Thus choosing one direct basis $\{\mathbf{v}\}$ for V (which is oriented, so this is meaningful) and one direct basis $\{\mathbf{w}\}$ for W, and declaring that the basis $(\{\mathbf{v}\}, \{\mathbf{w}\})$ of $V \times W$ is direct orients $V \times W$.

The thing we need to check is that if we had chosen some other direct bases $\{\mathbf{v}'\}, \{\mathbf{w}'\}$ for V and W we would get the same orientation of $V \times W$.

Since $\{\mathbf{v}'\}$ and $\{\mathbf{w}'\}$ are direct, the matrices P, Q such that $\{\mathbf{v}'\} = \{\mathbf{v}\}P$ and $\{\mathbf{w}'\} = \{\mathbf{w}\}Q$ both have positive determinant. Then

$$[\{\mathbf{v}'\}, \{\mathbf{w}'\}] = [\{\mathbf{v}\}, \{\mathbf{w}\}] \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix}$$

and $\det \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} = \det P \det Q = 1$. Thus $\{\mathbf{v}'\}, \{\mathbf{w}'\}$ is indeed direct.

Solutions for the Appendix

A1.1 Suppose x and y are k-close for all k and that $x < y$; moreover we can suppose that x and y are both positive. This means that for all k, either $[x]_{-k} = [y]_{-k}$ or $[y]_{-k} = [x]_{-k} + 10^{-k}$.

By "mth digit" we mean the mth digit to the right of the decimal point.

Since $x \neq y$, there must be a first digit in which they differ, say the mth. Saying that they are m-close means that the mth digit of x is one less than the mth digit of y; in particular, the mth digit of x is not a 9. What does saying that they are $(m+1)$-close tell us? Adding $10^{-(m+1)}$ to $[x]_{-(m+1)}$ can only produce a carry in the mth position if the $(m+1)$st digit of x is a 9 and the $(m+1)$st digit of y is a 0, and in that case they are $(m+1)$-close. Saying that they are $(m+2)$-close says that the $(m+2)$nd digit of x and y are respectively 9 and 0 as well.

Continuing this way, we see that all the digits of x after the mth are 9's, and all the digits of y after the mth are 0's; moreover, the mth digit of x is one less than the mth digit of y. That is exactly the condition for x and y to be equal.

A1.3 a. Let a and b be positive finite decimals. We want to consider a/b, as computed by long division, so we may assume (by multiplying both by an appropriate power of 10) that they are both integers. When writing the digits of the quotient after the decimal point, the next digit depends only on the current remainder, which is a number $< b$. Thus the same integer must appear twice, and the digits of the quotient between two such appearances will then be repeated indefinitely.

Suppose we carry out long division to k digits after the decimal point, obtaining a quotient q_k and a remainder r_k. These numbers then satisfy

$$a = q_k b + 10^{-k} r_k.$$

By equation A1.5, $b(a/b)$ is

$$\sup_k \inf_{l \geq k} b([a/b]_l) = \sup_k \inf_{l \geq k} bq_l = \sup_k bq_k.$$

A remainder is always smaller than the divisor, so $r_k < b$.

where the inf was dropped since the sequence $k \mapsto bq_k$ is increasing, so the inf is the first element. But we just saw that $a - q_k b = 10^{-k} r_k < 10^{-k} b$. Thus a is the least upper bound of the $q_k b$: it is an upper bound since $10^{-k} r_k \geq 0$, and it is a least upper bound since $10^{-k} r_k$ will be arbitrarily small for k large.

b. Suppose $x \in \mathbb{R}$ satisfies $x > 0$; then $x > 10^{-p}$ for some p, since it must have a first nonzero digit. Define

$$\operatorname{inv} x = \inf_{k > p} (1/[x]_k).$$

We need to show that $x \operatorname{inv}(x) = 1$. By equation A1.5, this product of real numbers means

$$x \operatorname{inv}(x) = \sup_k \inf_{l > k} [x]_l [\operatorname{inv}(1/x)]_l = \sup_k [x]_k [\operatorname{inv}(1/x)]_k,$$

since again the sequence $l \mapsto [x]_l[\mathrm{inv}(1/x)]_l$ is increasing. So we must show that

$$\sup_k [x]_k[\mathrm{inv}(1/x)]_k = 1.$$

Suppose $x < 10^n$, and choose an integer p. Since $\mathrm{inv}(x) = \inf_k 1/[x]_k$, there exists q such that $[\mathrm{inv}(x)]_k - 1/[x]_k] \le 10^{-(p+n)}$ when $k > q$. Then

$$\begin{aligned}\big|[x]_k[\mathrm{inv}(1/x)]_k - 1\big| &= \big|[x]_k[\mathrm{inv}(1/x)]_k - [x]_k(1/[x]_k)\big|\\ &= [x]_k\big|[\mathrm{inv}(1/x)]_k - (1/[x]_k)\big| \le x10^{-(p+n)} \le 10^{-p},\end{aligned}$$

when $k > q$. Thus $\sup_k [x]_k[\mathrm{inv}(1/x)]_k$ is p-close to 1 for every $p \in \mathbb{N}$, which means that it is 1.

c. If $x < 0$, define $\mathrm{inv}(x) = -\mathrm{inv}(-x)$. We need to check that for $x < 0$, we have $x\,\mathrm{inv}(x) = 1$, and there is nothing to it:

We are here assuming the rule of signs that

$$(-a)b = a(-b) = -(ab);$$

how do we know that? Since $(-a)b$ and $a(-b)$ are $\mathbb{D}$-continuous functions, and they agree on the finite decimals, they agree everywhere.

$$x\,\mathrm{inv}(x) = x\big(-\mathrm{inv}(-x)\big) = \underbrace{(-x)}_{\substack{+\text{ since}\\ x<0}}\ \mathrm{inv}\ \underbrace{(-x)}_{\substack{+\text{ since}\\ x<0}} = 1,$$

since $x\,\mathrm{inv}(x) = 1$ was proved for $x > 0$ in part b.

A1.5 This is more fun, after the deadly problems above.

a. Call $a \le b < c$ the lengths of the sides of original triangle, so that c is the length of the hypotenuse. Let $l_i(\underline{s}) = \big|\mathbf{x}_i(\underline{s}) - \mathbf{x}_{i+1}(\underline{s})\big|$. Then

$$\frac{l_i(\underline{s})}{l_{i-1}(\underline{s})} = \frac{a}{c} \text{ or } \frac{b}{c}$$

depending on whether s_i is 0 or 1 and i is even or odd. Indeed, drawing the segment from $\mathbf{x}_i(\underline{s})$ to $\mathbf{x}_{i+1}(\underline{s})$ is dropping the altitude to the hypotenuse of a triangle similar to the original triangle, with one side of length l_{i-1}. In the original triangle, the ratio of the height to a side is either a/c or b/c. In particular, the telescoping series

$$\mathbf{x}_0(\underline{s}) + \sum_{i=0}^{\infty} \mathbf{x}_{i+1}(\underline{s}) - \mathbf{x}_i(\underline{s})$$

is convergent, since the corresponding sequence of absolute values $\displaystyle\sum_{i=0}^{\infty} l_i$ is convergent.

b. Suppose the ith digit of $\underline{s}$ and $\underline{s}'$ is the first digit that is different. Then $\mathbf{x}_i(\underline{s}) = \mathbf{x}_i(\underline{s}')$, and after that you go either left-right-left-right- … , or right-left-right-left- … . Both paths lead back to $\mathbf{x}_{i-1}(\underline{s})$.

c. Suppose that $|t_1 - t_2| < 2^{-k}$. Then either

• the first k digits of t_1 and t_2 coincide; in that case the line from $\mathbf{x}_{k-1}(t_1)$ to $\mathbf{x}_k(t_2)$ is the altitude of a triangle in which both $\gamma(t_1)$ and $\gamma(t_2)$ lie. But the diameter of that triangle is at most $c(b/c)^k$.

or

• there exists $t_3 \in [t_1, t_2]$ that can be written in two ways, and the first k digits of t_1 coincide with the first k digits of one way of writing t_3, whereas the first k digits of t_2 coincide with the first k digits of the other way of writing t_3. By the same argument as above,

$$\left|\gamma(t_1) - \gamma(t_2)\right| \le \left|\gamma(t_1) - \gamma(t_3)\right| + \left|\gamma(t_3) - \gamma(t_2)\right| \le 2c\left(\frac{b}{c}\right)^k .$$

d. Take any point $\mathbf{x} \in T$. Draw the first altitude $[\mathbf{x}_0, \mathbf{x}_1]$, and see if $\mathbf{x}$ is to the left or to the right. Then in the "even/odd" table, look in the second row (since we are writing the first digit of t) to see whether the first digit should be 0 or 1. If $\mathbf{x}$ is on the altitude you just drew, then enter either a 0 or a 1 for the first digit. In any case, draw the second altitude $|\mathbf{x}_1, \mathbf{x}_2]$ according to the digit that was either imposed or chosen, and see whether $\mathbf{x}$ is to the left or right of this altitude t.

If either is the case, then the first row of the table tells you what the second digit should be. If $\mathbf{x}$ is on this second altitude, the second digit is free. Keep going this, using alternately the top and the bottom row of the table to decide on successive digits. This also creates a sequence of points $\mathbf{x}_1, \mathbf{x}_2, \ldots$ that corresponds to the string of digits chosen. This sequence converges to $\mathbf{x}$, since the successive segments $[\mathbf{x}_i, \mathbf{x}_{i+1}]$ have lengths that decrease at least like a geometric series with ratio b/c. The maximum k such that there exist $t_1 < \cdots < t_k$ with

$$\gamma(t_1) = \gamma(t_2) = \cdots = \gamma(t_k)$$

is 4, and even that happens only when a and b are chosen carefully (for instance, taking $a = b$). Usually, the maximum is 3, achieved at the feet of altitudes that are in the interior of T; you get 4 if such a point is a foot of two altitudes, one from each side, as shown in the figure in the margin.

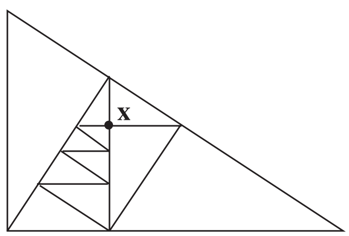


Figure for Solution A1.5. Suppose that the triangle T looks like the triangle above; more particularly, that the two horizontal lines meet exactly at the marked point $\mathbf{x}$. Then there exist four numbers $t_1 < t_2 < t_3 < t_4$ such that

$$\gamma(t_1) = \gamma(t_2) = \gamma(t_3) = \gamma(t_4).$$

It is quite instructive to figure out what they are for this specific example.

A2.1 The first step is to set $x = y + \frac{1}{3}$. The equation

$$x^3 - x^2 - x - 2 = 0 \quad \text{becomes} \quad y^3 - \frac{4}{3}y - \frac{65}{27} = 0.$$

The next step is to set

$$y = u - \frac{p}{3u} = u + \frac{4}{9u}.$$

Substituting this in the equation for y leads to

$$\left(u + \frac{4}{9u}\right)^3 - \frac{4}{3}\left(u + \frac{4}{9u}\right) - \frac{65}{27} = u^3 + \frac{4^3}{9^3 u^3} - \frac{65}{27} = 0.$$

Multiply by u^3 to find the quadratic equation for u^3

$$u^6 - -\frac{65}{27}u^3 + \frac{4^3}{27^2} = 0.$$

The discriminant of this quadratic equation is $3969/1729 > 0$, so the equation has exactly one real root, and it is simple. Solve the equation to find

$$u = \left(\frac{65 \pm \sqrt{3969}}{54}\right)^{1/3}.$$

This gives two expressions for x:

$$\left(\frac{65+\sqrt{3969}}{54}\right)^{1/3}+\frac{4}{9}\left(\frac{65+\sqrt{3969}}{54}\right)^{-1/3}+\frac{1}{3}=2 \quad \text{(surprise!)}$$

and

$$\left(\frac{65-\sqrt{3969}}{54}\right)^{1/3}+\frac{4}{9}\left(\frac{65-\sqrt{3969}}{54}\right)^{-1/3}+\frac{1}{3}=2. \quad \text{(another surprise!)}$$

It certainly isn't obvious that those funny expressions are equal to 2, or to each other, but it is true (try them on your calculator).

