Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach, 4th Edition

Notes and Errata

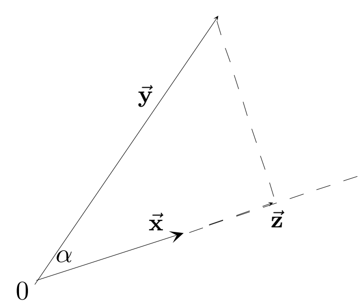
Complete as of March 4, 2011

This list is divided into errors, notes and amplifications, and minor typos.

Many thanks to Itai B., Jimmy Hung, Paul Kirk, David Klein, Manuel López Mateos, David Meyer, John Piliounis, Jeff Rabin, and Walter Alden Tackett, for their contributions to this list.

Errors

Page 70 Figure 1.4.3: **z** is in the wrong place. The figure should be as follows.



Page 283 Exercise 2.36: 2nd line, "satisfies $\sup_{k\neq j} |\lambda_k - \lambda_j| \geq m > 0$" should be "satisfies $\inf_{k\neq j} |\lambda_k - \lambda_j| \geq m > 0$."

Page 311 Proof of proposition 3.2.7: In the 1st line, "in some neighborhood V of $\mathbf{u} \in \mathbb{R}^n$" should be "in some neighborhood V of $\gamma(\mathbf{u}) \in \mathbb{R}^n$." In the 4th line, "open neighborhood of $\gamma^{-1}(\mathbf{u})$" should be "open neighborhood of $\mathbf{u}$."

Page 319 Sentence before example 3.3.11: "be continuous" should be "be differentiable". The partial derivatives in that example are continuous, but not differentiable.

Page 344 Exercise 3.5.18: We should have said that A is a real matrix. (Also, there should be an arrow on $\mathbf{x}^\top$.)

Page 369 Exercise 3.7.8: $x + y \leq 1$, not $x + y = 1$.

Page 388 Exercise 3.8.4: 'Hyperboloid" should be "ellipsoid".

Page 402 The proof of proposition 4.1.16 assumes that if f is integrable, then f^+ and f^- are also. This is only proved on page 425 (corollary 4.3.4). If and when the book is reprinted, we will move this corollary here, with a different (and easier) proof. Corollary 4.3.5 will be moved also, becoming part two of the new corollary. The text begining "Definition 4.1.15 allows us to reduce ... " and ending with the line before proposition 4.1.16 will be replaced by:

Part 1 of corollary 4.1.15 lets us reduce a problem about an arbitrary function to a problem about nonnegative functions. We will use it to prove proposition 4.1.16. If $f : \mathbb{R}^n \to \mathbb{R}$ is any function, we define f^+ and f^- by

$$f^+(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } f(\mathbf{x}) \geq 0 \\ 0 & \text{if } f(\mathbf{x}) < 0 \end{cases}, \qquad f^-(\mathbf{x}) = \begin{cases} -f(\mathbf{x}) & \text{if } f(\mathbf{x}) \leq 0 \\ 0 & \text{if } f(\mathbf{x}) > 0. \end{cases} \tag{4.1.39}$$

Clearly both f^+ and f^- are nonnegative functions, and $f = f^+ - f^-$.

Corollary 4.1.15. *1. A bounded function f with bounded support is integrable if and only if both f^+ and f^- are integrable.*

2. If f and g are integrable, so are $\sup(f, g)$ and $\inf(f, g)$.

Proof. 1. If f is integrable, then by part 4 of proposition 4.1.14, $|f|$ is integrable, and so are $f^+ = \frac{1}{2}(|f| + f)$ and $f^- = \frac{1}{2}(|f| - f)$, by parts 1 and 2. Now assume that f^+ and f^- are integrable. Then $f = f^+ - f^-$ is integrable, again by parts 1 and 2.

2. By part 1 and by proposition 4.1.14, the functions $(f - g)^+$ and $(f - g)^-$ are integrable, so by proposition 4.1.14,

$$\begin{aligned} \inf(f, g) &= \frac{1}