This is "explained" by the sentence following equation A2.13 and by Exercise A2.2.

The complex roots can be obtained from the complex cube roots:

$$\left(\frac{65+\sqrt{3969}}{54}\right)^{1/3}\left(\frac{-1\pm\sqrt{3}}{2}\right)+\frac{4}{9}\left(\frac{65+\sqrt{3969}}{54}\right)^{-1/3}\left(\frac{-1\mp\sqrt{3}}{2}\right)+\frac{1}{3}=\frac{-1+i\sqrt{3}}{2}.$$

We do not claim that this result is obvious, but a calculator will easily confirm it, and it could also be obtained by long division:

$$\frac{x^3-x^2-x-2}{x-2}=x^2+x+1,$$

whose roots are indeed the nonreal cubic roots of 1.

A2.3 We have

$$\begin{aligned}
&\left(y-\frac{a}{3}\right)^3+a\left(y-\frac{a}{3}\right)^2+b\left(y-\frac{a}{3}\right)+c=\left(y-\frac{a}{3}\right)\left(y^2-\frac{2ay}{3}+\frac{a^2}{9}\right)+ay^2-\frac{2a^2y}{3}+\frac{a^3}{9}+by-\frac{ab}{3}+c\\
&=y^3-\frac{2ay^2}{3}+\frac{a^2y}{9}-\frac{ay^2}{3}+\frac{2a^2y}{9}-\frac{a^3}{27}+ay^2-\frac{2a^2y}{3}+\frac{a^3}{9}+by-\frac{ab}{3}+c\\
&=y^3+\frac{a^2y}{9}+\frac{2a^2y}{9}-\frac{a^3}{27}-\frac{2a^2y}{3}+\frac{a^3}{9}+by-\frac{ab}{3}+c\\
&=y^3+y\left(b-\frac{a^2}{3}\right)+\frac{2a^3}{27}-\frac{ab}{3}+c.
\end{aligned}$$

A2.5 Here the appropriate substitution is $x=u+7/(3u)$, which leads to

$$\left(u+\frac{7}{3u}\right)^3-7\left(u+\frac{7}{3u}\right)+6=0.$$

As above, after multiplying out, simplifying, multiplying through by u^3, and setting $v=u^3$, you find the quadratic equation

$$v^2+6v+\frac{343}{27}=0.$$

This time, the roots of the quadratic are complex: $v_1=-3+i\frac{10\sqrt{3}}{9}$ and $v_2=-3-i\frac{10\sqrt{3}}{9}$. One approach (usually the only one) is to pass to polar coordinates; in this case, you might "just happen to observe" that the cube roots of $-3 + i\frac{10\sqrt{3}}{9}$ are

$$u_1 = 1 + i\frac{2\sqrt{3}}{3} \qquad u_2 = -\frac{3}{2} + i\frac{\sqrt{3}}{6} \quad \text{and} \quad u_3 = \frac{1}{2} - i\frac{5\sqrt{3}}{6}.$$

These lead to

$$u_1 + \frac{7}{3u_1} = 2, \quad u_2 + \frac{7}{3u_2} = -3 \quad \text{and} \quad u_3 + \frac{7}{3u_3} = 1.$$

Indeed, it is easy to check that $(x+3)(x-1)(x-2) = x^3 - 7x + 6$.

A2.7 a. This is a straightforward computation:

$$\left(x - \frac{a}{4}\right)^4 + a\left(x - \frac{a}{4}\right)^3 + b\left(x - \frac{a}{4}\right)^2 + c\left(x - \frac{a}{4}\right) + d$$

$$\begin{array}{rlllll}
= & x^4 & -ax^3 & +\frac{3}{8}a^2x^2 & -\frac{1}{16}a^3x & +\frac{1}{256}a^4 \\
 & & +ax^3 & -\frac{3}{4}a^2x^2 & +\frac{3}{16}a^3x & -\frac{1}{64}a^4 \\
 & & & +bx^2 & -\frac{1}{2}abx & +\frac{1}{16}ba^2 \\
 & & & & +cx & -\frac{1}{4}ac \\
 & & & & & +d \\
= & x^4 & & +(-\frac{3}{8}a^2+b)x^2 & +(\frac{a^3}{8} - \frac{ab}{2} + c)x & +(-\frac{3a^4}{256} + \frac{ba^2}{16} - \frac{ac}{4} + d)
\end{array}$$

So we see that

$$\begin{aligned}
p &= -\frac{3}{8}a^2 + b \\
q &= \frac{a^3}{8} - \frac{ab}{2} + c \\
r &= -\frac{3a^4}{256} + \frac{ba^2}{16} - \frac{ac}{4} + d.
\end{aligned}$$

b. The equation $x^2 - y + p/2 = 0$ is the definition of y, and if you substitute it into the second, you get

$$\left(x^2 + \frac{p}{2}\right)^2 + qx + r - \frac{p^2}{4} = x^4 + px^2 + qx + r = 0.$$

c. When $m = 1$, the curve is a circle; when $m < 0$, it is a hyperbola; when $m > 0$, it is an ellipse.

d. It is exactly the curve $y^2 + qx + r - \frac{p^2}{4} = 0$, which corresponds to $m = \infty$.

e. and f. If the equation $f_m\begin{pmatrix} x \\ y \end{pmatrix} = 0$ is a union of two lines, then at the point $\begin{pmatrix} x \\ y \end{pmatrix}$ where the lines intersect it cannot represent x implicitly as a function of y or y implicitly as a function of x. (We discuss this sort of thing in detail in Section 2.10.) It follows that both partial derivatives of

f_m must vanish at that point, so m and $\begin{pmatrix} x \\ y \end{pmatrix}$ satisfy the equations

$$f_m \begin{pmatrix} x \\ y \end{pmatrix} = y^2 + qx + r - \frac{p^2}{4} + m\left(x^2 - y + \frac{p}{2}\right) = 0$$
$$2y - m = 0$$
$$q + 2mx = 0.$$

This gives

$$y = \frac{m}{2} \quad \text{and} \quad x = -\frac{q}{2m},$$

and substituting these values into the first equation, multiplying through by $-4m$, and collecting terms gives

$$m^3 - 2pm^2 + (p^2 - 4r)m + q^2 = 0.$$

g. The two functions $y^2 + qx + r - \frac{p^2}{4}$ and $x^2 - y + \frac{p}{2}$, when restricted to a diagonal l_k, are two quadratic functions that vanish at the same two points. But any quadratic function of t that vanishes at two points a and b is of the form $A(t-a)(t-b)$. In other words, any two such functions are multiples of each other. Thus we have

$$\big(k(x - x_1) + y_1\big)^2 + qx + r - \frac{p^2}{4} = A\left(x^2 - k(x - x_1) - y_1 + \frac{p}{2}\right)$$

for some number A. It follows that the ratios of the coefficients of x^2, x, and constants must be equal, which leads to the equations given in the exercise. Use the first and second, and the first and third, to express k^3 as a quadratic polynomial in k, and equate these polynomials to get

"The ratios ... must be equal": If

$$\alpha_1 x^2 + \beta_1 x + \gamma_1 = A(\alpha_2 x^2 + \beta_2 x + \gamma_2),$$

then

$$\frac{\alpha_1}{\alpha_2} = \frac{\beta_1}{\beta_2} = \frac{\gamma_1}{\gamma_2}.$$

$$k^2\left(x_1^2 - y_1 + \frac{p}{2}\right) = y_1^2 + qx_1 - \frac{p^2}{4} + r.$$

h. The two parabolas to intersect are $y = x^2 - 2$ and $x = 3 - y^2$, represented in the figure below.

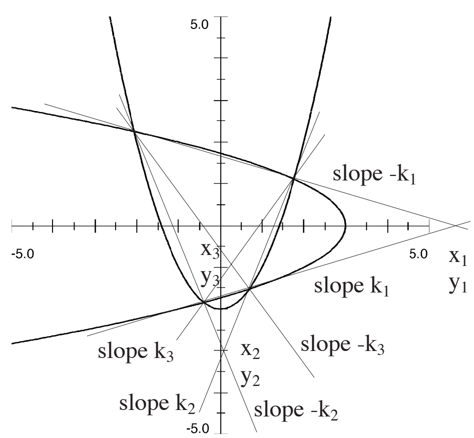


FIGURE FOR SOLUTION A2.7, part h. The parabolas to intersect in order to solve $x^4 - 4x^2 + x + 1 = 0$.

The resolvent cubic is $m^3 + 8m^2 + 12m + 1 = 0$. If we set $m = t - 8/3$ to eliminate the square term, we find the equation

$$t^3 - \frac{28}{3}t + \frac{187}{27} = t^3 + \alpha t + \beta = 0.$$

This is best solved by the trigonometric formula of Exercise A2.6, which gives

$$t = \sqrt{-\frac{4\alpha}{3}} \cos\left(\frac{1}{3}\arccos\left(\sqrt{-\frac{3}{4\alpha}}\frac{3\beta}{\alpha}\right)\right).$$

With our values of α and β, this gives the roots

$$t_1 \approx 2.57816998074169, \quad t_2 \approx -3.37429792821080, \quad t_3 \approx 0.79612794746911.$$

Since $m = t - 8/3$, this gives

$$m_1 \approx -0.08849668592498, \quad m_2 \approx -6.04096459487747, \quad m_3 \approx -1.87053871919756.$$

The corresponding points are

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \approx \begin{pmatrix} 5.64992908800994 \\ -0.04424834296249 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \approx \begin{pmatrix} 0.08276823877167 \\ -3.02048229743873 \end{pmatrix}, \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} \approx \begin{pmatrix} 0.26730267321838 \\ -0.93526935959878 \end{pmatrix}.$$

If we insert these values into the equation for k, we find

$$k_1 \approx 0.29748392549006, \quad k_2 \approx 2.45783738169910, \quad k_3 \approx 1.36767639418013.$$

To find the roots, we must intersect the lines given by the equations $y - y_i = \pm k_i(x - x_i)$, two at a time, of course.

For instance, if we intersect $y - y_1 = -k_1(x - x_1)$ and $y - y_2 = k_2(x - x_2)$, we find the point

$$\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix} \approx \begin{pmatrix} 1.76401492519458 \\ 1.11174865630925 \end{pmatrix}.$$

If we intersect $y - y_1 = -k_1(x - x_1)$ and $y - y_3 = -k_3(x - x_3)$, we find

$$\begin{pmatrix} X_2 \\ Y_2 \end{pmatrix} \approx \begin{pmatrix} -2.06149885068464 \\ 2.24977751137410 \end{pmatrix}.$$

If we intersect $y - y_1 = k_1(x - x_1)$ and $y - y_3 = k_3(x - x_3)$, we find

$$\begin{pmatrix} X_3 \\ Y_3 \end{pmatrix} \approx \begin{pmatrix} -0.39633853101445 \\ -1.84291576883330 \end{pmatrix}.$$

Finally, if we intersect $y - y_1 = k_1(x - x_1)$ and $y - y_2 = k_2(x - x_2)$, we find

$$\begin{pmatrix} X_4 \\ Y_4 \end{pmatrix} \approx \begin{pmatrix} 0.69382245650451 \\ -1.51861039885004 \end{pmatrix}.$$

The points are labeled by the quadrant they are in.

Indeed, the four numbers X_1, X_2, X_3, X_4 are (approximations to) the roots of the original polynomial.

Now let us solve the equation $x^4 + 4x^3 + x - 1 = 0$. If we set $x = y - 1$, this becomes $y^4 - 6y^2 + 9y - 5 = 0$, with resolvent cubic

$$m^3 + 12m^2 + 56m + 81 = 0.$$

To solve the resolvent cubic, we need to eliminate the term in m^2, by setting $m = t - 4$, giving the equation

$$t^3 + 8t + 15 = t^3 + Pt + Q = 0.$$

Then set $t = u - P/(3u)$ and $v = u^3$; as in equation A2.11, this leads to the quadratic equation

$$v^2 + Qv - \frac{P^3}{27} = 0. \tag{1}$$

One root of this equation is

$$v \approx 16.17254074438183,$$

and the three cube roots of v are

$$\begin{aligned} u_1 &\approx 2.52886755571821, \\ u_2 &\approx -1.26443377785910 + 2.19006354605823\,i, \\ u_3 &\approx -1.26443377785910 - 2.19006354605823\,i. \end{aligned}$$

(If we choose the other root of equation 1, we would get different values of u_i but the same values of t_i; see the text immediately following equation A2.13).

These now give

$$\begin{aligned} t_1 &\approx 1.47437711368740, \\ t_2 &\approx -0.73718855684370 + 3.10327905690479i, \\ t_3 &\approx -0.73718855684370 - 3.10327905690479i, \end{aligned}$$

and finally, using $m = t - 4$, we get

$$\begin{aligned} m_1 &\approx -2.52562288631260, \\ m_2 &\approx -4.73718855684370 + 3.10327905690479\,i, \\ m_3 &\approx -4.73718855684370 - 3.10327905690479\,i. \end{aligned}$$

We have now solved the resolvent cubic, and can give the points of intersection of the diagonals of the quadrilateral formed by the roots:

$$\begin{aligned} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} &\approx \begin{pmatrix} 1.78173868489526 \\ -1.26281144315630 \end{pmatrix}, \\ \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} &\approx \begin{pmatrix} 0.66468621310792 + 0.43542847826296i \\ -2.36859427842185 + 1.55163952845240i \end{pmatrix}, \\ \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} &\approx \begin{pmatrix} 0.66468621310792 - 0.43542847826296i \\ -2.36859427842185 - 1.55163952845240i \end{pmatrix}. \end{aligned}$$

We can now compute the slopes of two of the pairs of diagonals, given by

$$\begin{aligned} k_{1,1} &\approx 1.58922084252397, \\ k_{1,2} &= -k_{1,1}, \\ k_{2,1} &\approx 2.28038824159147 - 0.68042778863371\, i, \\ k_{2,2} &= -k_{2,1}. \end{aligned}$$

We are finally in a position to intersect the pairs of diagonals; the x-coordinate of each point of intersection is given by the formula

$$\frac{k_{1,i}\, x_1 - k_{2,j} x_2 - y_1 + y_2}{k_{1,i} - k_{2,j}}, \quad \text{for appropriate values of } i \text{ and } j;$$

the four values given by $i = 1, 2,\ j = 1, 2$ give the four roots of the equation

$$y^4 - 6y^2 + 9y - 5 = 0.$$

They are

$$\begin{aligned} &0.79461042126199 + 0.68042778863371i, \\ &0.79461042126199 - 0.68042778863371i, \\ &1.48577782032949, \\ -&\,3.07499866285346, \end{aligned}$$

and finally, the roots of the original equation

$$x^4 + 4x^3 + x - 1 = 0$$

are found from $x = y - 1$ to be

$$\begin{aligned} &-0.20538957873801 + 0.68042778863371i, \\ &-0.20538957873801 - 0.68042778863371i, \\ &0.48577782032949, \\ &-4.07499866285346. \end{aligned}$$

A5.1 First note that

$$\begin{aligned} \Big|\mathbf{f}(\mathbf{u}+\mathbf{h}) - \mathbf{f}(\mathbf{u}) - [\mathbf{Df}(\mathbf{u})]\mathbf{h}\Big| &= \left| \int_0^1 \Big([\mathbf{Df}(\mathbf{u}+t\mathbf{h})]\mathbf{h} - [\mathbf{Df}(\mathbf{u})]\mathbf{h} \Big) dt \right| \\ &\le \int_0^1 M|t\mathbf{h}|^\alpha\, |\mathbf{h}|\, dt = \frac{M}{\alpha+1}|\mathbf{h}|^{\alpha+1}. \end{aligned}$$

To show that $[\mathbf{Df}(\mathbf{a}_1)]$ is invertible, we will use the following lemma:

Lemma 1. *If P and Q are $n \times n$ matrices with Q invertible, and*

$$|I - PQ^{-1}| \le k < 1,$$

then P is invertible, and

$$|P^{-1}| \le \frac{1}{1-k}|Q^{-1}|.$$

Proof. Set $X = I - PQ^{-1}$; then $I - X = PQ^{-1}$ is invertible by geometric series. Thus P^{-1} exists, and

$$P^{-1} = Q^{-1}(I + X + X^2 + \cdots),$$

so

$$\begin{aligned}|P^{-1}| &\le |Q^{-1}| + |Q^{-1}X| + \cdots \le |Q^{-1}|(1 + |X| + \cdots) \\ &= \frac{|Q^{-1}|}{1-|X|} \le \frac{1}{1-k}|Q^{-1}|. \quad \square\end{aligned}$$

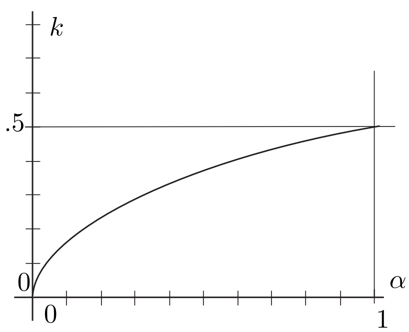
To apply this to show that $[\mathbf{Df}(\mathbf{a}_1)]$ is invertible, note that

$$\begin{aligned}\left|I - [\mathbf{Df}(\mathbf{a}_1)][\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| &= \left|\Big([\mathbf{Df}(\mathbf{a}_0)] - [\mathbf{Df}(\mathbf{a}_1)]\Big)[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \\ &\le M\,|\mathbf{h}_0|^\alpha\,\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \\ &\le M\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\mathbf{f}(\mathbf{a}_0)\right|^\alpha\,\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \\ &\le k.\end{aligned}$$

The quantity on the third line is $\le k$ by hypothesis.

Moreover, note that k as defined in the exercise is necessarily strictly less than 1; in fact, it is less than 1/2; see the figure in the margin. So Lemma 1 says that $[\mathbf{Df}(\mathbf{a}_1)]^{-1}$ exists and

$$[\mathbf{Df}(\mathbf{a}_1)]^{-1} \le \frac{1}{1-k}[\mathbf{Df}(\mathbf{a}_0)]^{-1}.$$



The equation

$$\left(\frac{k}{(\alpha+1)(1-k)}\right)^\alpha = 1-k$$

represents k implicitly as a function of α for $0 \le \alpha \le 1$ (and in fact for all $\alpha \ge 0$). By implicit differentiation, one can show that this function is increasing; more specifically, it increases from 0 to 1/2 as α increases from 0 to 1.

It follows that we can define

$$\mathbf{h}_1 = -[\mathbf{Df}(\mathbf{a}_1)]^{-1}\mathbf{f}(\mathbf{a}_1) \quad \text{and} \quad \mathbf{a}_2 = \mathbf{a}_1 + \mathbf{h}_1.$$

We need three estimates for the size of $\mathbf{f}(\mathbf{a}_1)$:

$$|\mathbf{f}(\mathbf{a}_1)| = \left|\mathbf{f}(\mathbf{a}_1) - \mathbf{f}(\mathbf{a}_0) - [\mathbf{Df}(\mathbf{a}_0)]\mathbf{h}_0\right| \le \begin{cases} \dfrac{M}{\alpha+1}|\mathbf{h}_0|^{\alpha+1} \\[2ex] \left|\dfrac{M}{\alpha+1}[\mathbf{Df}(\mathbf{a}_0^{-1})]\mathbf{f}(\mathbf{a}_0)\right|^{\alpha+1} \\[2ex] \dfrac{M}{\alpha+1}\left|[\mathbf{Df}(\mathbf{a}_0^{-1})]\mathbf{f}(\mathbf{a}_0)\right|^\alpha |\mathbf{h}_0|, \end{cases}$$

where the three lines on the right after the inequality are all equal.

We find the size of $\mathbf{h}_1$:

$$\begin{aligned}|\mathbf{h}_1| &= \left|[\mathbf{Df}(\mathbf{a}_1)]^{-1}\right|\,|\mathbf{f}(\mathbf{a}_1)| \\ &\le \frac{[\mathbf{Df}(\mathbf{a}_0)]^{-1}}{1-k}\,\frac{M}{\alpha+1}\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\mathbf{f}(\mathbf{a}_0)\right|^\alpha\,|\mathbf{h}_0| \\ &\le \frac{k}{1-k}\,\frac{|\mathbf{h}_0|}{\alpha+1}.\end{aligned}$$

Next we show that the modified Kantorovich inequality

$$\left|[\mathbf{D}\vec{\mathbf{f}}(\mathbf{a}_0)]^{-1}\right|^{\alpha+1}|\vec{\mathbf{f}}(\mathbf{a}_0)|^\alpha M \le k$$

holds for $\mathbf{a}_1$:

$$\begin{aligned}\left|[\mathbf{Df}(\mathbf{a}_1)]^{-1}\right|^{1+\alpha}|\mathbf{f}(\mathbf{a}_1)|^{\alpha}M &\le \left(\frac{|[\mathbf{Df}(\mathbf{a}_0)]^{-1}|}{1-k}\right)^{1+\alpha}\left(\frac{M}{\alpha+1}\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\mathbf{f}(\mathbf{a}_0)\right|^{\alpha+1}\right)^{\alpha}M\\ &\le \left(\frac{|[\mathbf{Df}(\mathbf{a}_0)]^{-1}|}{1-k}\right)^{1+\alpha}\left(\frac{k|\mathbf{f}(\mathbf{a}_0)|}{\alpha+1}\right)^{\alpha}M\\ &\le \left(\frac{1}{1-k}\right)^{1+\alpha}\left(\frac{k}{\alpha+1}\right)^{\alpha}k\\ &\le \frac{1}{1-k}\left(\frac{k}{(\alpha+1)(1-k)}\right)^{\alpha}k \le \frac{1-k}{1-k}k = k.\end{aligned}$$

So we have shown that the method may be iterated; that

$$|\mathbf{h}_{m+1}| \le \frac{k}{(1-k)(\alpha+1)}|\mathbf{h}_m|,$$

so $\lim \mathbf{a}_j = \mathbf{a}_0 + \sum \mathbf{h}_j$ converges to a point $\mathbf{a}$; and that

$$|\mathbf{f}(\mathbf{a}_{m+1})| \le \frac{M}{\alpha+1}|\mathbf{h}_m|^{\alpha+1},$$

so $\mathbf{f}(\mathbf{a}) = \mathbf{0}$.

Finally, we show the rate of convergence:

$$1 - [\mathbf{Df}(\mathbf{a}_{m+1})][\mathbf{Df}(\mathbf{a}_0)]^{-1} = \sum_{j=0}^{m}\Big([\mathbf{Df}(\mathbf{a}_j)] - [\mathbf{Df}(\mathbf{a}_{j+1})]\Big)[\mathbf{Df}(\mathbf{a}_0)]^{-1},$$

so

$$\begin{aligned}\left|1 - [\mathbf{Df}(\mathbf{a}_{m+1})][\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| &\le \sum_{j=0}^{m} M|\mathbf{h}_j|^{\alpha}\,\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \le \frac{M|\mathbf{h}_0|^{\alpha}\,\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right|}{1-\left(\frac{k}{(1-k)}\frac{1}{(\alpha+1)}\right)^{\alpha}}\\ &= \frac{M\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\mathbf{f}(\mathbf{a}_0)\right|^{\alpha}\left|[\mathbf{Df}(-\mathbf{a}_0)]^{-1}\right|}{1-\left(\frac{k}{(1-k)(\alpha+1)}\right)^{\alpha}}\\ &\le \frac{k}{1-\left(\frac{k}{(1-k)(\alpha+1)}\right)^{\alpha}} < \frac{k}{1-(1-k)} = 1.\end{aligned}$$

So, again by Lemma 1,

$$\left|[\mathbf{Df}(\mathbf{a}_{m+1})]^{-1}\right| \le \frac{1}{1-\dfrac{k}{1-\left(\dfrac{k}{(1-k)(\alpha+1)}\right)^{\alpha}}} \le C\frac{M}{\alpha+1}|\mathbf{h}_m|^{\alpha+1}.$$

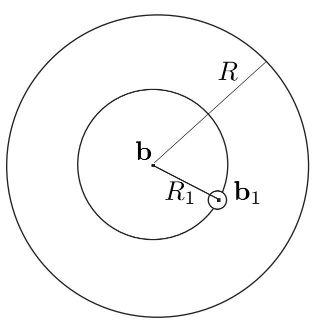


FIGURE FOR SOLUTION A8.1.

Since R_1 is the largest number such that $\mathbf{g}$ and $\mathbf{g}_1$ coincide on $B_{R_1}(\mathbf{b})$, and $|\mathbf{b}_1 - \mathbf{b}| = R_1$, we have $\mathbf{g}_1(\mathbf{b}_1) = \mathbf{g}(\mathbf{b}_1)$, but any neighborhood of $\mathbf{b}_1$ includes points in $B_R(\mathbf{b}) - B_{R_1}(\mathbf{b})$, where $\mathbf{g}$ and $\mathbf{g}_1$ do not coincide.

A8.1 Suppose there were two such continuous functions

$$\mathbf{g}, \mathbf{g}_1 : B_R(\mathbf{b}) \to \mathbb{R}^n \quad \text{with} \quad \mathbf{g}(\mathbf{b}) = \mathbf{g}_1(\mathbf{b}) = \mathbf{a}.$$

Let R_1 be the largest number such that $\mathbf{g}$ and $\mathbf{g}_1$ coincide on $B_{R_1}(\mathbf{b})$. If $R_1 = R$, we are done, so we may assume that $R_1 < R$. As shown in the picture in the margin, let $\mathbf{b}_1$ be a point with $|\mathbf{b}_1 - \mathbf{b}| = R_1$ and such that $\mathbf{g}$

and $\mathbf{g}_1$ do not coincide on any neighborhood of $\mathbf{b}_1$; however, we must have $\mathbf{g}(\mathbf{b}_1) = \mathbf{g}_1(\mathbf{b}_1)$ by the continuity of $\mathbf{g}_1$; call $\mathbf{a}_1 = \mathbf{g}(\mathbf{b}_1)$, and $\mathbf{c}_1 = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{b}_1 \end{pmatrix}$.

Our proof of Theorem 2.10.7 (the implicit function theorem) shows that at $\mathbf{b}_1$, the matrix

$$[D_1F(\mathbf{c}_1), \dots, D_nF(\mathbf{c}_1)]$$

is invertible, hence near $\mathbf{c}_1$, the set

$$\left\{ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \;\middle|\; F\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{0} \right\}$$

is the graph of a function expressing $\mathbf{x}$ as a function of $\mathbf{y}$. Thus there exist neighborhoods U_1 of $\mathbf{a}_1$ and V_1 of $\mathbf{b}_1$ such that for every $\mathbf{y} \in V_1$, there is a *unique* $\mathbf{x} \in U_1$ such that $F\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{0}$. So this point $\mathbf{x}$ must be the value $\mathbf{g}(\mathbf{y})$ and the value $\mathbf{g}_1(\mathbf{y})$, meaning that $\mathbf{g}$ and $\mathbf{g}_1$ coincide on some neighborhood of $\mathbf{b}_1$. This contradicts the hypothesis that $R_1 < R$.

A9.1 The final equality in equation A9.7 is justified by the following computation:

$$\begin{aligned}
\lim_{t\to 0} \frac{g'(c_t)}{t} &= \lim_{t\to 0} \frac{D_if(\mathbf{a} + c_t\vec{\mathbf{e}}_i + t\vec{\mathbf{e}}_j) - D_if(\mathbf{a} + c_t\vec{\mathbf{e}}_i)}{t} \\
&= \lim_{t\to 0} \frac{D_if(\mathbf{a} + c_t\vec{\mathbf{e}}_i + t\vec{\mathbf{e}}_j) - D_if(\mathbf{a} + c_t\vec{\mathbf{e}}_i)}{t} \\
&\quad - \overbrace{\left(\lim_{t\to 0} \frac{D_if(\mathbf{a} + c_t\vec{\mathbf{e}}_i + t\vec{\mathbf{e}}_j) - D_if(\mathbf{a}) - c_tD_i^2f(\mathbf{a}) - tD_jD_if(\mathbf{a})}{t}\right)}^{0 \text{ by equation A9.5}} \\
&\quad + \overbrace{\left(\lim_{t\to 0} \frac{D_if(\mathbf{a} + c_t\vec{\mathbf{e}}_i) - D_if(\mathbf{a}) - c_tD_i^2f(\mathbf{a})}{t}\right)}^{0 \text{ by equation A9.6}} \qquad (1) \\
&= \lim_{t\to 0} \frac{D_if(\mathbf{a}) + c_tD_i^2f(\mathbf{a}) + tD_jD_if(\mathbf{a}) - D_if(\mathbf{a}) - c_tD_i^2f(\mathbf{a})}{t} \\
&= D_jD_if(\mathbf{a}).
\end{aligned}$$

Solution A9.1: This argument would not work if we required only that all second partial derivatives of f exist, instead of requiring that the partials be differentiable at $\mathbf{a}$. If D_if were a nondifferentiable function with all partial derivatives existing (like the function of Example 1.9.3), then we could still write the Jacobian matrix of partial derivatives

$$[D_1D_if(\mathbf{a}) \dots D_nD_if(\mathbf{a})],$$

but we could not substitute it for the derivative L in equation 2.

The most delicate point is that the third line equals 0; this was explained briefly in the textbook. The first line of equation A9.5:

$$0 = \lim_{\sqrt{c_t^2+t^2}\to 0} \frac{1}{\sqrt{c_t^2 + t^2}} \Big(D_if(\mathbf{a}+c_t\mathbf{e}_i+t\mathbf{e}_j) - D_if(\mathbf{a}) - [\mathbf{D}D_if(\mathbf{a})](c_t\vec{\mathbf{e}}_i+t\vec{\mathbf{e}}_j) \Big)$$

We can replace the $\sqrt{c_t^2 + t^2}$ in the denominator of the first line of equation A9.5 by $|t|$ in the denominator of the second line because c_t is between 0 and t, so $\sqrt{c_t^2 + t^2} < \sqrt{2}|t|$, which goes to 0 with t.

is the equation

$$\lim_{\vec{\mathbf{h}}\to \mathbf{0}} \frac{1}{|\vec{\mathbf{h}}|} \Big(\big(\mathbf{f}(\mathbf{a} + \vec{\mathbf{h}}) - \mathbf{f}(\mathbf{a})\big) - \big(L(\vec{\mathbf{h}})\big) \Big) = \mathbf{0} \qquad (2)$$

(see Theorem 1.7.10), where $\vec{\mathbf{h}}$ is replaced by the vector $c_t\vec{\mathbf{e}}_i + t\vec{\mathbf{e}}_j$, the function $\mathbf{f}$ is replaced by the function D_if (differentiable by hypothesis), and L is the derivative of D_if, i.e., the Jacobian matrix $[\mathbf{D}D_if(\mathbf{a})]$ made of the partial derivatives of $D_i f$. Because the derivative is linear, we can write

Equation 3: The ith column of $[\mathbf{D}D_i f(\mathbf{a})]$ is $[\mathbf{D}D_i f(\mathbf{a})](\vec{\mathbf{e}}_i)$, by Theorem 1.3.4. Or note simply the definition of the Jacobian matrix, Definition 1.7.7.

$$\begin{aligned}[\mathbf{D}D_i f(\mathbf{a})](c_t\vec{\mathbf{e}}_i + t\vec{\mathbf{e}}_j) &= c_t[\mathbf{D}D_i f(\mathbf{a})]\vec{\mathbf{e}}_i + t[\mathbf{D}D_i f(\mathbf{a})]\vec{\mathbf{e}}_j \\ &= c_t(D_i^2 f)(\mathbf{a}) + tD_j(D_i f)(\mathbf{a}).\end{aligned} \tag{3}$$

A similar but easier computation justifies the 0 above the brackets in the fourth line in equation 1.

A10.1 We give two solutions: the solution to Exercise A10.1 in the first printing of the 5th edition of the text, and the solution to Exercise A10.1 in the second printing.

Solution to Exercise A10.1 in the first printing

It is actually easier to start the induction at $k = 0$: the beginning of the induction is then simply that if f is continuous at $\mathbf{a}$ and $f(\mathbf{a}) = 0$, then

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} f(\mathbf{a} + \vec{\mathbf{h}}) = 0,$$

which is the definition of continuity.

The inductive step as written is just as valid to go from $k = 0$ to $k = 1$ as for any other values; more particularly, in equations A10.6–A10.8, we use the fact that if f is of class C^1 with derivatives vanishing at $\mathbf{a}$, then

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} D_i f(\mathbf{a} + \vec{\mathbf{c}}) = 0,$$

which is the definition of continuity for the derivatives.

Why does that imply Theorem 1.9.8? There the issue is that if the first partials of f are continuous, then

$$\lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{1}{|\vec{\mathbf{h}}|}\left(f(\mathbf{a} + \vec{\mathbf{h}}) - f(\mathbf{a}) - \sum_{i=1}^n h_i D_i(f)(\mathbf{a})\right) = 0.$$

The entire expression $f(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a}) - \sum_{i=1}^n h_i D_i(f)(\mathbf{a})$ is a function whose 0th and first partials are continuous and vanish at $\mathbf{a}$, so Proposition 3.3.17 applies (and is now proved without using Theorem 1.9.8, so the proof isn't circular) and shows that the limit is 0.

Solution to Exercise A10.1 in the second printing

Set $h_i(t) \overset{\text{def}}{=} D_i g(t\mathbf{x})$, so $f'_{\mathbf{x}}(t) = \sum_{i=1}^n h_i(t\mathbf{x})x_i$ by the chain rule. We will work by induction on k; since h_i is of class C^{k-1} on U we may assume that the result is true for each h_i, i.e.,

$$h_j^{(i-1)}(0) = \sum_{J\in\mathcal{I}_n^{i-1}} \frac{(i-1)!}{J!} D_J D_j g(\mathbf{0})\, \mathbf{x}^J$$

for all $i \le k$, so $i-1 \le k-1$. Write

$$\begin{aligned} f_{\mathbf{x}}^{(i)}(0) &= (f'_{\mathbf{x}})^{(i-1)}(0) = \left(\sum_{j=1}^{n} x_j h_j\right)^{(i-1)}(0) \\ &= \sum_{j=1}^{n} \left(h_j^{(i-1)}(0)\right) x_j \quad (x_j \text{ is a constant}) \\ &= \sum_{j=1}^{n} \Big(\sum_{J \in \mathcal{I}_n^{i-1}} \frac{(i-1)!}{J!} D_J D_j g(\mathbf{0})\, \mathbf{x}^J \Big) x_j \\ &= \sum_{I \in \mathcal{I}_n^i} \frac{i!}{I!} D_I g(0)\, \mathbf{x}^I. \end{aligned}$$

The fourth equality is induction, but the fifth is a more delicate combinatorial computation. Choose some $I = (i_1, \dots, i_n) \in \mathcal{I}_n^i$. Set

$$\begin{aligned} J_1 &\overset{\text{def}}{=} (i_1 - 1, i_2, \dots, i_n) \\ J_2 &\overset{\text{def}}{=} (i_1, i_2 - 1, \dots, i_n), \dots \\ &\vdots \\ J_n &\overset{\text{def}}{=} (i_1, i_2, \dots, i_n - 1) \end{aligned}$$

Then J_p contributes to $\mathbf{x}^I$ if $i_p \neq 0$, in the sense that $\mathbf{x}^{J_p} x_j = \mathbf{x}^I$. In that case the multi-index J_p contributes

$$\frac{(i-1)!}{J_p!} = \frac{(i-1)!\, i_p}{I!},$$

so that together all the J_p contribute

$$\frac{(i-1)!}{I!}(i_1 + \dots + i_n) = \frac{i!}{I!}.$$

A11.1

a. The function $|x|^{3/2}$ is in $o(|x|)$, because

$$\lim_{x \to 0} \frac{|x|^{3/2}}{|x|} = \lim_{x \to 0} \frac{1}{|x|1/2} = 0.$$

b. It is also in $O(|x|)$, because it is in $o(|x|)$; see part a.

c. The function $|x|^{3/2}$ is not in $o(|x|^2)$ because it isn't even in $O(|x|^2)$; see part d.

Recall (equation A11.1) that

$$g \in o(f) \implies g \in O(f).$$

d. The function $|x|^{3/2}$ is not in $O(|x|^2)$, because

$$\lim_{x \to 0} \frac{|x|^{3/2}}{|x|^2} = \lim_{x \to 0} \frac{1}{|x|^{1/2}} = \infty.$$

A11.3 The function $x \log |x|$ is not in $O(x)$; indeed,

$$\lim_{x \to 0} \frac{x \log |x|}{x} = \lim_{x \to 0} \log |x| = -\infty.$$

So in particular, it is not in $o(x)$, or in $O(x^2)$, or in $o(x^2)$ either.

A12.1 a. This follows immediately from Corollary A12.3. We have (Proposition 3.4.2)

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^k}{k!} + \frac{e^c}{(k+1)!}x^{k+1}$$

for some $0 \le c \le x$. Apply this when $x = 1$, so that $e^c < 3$.

b. Multiply the inequality

$$\left|\frac{a}{b} - \Big(1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{k!}\Big)\right| \le \frac{3}{(k+1)!}$$

by $k!b$ to get

$$\left|k!a - b\Big(k! + k! + \frac{k!}{2!} + \cdots + \frac{k!}{k!}\Big)\right| \le \frac{3k!b}{(k+1)!} = \frac{3b}{k+1}.$$

Since $k!a - bm$ is an integer whose absolute value is arbitrarily small, it is 0.

c. As the hint indicates, if you take k to be a prime number $> b$, you have a contradiction: k divides $k!a$ evenly, of course, but it divides neither b nor m. Thus e is not rational.

A12.3 a. Using Theorem A12.1 (and equation 3.4.6), and setting $k = 2$, we get

$$\sin h = h + \frac{1}{2!}\int_0^h (h-t)^2(-\cos t)\,dt$$

(where $-\cos t$ is the third derivative of $\sin t$). Thus

$$\sin xy = xy + \frac{1}{2!}\int_0^{xy} (xy-t)^2(-\cos t)\,dt.$$

b. Using part a, the bound is

$$\begin{aligned}\left|\frac{1}{2!}\int_0^{xy}(xy-t)^2\cos t\,dt\right| &\le \frac{1}{2!}\int_0^{xy}(xy-t)^2\,dt\\ &= \frac{1}{2!}\left[-\frac{(xy-t)^3}{3}\right]_0^{xy} = \frac{1}{6}(xy)^3\\ &\le \frac{1}{6}\left(\frac{x^2+y^2}{2}\right)^3 \le \frac{1}{6}\left(\frac{1}{8}\right)^3.\end{aligned}$$

Solution A12.3, part b: The first inequality in the third line uses

$$0 \le (x-y)^2 = x^2 - 2xy + y^2,$$

hence $xy \le (x^2+y^2)/2$. The second inequality uses our assumption that $x^2 + y^2 \le 1/4$.

A12.5 a. Clearly $\operatorname{sgn} y$ is differentiable when $y \ne 0$. The chain rule guarantees that f is continuously differentiable unless $y = 0$ or the quantity under the square root is ≤ 0. But $-x + \sqrt{x^2+y^2} \ge 0$ and it only vanishes if $y = 0$ and $x \ge 0$.

So to show that f is continuously differentiable on the complement of the half-line $y = 0, x \le 0$, we need to show that f is continuously differentiable on the open half-line where $y = 0$ and $x > 0$. In a neighborhood of a point $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ satisfying $x_0 > 0$ and $y_0 = 0$, we can write (using the Taylor polynomial given in equation 3.4.9, with $m = 1/2$ and $n = 1$),

We can factor out the x to write
$$-x+\sqrt{x^2+y^2} = -x+x\sqrt{1+\frac{y^2}{x^2}}$$
because x is positive in a neighborhood of $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$.

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

Since $\sqrt{y^2} = (\operatorname{sgn} y)y$, we see that in a neighborhood of $\begin{pmatrix} x_0 \\ 0 \end{pmatrix}$ with $x_0 > 0$ we have

$$f\begin{pmatrix} x \\ y \end{pmatrix} = (\operatorname{sgn} y)(\operatorname{sgn} y)\left(\frac{y}{2\sqrt{x}}+o(y)\right) = \frac{y}{2\sqrt{x}}+o(y).$$

The function f vanishes identically on the positive x-axis, so we see that f is differentiable on the positive x-axis, with

$$\left[\mathbf{D}f\begin{pmatrix} x \\ 0 \end{pmatrix}\right] = \begin{bmatrix} 0 & \dfrac{1}{2\sqrt{x}} \end{bmatrix}.$$

To see that f is of class C^1 on the positive x-axis, we need to show that the partial derivatives are continuous there. This is a straightforward but messy computation using the chain rule: if $y \neq 0$ we find

$$D_1 f\begin{pmatrix} x \\ y \end{pmatrix} = \operatorname{sgn} y \frac{x-\sqrt{x^2+y^2}}{4\sqrt{x^2+y^2}\sqrt{\frac{1}{2}(-x+\sqrt{x^2+y^2})}}$$

which tends to 0 when y tends to 0 since after cancellations there is a $\sqrt{x-\sqrt{x^2+y^2}}$ left over. The partial derivative with respect to y is

$$D_2 f\begin{pmatrix} x \\ y \end{pmatrix} = \operatorname{sgn} y \frac{y}{4\sqrt{x^2+y^2}\sqrt{\frac{1}{2}(-x+\sqrt{x^2+y^2})}} = \operatorname{sgn} y \frac{y}{4x\frac{(\operatorname{sgn} y)y}{2\sqrt{x}}} + o(y).$$

Again the sgn(y)'s cancel, the y's cancel, one power of 2 cancels, as well as an $\sqrt{x}$, leaving $\dfrac{1}{2\sqrt{x}}$. So f is continuously differentiable on the open half-line $y = 0$ and $x > 0$, i.e., on the complement of the half-line $y = 0$, $x \leq 0$.

Here we are dealing with the case $n = 2$, $k = 0$, so Theorem A12.5 says that the remainder should be
$$\sum_{I\in\mathcal{I}_2^1} \frac{1}{I!} D_I f(\mathbf{c})\, \vec{\mathbf{h}}^I$$
$$= [D_1 f(\mathbf{c})\ D_2 f(\mathbf{c})]\vec{\mathbf{h}}.$$

b. We have

$$f(\mathbf{a}) = -\sqrt{\frac{1+\sqrt{1+\epsilon^2}}{2}}, \quad f(\mathbf{a}+\vec{\mathbf{h}}) = +\sqrt{\frac{1+\sqrt{1+\epsilon^2}}{2}}.$$

So

$$f(\mathbf{a}+\vec{\mathbf{h}}) - f(\mathbf{a}) = 2\sqrt{\frac{1+\sqrt{1+\epsilon^2}}{2}},$$

which tends to 2 as $\epsilon \to 0$, and cannot be $[\mathbf{D}f(\mathbf{c})]\mathbf{h}$ for any $\mathbf{c} \in [\mathbf{a}, \mathbf{a}+\vec{\mathbf{h}}]$, i.e., any $\mathbf{c} \in \left[\begin{pmatrix} -1 \\ -\epsilon \end{pmatrix}, \begin{pmatrix} -1 \\ \epsilon \end{pmatrix}\right]$, since $[\mathbf{D}f(\mathbf{c})]$ remains bounded. To see that

$[\mathbf{D}f(\mathbf{c})]$ remains bounded, first we compute

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \end{pmatrix}\right] = \left[\operatorname{sgn} y \frac{-\frac{1}{2} + \frac{x}{\sqrt{x^2+y^2}}}{2\sqrt{\frac{-x+\sqrt{x^2+y^2}}{2}}}, \quad \operatorname{sgn} y \frac{\frac{y}{\sqrt{x^2+y^2}}}{2\sqrt{\frac{-x+\sqrt{x^2+y^2}}{2}}}\right].$$

For $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \in [\mathbf{a}, \mathbf{a} + \vec{\mathbf{h}}]$, we have $c_1 = -1$, and the derivative does not blow up because the denominators do not tend to 0 as $c_2 \to 0$:

$$\sqrt{\frac{-(-1) + \sqrt{(-1)^2 + c_2^2}}{2}} \underset{\text{as } c_2 \to 0}{\longrightarrow} \sqrt{\frac{1+\sqrt{1}}{2}} \neq 0$$

c. Theorem A12.5 requires that f be of class C^{k+1} on $[\mathbf{a}, \mathbf{a} + \mathbf{h}]$; in our case $k = 0$ and we need f of class C^1 on $[\mathbf{a}, \mathbf{a} + \vec{\mathbf{h}}]$, which it isn't: the line of discontinuity of f crosses $[\mathbf{a}, \mathbf{a} + \vec{\mathbf{h}}]$, so we do not have $[\mathbf{a}, \mathbf{a} + \vec{\mathbf{h}}] \subset U$.

A14.1 a. We have $f\begin{pmatrix} x \\ y \end{pmatrix} = y$ and $F\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 - 1$, so at $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ we have $\lambda = 1/2$, and at $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ we have $\lambda = -1/2$. We have

$$L_{f,F}\begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = y - \lambda(x^2 + y^2 - 1)$$

$$\Psi\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x^2 + y^2 - 1 \end{pmatrix}$$

$$\left[\mathbf{D}\Psi\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \qquad \left[\mathbf{D}\Psi\begin{pmatrix} 0 \\ -1 \end{pmatrix}\right] = \begin{bmatrix} 1 & 0 \\ 2 & -2 \end{bmatrix}.$$

The two functions Φ_1 and Φ_2 are

$$\Phi_1\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} x \\ \sqrt{u + 1 - x^2} \end{pmatrix} \quad \text{for } \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and

$$\Phi_2\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} x \\ -\sqrt{u + 1 - x^2} \end{pmatrix} \quad \text{for } \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

This gives

$$\Phi_1\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ \sqrt{1 - x^2} \end{pmatrix} = \widetilde{\mathbf{g}_1}(x) \quad \text{and} \quad \Phi_2\begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ -\sqrt{1 - x^2} \end{pmatrix} = \widetilde{\mathbf{g}_2}(x),$$

where

$$\mathbf{g}_1(x) = +\sqrt{1 - x^2}, \qquad \mathbf{g}_2(x) = -\sqrt{1 - x^2}.$$

Then

$$\widetilde{f}_1\begin{pmatrix} x \\ u \end{pmatrix} = f \circ \Phi_1 \begin{pmatrix} x \\ u \end{pmatrix} = +\sqrt{u+1-x^2}, \qquad \widetilde{f}_2 = f \circ \Phi_2 = -\sqrt{u+1-x^2}$$

$$\widetilde{F}_1\begin{pmatrix} x \\ u \end{pmatrix} = F \circ \Phi_1 \begin{pmatrix} x \\ u \end{pmatrix} = F\begin{pmatrix} x \\ \sqrt{u+1-x^2} \end{pmatrix} = x^2 + \left(\sqrt{u+1-x^2}\right)^2 - 1 = u$$

$$\widetilde{F}_2\begin{pmatrix} x \\ u \end{pmatrix} = F \circ \Phi_2 \begin{pmatrix} x \\ u \end{pmatrix} = F\begin{pmatrix} x \\ -\sqrt{u+1-x^2} \end{pmatrix} = x^2 + \left(-\sqrt{u+1-x^2}\right)^2 - 1 = u$$

and

$$L_{\widetilde{f}_1,\widetilde{F}_1}\begin{pmatrix} x \\ u \\ \lambda \end{pmatrix} = \widetilde{f}_1\begin{pmatrix} x \\ u \end{pmatrix} - \lambda u = \sqrt{u+1-x^2} - \lambda u = L_{f,F}\begin{pmatrix} \Phi_1\begin{pmatrix} x \\ u \end{pmatrix} \\ \lambda \end{pmatrix}$$

$$L_{\widetilde{f}_2,\widetilde{F}_2}\begin{pmatrix} x \\ u \\ \lambda \end{pmatrix} = \widetilde{f}_2\begin{pmatrix} x \\ u \end{pmatrix} - \lambda u = -\sqrt{u+1-x^2} - \lambda u = L_{f,F}\begin{pmatrix} \Phi_2\begin{pmatrix} x \\ u \end{pmatrix} \\ \lambda \end{pmatrix}.$$

To find the Hessian matrix of $L_{\widetilde{f}_1,\widetilde{F}_1}$, it is easiest to develop the function as a Taylor polynomial

$$1 + \frac{1}{2}(u - x^2) - \frac{1}{8}u^2 - \lambda u$$

to degree 2 (rather than computing the second derivatives of $L_{\widetilde{f}_1,\widetilde{F}_1}$ directly).

The Hessian matrix of $L_{\widetilde{f}_1,\widetilde{F}_1}$ at $\begin{pmatrix} 0 \\ 0 \\ 1/2 \end{pmatrix}$ is $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1/4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$, giving

$$-A^2 - \frac{B^2}{4} - 2BC = -A^2 - \left(\frac{B}{2} + 2C\right)^2 + 4C^2,$$

with signature $(1,2)$.

The Hessian matrix of $L_{\widetilde{f}_2,\widetilde{F}_2}$ at $\begin{pmatrix} 0 \\ 0 \\ -1/2 \end{pmatrix}$ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$, giving

$$A^2 + \frac{B^2}{4} - 2BC = A^2 + \left(\frac{B}{2} - 2C\right)^2 - 4C^2,$$

with signature $(2,1)$.

The Hessian matrix of $L_{f,F}$ at $\begin{pmatrix} 0 \\ 1 \\ 1/2 \end{pmatrix}$ is $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & -2 \\ 0 & -2 & 0 \end{bmatrix}$, with signature $(1,2)$.

The Hessian matrix of $L_{f,F}$ at $\begin{pmatrix} 0 \\ -1 \\ -1/2 \end{pmatrix}$ is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 0 \end{bmatrix}$, with signature $(2,1)$.

The point $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is a maximum, with signature $(p = 0,\ q = 1)$, so $(p+1, q+1) = (1,2)$. The point $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ is a minimum, with signature $(p = 1,\ q = 0)$, so $(p+1, q+1) = (2,1)$.

b. We have $f\begin{pmatrix} x \\ y \end{pmatrix} = y - ax^2$ and $F\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 - 1$.

The Lagrange multiplier equations for the critical points are

$$[-2ax \quad 1] = \lambda[2x \quad 2y] \quad \text{and} \quad x^2 + y^2 = 1.$$

The equation $-2ax = 2\lambda x$ leads to two possibilities, $x = 0$ or $\lambda = -a$. The first possibility leads to the same two critical points as in part a:

$$\begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1/2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -1/2 \end{pmatrix}. \tag{1}$$

These solutions exist for all values of a.

In addition, when $|a| > 1/2$, there are two more solutions, coming from the equality $\lambda = -a$:

$$\begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} +\sqrt{1 - 1/(4a^2)} \\ -1/2a \\ -a \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} -\sqrt{1 - 1/(4a^2)} \\ -1/2a \\ -a \end{pmatrix}. \tag{2}$$

Let us begin by treating the solutions in (1).

We have

$$L_{f,F} \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = y - ax^2 - \lambda(x^2 + y^2 - 1)$$

$$\Psi \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x^2 + y^2 - 1 \end{pmatrix}$$

$$\left[\mathbf{D}\Psi\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix}, \qquad \left[\mathbf{D}\Psi\begin{pmatrix} 0 \\ -1 \end{pmatrix}\right] = \begin{bmatrix} 1 & 0 \\ 2 & -2 \end{bmatrix}.$$

Equations 3–6 involve only the constraint and are identical with part a: the two functions Φ_1 and Φ_2 are

$$\Phi_1 \begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} x \\ \sqrt{u + 1 - x^2} \end{pmatrix} \quad \text{for } \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3}$$

and

$$\Phi_2 \begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} x \\ -\sqrt{u + 1 - x^2} \end{pmatrix} \quad \text{for } \begin{pmatrix} 0 \\ -1 \end{pmatrix} \tag{4}$$

This gives

$$\Phi_1 \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ \sqrt{1 - x^2} \end{pmatrix} = \widetilde{g_1}(x) \quad \text{and} \quad \Phi_2 \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ -\sqrt{1 - x^2} \end{pmatrix} = \widetilde{g_2}(x), \tag{5}$$

where

$$g_1(x) = +\sqrt{1 - x^2}, \qquad g_2(x) = -\sqrt{1 - x^2}. \tag{6}$$

Then

$$\widetilde{f_1} \begin{pmatrix} x \\ u \end{pmatrix} = f \circ \Phi_1 \begin{pmatrix} x \\ u \end{pmatrix} = +\sqrt{u + 1 - x^2} - ax^2,$$

$$\widetilde{f_2} \begin{pmatrix} x \\ u \end{pmatrix} = f \circ \Phi_2 \begin{pmatrix} x \\ u \end{pmatrix} = -\sqrt{u + 1 - x^2} - ax^2$$

$$\widetilde{F_1} \begin{pmatrix} x \\ u \end{pmatrix} = F \circ \Phi_1 \begin{pmatrix} x \\ u \end{pmatrix} = F \begin{pmatrix} x \\ \sqrt{u + 1 - x^2} \end{pmatrix} = x^2 + \left(\sqrt{u + 1 - x^2}\right)^2 - 1 = u$$

$$\widetilde{F_2} \begin{pmatrix} x \\ u \end{pmatrix} = F \circ \Phi_2 \begin{pmatrix} x \\ u \end{pmatrix} = F \begin{pmatrix} x \\ -\sqrt{u + 1 - x^2} \end{pmatrix} = x^2 + \left(-\sqrt{u + 1 - x^2}\right)^2 - 1 = u$$

and

$$L_{\widetilde{f}_1,\widetilde{F}_1}\begin{pmatrix} x\\ u\\ \lambda\end{pmatrix} = \widetilde{f}_1\begin{pmatrix} x\\ u\end{pmatrix} - \lambda u = \sqrt{u+1-x^2} - ax^2 - \lambda u$$
$$= L_{f,F}\begin{pmatrix} \Phi_1\begin{pmatrix} x\\ u\end{pmatrix}\\ \lambda\end{pmatrix}$$
$$L_{\widetilde{f}_2,\widetilde{F}_2}\begin{pmatrix} x\\ u\\ \lambda\end{pmatrix} = \widetilde{f}_2\begin{pmatrix} x\\ u\end{pmatrix} - \lambda u = -\sqrt{u+1-x^2} - ax^2 - \lambda u$$
$$= L_{f,F}\begin{pmatrix} \Phi_2\begin{pmatrix} x\\ u\end{pmatrix}\\ \lambda\end{pmatrix}.$$

To compute the second derivatives of

$$L_{\widetilde{f}_1,\widetilde{F}_1}$$

at the origin, it is easiest to develop the function as a Taylor polynomial

$$1+\frac{1}{2}(u-x^2)-\frac{1}{8}u^2-ax^2-\lambda u$$

to degree 2.

The Hessian matrix of $L_{\widetilde{f}_1,\widetilde{F}_1}$ at $\begin{pmatrix} 0\\ 0\\ 1/2\end{pmatrix}$ is

$$\begin{bmatrix} -1-2a & 0 & 0\\ 0 & -1/4 & -1\\ 0 & -1 & 0\end{bmatrix}, \quad \text{giving}$$

$$-(1+2a)A^2 - \frac{B^2}{4} - 2BC = -(1+2a)A^2 - \left(\frac{B}{2}+2C\right)^2 + 4C^2,$$

with signature $(1,2)$ if $a > -1/2$, and signature $(2,1)$ if $a < -1/2$. For $a = -1/2$, the form is degenerate.

The Hessian matrix of $L_{\widetilde{f}_2,\widetilde{F}_2}$ at $\begin{pmatrix} 0\\ 0\\ -1/2\end{pmatrix}$ is

$$\begin{bmatrix} 1-2a & 0 & 0\\ 0 & 1/4 & -1\\ 0 & -1 & 0\end{bmatrix}, \quad \text{giving}$$

$$(1-2a)A^2 + \frac{B^2}{4} - 2BC = (1-2a)A^2 + \left(\frac{B}{2}-2C\right)^2 - 4C^2,$$

with signature $(1,2)$ if $a > 1/2$, and signature $(2,1)$ if $a < 1/2$. For $a = 1/2$, the form is degenerate.

The Hessian matrix of $L_{f,F}$ at $\begin{pmatrix} 0\\ 1\\ 1/2\end{pmatrix}$ is $\begin{bmatrix} -1-2a & 0 & 0\\ 0 & -1 & -2\\ 0 & -2 & 0\end{bmatrix}$, again with signature $(1,2)$ if $a > -1/2$, and signature $(2,1)$ if $a < -1/2$. For $a = -1/2$, the form is degenerate.

The Hessian matrix of $L_{f,F}$ at $\begin{pmatrix} 0\\ -1\\ -1/2\end{pmatrix}$ is $\begin{bmatrix} 1-2a & 0 & 0\\ 0 & 1 & 2\\ 0 & 2 & 0\end{bmatrix}$, again with signature $(1,2)$ if $a > 1/2$, and signature $(2,1)$ if $a < 1/2$. For $a = 1/2$, the form is degenerate.

Thus, for the critical points in (1), the signatures of the quadratic forms given by the Hessians of $L_{f,F}$ and $L_{\widetilde{f},\widetilde{F}}$ agree at both $\binom{0}{1}$ and $\binom{0}{-1}$, as Theorem 3.7.13 demands. For $a > 1/2$, both points are maxima, and for $a < -1/2$ both points are minima. For $-1/2 < a < 1/2$, the point $\binom{0}{1}$ is a maximum and the point $\binom{0}{-1}$ is a minimum.

Now we will consider the critical points

$$\begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} +\sqrt{1-\frac{1}{4a^2}} \\ -1/2a \\ -a \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} = \begin{pmatrix} -\sqrt{1-\frac{1}{4a^2}} \\ -1/2a \\ -a \end{pmatrix}$$

of equation 2, which exist only if $|a| > 1/2$.

The Hessian matrix of $L_{f,F}$ at $\begin{pmatrix} +\sqrt{1-\frac{1}{4a^2}} \\ -1/2a \\ -a \end{pmatrix}$ is

$$\begin{bmatrix} -2a-2\lambda & 0 & -2x \\ 0 & -2\lambda & -2y \\ -2x & -2y & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \mp 2\sqrt{1-\frac{1}{4a^2}} \\ 0 & 2a & 1/a \\ \mp 2\sqrt{1-\frac{1}{4a^2}} & 1/a & 0 \end{bmatrix}.$$

The corresponding quadratic form

$$\begin{aligned} 2aB^2 \mp 4\sqrt{1-\frac{1}{4a^2}}AC + \frac{2}{a}BC = 2a\left(B^2 + \frac{1}{a^2}BC + \frac{1}{4a^4}C^2\right) \\ - \frac{1}{2a^3}\left(C^2 \pm 2a^3\sqrt{1-\frac{1}{4a^2}}AC + a^6\left(1-\frac{1}{4a^2}\right)A^2\right) \\ + \frac{a^3}{2}\left(1-\frac{1}{4a^2}\right)A^2 \end{aligned}$$

If the square root on the left is the positive square root, then that on the right is the negative square root, and vice versa: the top sign in $\mp$ goes with the top sign in $\pm$.

has signature $(2,1)$ if $a > 1/2$ and signature $(1,2)$ if $a < -1/2$. Thus both of the new critical points are minima if $a > 1/2$ and maxima if $a < -1/2$.

Now the exercise requires us to find the signature of $L_{\widetilde{f},\widetilde{F}}$. If $a > 1/2$, then for both critical points, we have

$$\Phi\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} x \\ -\sqrt{1+u-x^2} \end{pmatrix},$$

since in that case $y = -1/2a < 0$. Then

$$\widetilde{f}\binom{x}{u} = f \circ \Phi\binom{x}{u} = -\sqrt{u+1-x^2} - ax^2,$$

so

$$L_{\widetilde{f},\widetilde{F}}\begin{pmatrix} x \\ u \\ \lambda \end{pmatrix} = \widetilde{f}_2\binom{x}{u} - \lambda u = -\sqrt{u+1-x^2} - ax^2 - \lambda u,$$

and we need to compute its Hessian at the two critical points; in this case it is easier to compute the Taylor polynomial to degree 2 at the critical point, checking along the way that the terms of degree 1 drop out as they should.[8] Thus set

$$x = \sqrt{1 - \frac{1}{4a^2}} + A, \quad u = 0 + B, \quad \lambda = -a + C,$$

to find

$$\begin{aligned}
&-\sqrt{B + 1 - \left(\sqrt{1 - \frac{1}{4a^2}} + A\right)^2} - a\left(\sqrt{1 - \frac{1}{4a^2}} + A\right)^2 - (-a + C)B)\\
&= -\sqrt{\frac{1}{4a^2} + B - 2A\sqrt{1 - \frac{1}{4a^2}} - A^2} - a\left(1 - \frac{1}{4a^2} + 2A\sqrt{1 - \frac{1}{4a^2}} + A^2\right) + aB - BC\\
&= -\frac{1}{2a}\left(1 + \frac{1}{2}\left(4a^2B - 8a^2A\sqrt{1 - \frac{1}{4a^2}} - 4a^2A^2\right) - \frac{1}{8}\left(4a^2B - 8a^2A\sqrt{1 - \frac{1}{4a^2}}\right)^2\right)\\
&\qquad - a + \frac{1}{4a} - 2aA\sqrt{1 - \frac{1}{4a^2}} - aA^2 + aB - BC + o(A^2 + B^2 + C^2).
\end{aligned}$$

Indeed, the linear terms in A and B do vanish, and the quadratic terms are (after a bit of computation)

$$\begin{aligned}
&a(4a^2 - 1)A^2 + a^3B^2 - 2a^2\sqrt{4a^2 - 1}AB + a^3B^2 - BC\\
&= a(4a^1 - 1)\left(A - \frac{a}{\sqrt{4a^2 - 1}}\right)^2 - \frac{1}{4}(B + C)^2 + \frac{1}{4}(B - C)^2.
\end{aligned}$$

Recall that $a > 1/2$, so this quadratic form has signature $(2, 1)$, representing a minimum, as was required. Note that for the point $-\sqrt{1 + \frac{1}{4a^2}}$ the computation is the same, just interpreting the square root throughout as the negative square root.

In the case $a < -1/2$, we need the other possible Φ, namely

$$\Phi\begin{pmatrix} x \\ u \end{pmatrix} = \begin{pmatrix} x \\ \sqrt{1 + u - x^2} \end{pmatrix}, \quad \text{since} \quad y = -1/2a > 0.$$

This leads to

$$L_{\tilde{f}, \tilde{F}}\begin{pmatrix} x \\ u \\ \lambda \end{pmatrix} = +\sqrt{u + 1 - x^2} - ax^2 - \lambda u.$$

Rather surprisingly, this leads to the Taylor polynomial

$$\begin{aligned}
&-\frac{1}{2a}\left(1 + \frac{1}{2}\left(4a^2B - 8a^2A\sqrt{1 - \frac{1}{4a^2}} - 4a^2A^2\right) - \frac{1}{8}\left(4a^2B - 8a^2A\sqrt{1 - \frac{1}{4a^2}}\right)^2\right)\\
&\qquad - a + \frac{1}{4a} - 2aA\sqrt{1 - \frac{1}{4a^2}} - aA^2 + aB - BC + o(A^2 + B^2 + C^2);
\end{aligned}$$

[8] In the earlier part of part b we did not use Taylor polynomials, but it probably would have been easier to do so.

the minus sign in front of the first term is still there. Indeed, the square root $+\sqrt{u+1-x^2}$ is the positive one, and since $a < -1/2$ (and in particular, $a < 0$), we must take the negative square root of $4a^2$ when factoring it out of the square root. Thus the quadratic form is again

$$a(4a^2-1)\left(A - \frac{a}{\sqrt{4a^2-1}}\right)^2 - \frac{1}{4}(B+C)^2 + \frac{1}{4}(B-C)^2,$$

which this time has signature $(1, 2)$, since $a < -1/2$. Thus again we confirm Theorem 3.7.13.

A15.1 By Definition 3.9.7, $H = (1/2)(A_{2,0} + A_{0,2})$. Equation 3.9.36 gives us expressions for $A_{2,0}$ and $A_{0,2}$. It is now straightforward to show that H is given by the desired expression. Remember we have set $c = \sqrt{a_1^2 + a_2^2}$, so that $a_1^2 + a_2^2 = c^2$. Thus we have

$$\begin{aligned}
H &= \frac{1}{2}\left(A_{2,0} + A_{0,2}\right) \\
&= \frac{-1}{2c^2(1+c^2)^{1/2}}\left(a_{2,0}a_2^2 - 2a_{1,1}a_1a_2 + a_{0,2}a_1^2\right)\frac{-1}{2c^2(1+c^2)^{3/2}}\left(a_{2,0}a_1^2 + 2a_{1,1}a_1a_2 + a_{0,2}a_2^2\right) \\
&= \frac{-1}{2c^2\left(1+c^2\right)^{3/2}}\Big(\left(1+c^2\right)\left(a_{2,0}a_2^2 - 2a_{1,1}a_1a_2 + a_{0,2}a_1^2\right) + a_{2,0}a_1^2 + 2a_{1,1}a_1a_2 + a_{0,2}a_2^2\Big) \\
&= \frac{-1}{2c^2\left(1+c^2\right)^{3/2}}\Big(a_{2,0}(a_2^2 + a_1^2) + a_{0,2}(a_1^2 + a_2^2) + c^2a_{2,0}a_2^2 - c^2 2a_{1,1}a_1a_2 + c^2a_{0,2}a_1^2\Big) \\
&= \frac{-1}{2\left(1+c^2\right)^{3/2}}\Big(a_{2,0} + a_{0,2} + a_{2,0}a_2^2 + a_{0,2}a_1^2 - 2a_{1,1}a_1a_2\Big) \\
&= \frac{-1}{2\left(1+c^2\right)^{3/2}}\Big(a_{2,0}(1+a_2^2) - 2a_{1,1}a_1a_2 + a_{0,2}(1+a_1^2)\Big).
\end{aligned}$$

In going to the next-to-last line, one c^2 in the $2c^2$ in the denominator cancels the c^2 and $a_1^2 + a_2^2$ in the numerator.

A16.1 a. Since $0 \le \sin x \le 1$ for $x \in [0, \pi]$, it follows that

$$0 \le \sin^n x \le \sin^{n-1} x \le 1$$

for $x \in [0, \pi]$. So

$$c_n = \int_0^\pi \sin^n x\, dx < \int_0^\pi \sin^{n-1} x\, dx = c_{n-1}.$$

b. We have

$$\begin{aligned}
c_n &= \int_0^\pi \sin^n x\, dx = \int_0^\pi \sin x \sin^{n-1} x\, dx \\
&= \left[-\cos x \sin^{n-1} x\right]_0^\pi + \int_0^\pi \cos x (n-1)\sin^{n-2} x \cos x\, dx.
\end{aligned}$$

But

$$\left[-\cos x \sin^{n-1} x\right]_0^\pi = 0 \quad \text{for } n \ge 2,$$

so

$$c_n = \int_0^\pi (n-1)\left(\sin^{n-2} x - \sin^n x\right)\, dx.$$

Therefore,

$$c_n = (n-1)(c_{n-2} - c_n), \quad \text{so} \quad nc_n = (n-1)c_{n-2}, \quad \text{i.e.,} \quad c_n = \frac{n-1}{n}c_{n-2}.$$

c. The formulas $c_0 = \pi$ and $c_1 = 2$ follow immediately from simple integrals. Since

$$(2n-1)(2n-3)\cdots(1) = \frac{(2n)(2n-1)(2n-2)\cdots(1)}{(2n)(2n-2)\cdots(2)}$$

and

$$(2n)(2n-2)\cdots(2) = (2n)\big(2(n-1)\big)\cdots(2\cdot 1) = 2^n\, n!,$$

we have

$$c_{2n} = \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdots\frac{1}{2}\pi = \frac{(2n)!}{(2^n n!)^2}\,\pi = \frac{(2n)!}{2^{2n}(n!)^2}\,\pi$$

$$c_{2n+1} = \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdots\frac{2}{3}\,2 = \frac{(2^n n!)^2}{(2n+1)!}\,2 = \frac{2^{2n}(n!)^2}{(2n+1)!}2.$$

d. Noting that

$$\frac{1+o(1)}{1+o(1)} = 1+o(1),$$

we get by simple substitution that

$$c_{2n} = \frac{1}{C}\sqrt{\frac{2}{n}}\,\pi\big(1+o(1)\big) \quad \text{and} \quad c_{2n+1} = \frac{C}{\sqrt{2n+1}}\big(1+o(1)\big).$$

So since $c_{2n} > c_{2n+1}$, we have

$$\sqrt{2}\sqrt{\frac{2n+1}{n}}\pi\big(1+o(1)\big) \geq C^2\big(1+o(1)\big),$$

but

$$\sqrt{\frac{2n+1}{n}} = \sqrt{2}\big(1+o(1)\big), \quad \text{so} \quad C^2 \leq 2\pi\big(1+o(1)\big).$$

We also have

$$c_{2n} < c_{2n-1}, \quad \text{so } C^2 \geq 2\pi\big(1+o(1)\big).$$

So if there exists C such that $n! = C\sqrt{n}\left(\frac{n}{e}\right)^n\big(1+o(1)\big)$, then $C = \sqrt{2\pi}$.

A16.3 It is not really the object of this exercise, but let us see by induction that the number of ways of picking k things from m, called "m choose k," is given by the formula

$$\binom{m}{k} = \frac{m!}{k!(m-k)!}$$

which you probably saw in high school and in Theorem 6.1.10. One way of seeing this is the relation leading to Pascal's triangle:

$$\binom{m}{k} = \binom{m-1}{k-1} + \binom{m-1}{k},$$

which expresses the fact that to choose k things among m, you must either choose $k-1$ things among the first $m-1$, and then choose the last also, or choose k things among the first $m-1$, and then not choose the last.

Suppose that the formula is true for all $m-1$ and all k, with the convention that $\binom{m}{k} = 0$ if $k < 0$ or $k > m$. Then the inductive step is

$$\begin{aligned}\binom{m}{k} &= \binom{m-1}{k-1} + \binom{m-1}{k} \\ &= \frac{(m-1)!}{(k-1)!(m-k)!} + \frac{(m-1)!}{k!(m-k-1)!} \\ &= \frac{(m-1)!}{(k-1)!(m-k-1)!}\left(\frac{1}{m-k} + \frac{1}{k}\right) \\ &= \frac{(m-1)!}{(k-1)!(m-k-1)!}\frac{m}{k(m-k)} = \frac{m!}{k!(m-k)!}.\end{aligned}$$

Now to our question.

Since the coin is being tossed $2n$ times, there are 2^{2n} possible sequences of tosses; saying that the coin is a fair coin is exactly saying that all such sequences have the same probability $1/2^{2n}$.

The number of sequences corresponding to $n+k$ heads is exactly the number of ways of choosing $n+k$ tosses among the $2n$, i.e., $2n$ choose $n+k$:

$$\binom{2n}{n+k} = \frac{(2n)!}{(n+k)!(n-k)!}.$$

We need to sum this over all k such that $n + a\sqrt{n} \le n+k \le n + b\sqrt{n}$.

A21.1 a. For any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, we have

$$0 \le f(\mathbf{x}_2) = \inf_{\mathbf{y}\in X} |\mathbf{x}_2 - \mathbf{y}| \le \inf_{\mathbf{y}\in X}\Big(|\mathbf{x}_2 - \mathbf{x}_1| + |\mathbf{x}_1 - \mathbf{y}|\Big) = |\mathbf{x}_2 - \mathbf{x}_1| + f(\mathbf{x}_1).$$

The same statement holds if we exchange $\mathbf{x}_1$ and $\mathbf{x}_2$, so

$$|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \le |\mathbf{x}_1 - \mathbf{x}_2|.$$

This proves that f is continuous.

If $\mathbf{x} \in \overline{X}$, there exists a sequence $i \mapsto \mathbf{y}_i$ converging to $\mathbf{x}$, so $\inf_{\mathbf{y}\in X} |\mathbf{x} - \mathbf{y}_i| = 0$, hence $f(\mathbf{x}) = 0$. But if $\mathbf{x} \notin \overline{X}$, then there exists $\epsilon > 0$ such that $|\mathbf{x} - \mathbf{y}| > \epsilon$ for all $\mathbf{y} \in X$, hence $f(\mathbf{x}) \ge \epsilon$.

b. There isn't much to show: this function is bounded by 1 and has support in $[0,1]$; it is continuous, (that was the point of including 0 and 1 in X'), it satisfies $f \le 0$ everywhere, and $f(x) = 0$ precisely if $x \notin [0,1]$ or if x is in the unpavable set X'.

A21.3 This has nothing to do with what $\mathbb{R}^n - X_\epsilon$ might be; to lighten notation, set $Z \stackrel{\text{def}}{=} \mathbb{R}^n - X$ and denote by $d(\mathbf{x}, Z)$ the distance between $\mathbf{x}$ and the nearest point in Z: $d(\mathbf{x}, Z) \stackrel{\text{def}}{=} \inf_{\mathbf{y}\in Z} |\mathbf{x}-\mathbf{y}|$, Then by the triangle inequality

$$d(\mathbf{x}, Z) \le |\mathbf{x}-\mathbf{y}| + d(\mathbf{y}, Z)$$
$$d(\mathbf{y}, Z) \le |\mathbf{y}-\mathbf{x}| + d(\mathbf{x}, Z)$$

so

$$d(\mathbf{x}, Z) - d(\mathbf{y}, Z) \le |\mathbf{x}-\mathbf{y}|, \quad d(\mathbf{y}, Z) - d(\mathbf{x}, Z) \le |\mathbf{x}-\mathbf{y}|,$$

and finally $|d(\mathbf{x}, Z) - d(\mathbf{y}, Z)| \le |\mathbf{x}-\mathbf{y}|$. It then follows immediately that

$$\Big|(1 - d(\mathbf{x}, Z)) - (1 - d(\mathbf{y}, Z))\Big| = \Big|d(\mathbf{y}, Z) - d(\mathbf{x}, Z)\Big| \le |\mathbf{x}, \mathbf{y}|.$$

If $f, g : \mathbb{R}^n \to \mathbb{R}$ satisfy

$$|f(\mathbf{x}) - f(\mathbf{y})| \le |\mathbf{x}-\mathbf{y}| \quad \text{and} \quad |g(\mathbf{x}) - g(\mathbf{y})| \le |\mathbf{x}-\mathbf{y}|,$$

then $|\sup(f,g)(\mathbf{x}) - \sup(f,g)(\mathbf{y})| \le |\mathbf{x}-\mathbf{y}|$, as you can check by looking at the four cases where each of $\sup(f,g)(\mathbf{x})$ and $\sup(f,g)(\mathbf{y})$ is realized by f or g. Applying this to the case $f(\mathbf{z}) = 1 - d(\mathbf{x}, Z)$ and $g(\mathbf{z}) = 0$, we get $h(\mathbf{x}) - h(\mathbf{y})| \le |\mathbf{x}-\mathbf{y}|$.

Now we will show that $h(\mathbf{x}) = 1$ when $\mathbf{x} \notin X_\epsilon$, and $0 \le h(\mathbf{x}) < 1$ when $\mathbf{x} \in X_\epsilon$. By definition, $h = 1$ on $Z = \mathbb{R}^n - X_\epsilon$, which is closed. Thus if $\mathbf{z} \notin \mathbb{R}^n - X_\epsilon$, there is a ball $B_{\mathbf{z}}(r)$ of some radius $0 < r < 1$ around $\mathbf{z}$ such that $B_{\mathbf{z}}(r) \cap \mathbb{R}^n - X_\epsilon = \phi$, and $h(\mathbf{z}) \le 1 - r$. Thus $\mathbb{R}^n - X_\epsilon$ is the set of $\mathbf{z} \in \mathbb{R}^n$ such that $h(\mathbf{z}) = 0$. The number $h(\mathbf{z})$ is the supremum of $1 - d(\mathbf{z}, Z)$ and 0, so of course $h(\mathbf{z}) \ge 0$, and since $d(\mathbf{z}, Z) \ge 0$, we have $h(\mathbf{z}) \le 1$.

A21.5 The hypotheses of Theorem 4.11.20 imply that Φ^{-1} is bijective and of class C^1 with Lipschitz derivative. So the direction "$\Longrightarrow$" asserts that if

$$\underbrace{(f \circ \Phi)|\det[\mathbf{D}\Phi]|}_{\text{like } f \text{ in original statement}}$$

is integrable on U, then

$$\Big(\big((f \circ \Phi)|\det[\mathbf{D}\Phi]\big) \circ \Phi^{-1}\Big)|\det[\mathbf{D}\Phi^{-1}]| = \big((f \circ \Phi) \circ \Phi^{-1}\big)|\det[\mathbf{D}\Phi] \circ \Phi^{-1}||\det[\mathbf{D}\Phi^{-1}]|$$

is integrable on V, and

$$\int_U (f \circ \Phi)(\mathbf{x})\big|\det[\mathbf{D}\Phi(\mathbf{x})]\big|\,|d^n\mathbf{x}| = \int_V (f \circ \Phi \circ \Phi^{-1})(\mathbf{y})\ \Big|\det\Big[\mathbf{D}\Phi\Big(\Phi^{-1}(\mathbf{y})\Big)\Big]\Big|\ |\det[\mathbf{D}\Phi^{-1}(\mathbf{y})]|\,|d^n\mathbf{y}|$$
$$= \int_V f(\mathbf{y})\,\big|\det[\mathbf{D}(\Phi \circ \Phi^{-1})(\mathbf{y})]\big|\,|d^n\mathbf{y}| = \int_V f(\mathbf{y})\,|d^n\mathbf{y}|.$$

A22.1 To simplify notation, we will set $\varphi = dx_{i_1} \wedge \cdots \wedge dx_{i_k}$. Thus we want to show that the definition of the wedge product justifies going from the last line of equation A22.5 to the preceding line, i.e.,

$$(\mathbf{d}f \wedge \varphi)\big(P_0(\vec{\mathbf{v}}_1, \ldots, \vec{\mathbf{v}}_{k+1})\big) = \sum_{i=1}^{k+1} (-1)^{i-1}\big([\mathbf{D}f(\mathbf{0})]\vec{\mathbf{v}}_i\big)\varphi(\vec{\mathbf{v}}_1, \ldots, \widehat{\vec{\mathbf{v}}}_i, \ldots, \vec{\mathbf{v}}_{k+1}).$$

Since f is a 0-form, $\mathbf{d}f$ is a 1-form, and φ is a k-form. So Definition 6.1.12 of the wedge product says

$$\mathbf{d}f \wedge \varphi(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{k+1}) = \sum_{\substack{\text{shuffles} \\ \sigma \in \text{Perm}(1,k)}} \operatorname{sgn}(\sigma)\, \mathbf{d}f(\vec{\mathbf{v}}_{\sigma(1)})\varphi(\vec{\mathbf{v}}_{\sigma(2)}, \dots, \vec{\mathbf{v}}_{\sigma(k+1)}). \quad (1)$$

Equation 1: The $\sigma(i)$ indicate that these shuffles are those permutations such that within each group, the vectors come with indices in ascending order.

We have $1+k$ possible $(1,k)$-shuffles. As in the discussion following Definition 6.1.12, we use a vertical bar to separate the vector on which the 1-form $\mathbf{d}f$ is being evaluated from the vectors on which the k-form φ is being evaluated, and we use a hat to denote an omitted vector:

$$\begin{gathered}
\vec{\mathbf{v}}_1 \big| \widehat{\vec{\mathbf{v}}}_1, \vec{\mathbf{v}}_2, \dots, \vec{\mathbf{v}}_{k+1} \\
\vec{\mathbf{v}}_2 \big| \vec{\mathbf{v}}_1, \widehat{\vec{\mathbf{v}}}_2, \dots, \vec{\mathbf{v}}_{k+1} \\
\vdots \\
\vec{\mathbf{v}}_{k+1} \big| \vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \dots, \widehat{\vec{\mathbf{v}}}_{k+1},
\end{gathered}$$

which we can write

$$\vec{\mathbf{v}}_i \big| \vec{\mathbf{v}}_1, \dots, \widehat{\vec{\mathbf{v}}}_i, \dots, \vec{\mathbf{v}}_{k+1}$$

for $i = 1, \dots, k+1$. When i is odd, the signature of the permutation is $+1$; when i is even, the signature is -1. So we can rewrite equation 1 as

$$\mathbf{d}f \wedge \varphi(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i-1}\, \mathbf{d}f(\vec{\mathbf{v}}_i)\varphi(\vec{\mathbf{v}}_1, \dots, \widehat{\vec{\mathbf{v}}}_i, \dots, \vec{\mathbf{v}}_{k+1}). \quad (2)$$

Now we can rewrite $\mathbf{d}f$ as $[\mathbf{D}f]$. To evaluate $\mathbf{d}f$ at $\mathbf{0}$, we change the $(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{k+1})$ on the left side of equation 2 to $P_{\mathbf{0}}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{k+1})$:

$$\mathbf{d}f \wedge \varphi\big(P_{\mathbf{0}}(\vec{\mathbf{v}}_1, \dots, \vec{\mathbf{v}}_{k+1})\big) = \sum_{i=1}^{k+1} (-1)^{i-1} \left([\mathbf{D}f(\mathbf{0})]\vec{\mathbf{v}}_i\right)\varphi(\vec{\mathbf{v}}_1, \dots, \widehat{\vec{\mathbf{v}}}_i, \dots, \vec{\mathbf{v}}_{k+1}).$$

Solution A23.1: The $[0]$ in

$$F^{-1}([0]) \subset \mathbb{R}^n \times \operatorname{Mat}(n,n)$$

is the zero matrix among symmetric $n \times n$ matrices.

A23.1 By definition, the space of rigid motions of $\mathbb{R}^n$ is the space of maps $T : \mathbb{R}^n \to \mathbb{R}^n$ of the form $\mathbf{x} \mapsto A\mathbf{x} + \mathbf{b}$, where $A \in \operatorname{Mat}(n,n)$ satisfies $A^\top A = I$. Thus we can think of the space of rigid motions as the subset $F^{-1}([0]) \subset \mathbb{R}^n \times \operatorname{Mat}(n,n)$, where $F(\mathbf{b}, A) = A^\top A - I$, and it is important to consider the codomain of F as the space of symmetric $n \times n$ matrices, not the space of all $n \times n$ matrices. Then F is a map from a space of dimension $n^2 + n$ to a space of dimension $n(n+1)/2$. By part d of Exercise 3.2.11, we know that the derivative of F is surjective at all the points where $A^\top A = I$. So, by Theorem 3.1.10, $F^{-1}([0])$ is a manifold of dimension

$$n + n^2 - \frac{n(n+1)}{2} = \frac{n(n+1)}{2}.$$

A23.3 This is immediate from Heine-Borel. The manifold M is locally a graph. We have seen many times what this means: for every $\mathbf{x} \in M$, there exist

- subspaces $E_1, E_2 \subset \mathbb{R}^n$ with E_1 spanned by p standard basis vectors and E_2 spanned by the other $n-p$ standard basis vectors,
- open subsets $U \subset E_1$ and $V \subset E_2$,
- A C^1 mapping $\mathbf{g}: U \to V$,

such that $M \cap (U \times V) = \Gamma(\mathbf{g})$. Renumber the variables so that E_1 corresponds to the first basis vectors and E_2 to the last, and write $\mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix}$. Choose $r(\mathbf{x}) > 0$ such that $B_{r(\mathbf{x})}(\mathbf{y}) \subset U$, and call $Z_{\mathbf{x}}$ the graph of the restriction of the graph of $\mathbf{g}$ to the concentric ball of radius $r/2$. Then the closure of $Z_{\mathbf{x}}$, which is the graph of the restriction of $\mathbf{g}$ to the closed ball of radius $r/2$, is a compact subset of M.

The $Z_{\mathbf{y}}$ form an open cover of Y, so by Heine-Borel there is a finite subcover: there exist $\mathbf{x}_1, \dots, \mathbf{x}_m$ such that

$$Y \subset Z_{\mathbf{x}_1} \cup \cdots \cup Z_{\mathbf{x}_m}.$$

If each of the $f(\overline{Z}_{\mathbf{x}_i})$ has q-dimensional volume 0, then so does their union, which contains $f(Y)$.

Index

Bold page numbers indicate a page where a term is defined, formally or informally.
Page numbers in italics indicate that the term is used in a theorem, proposition, lemma, or corollary.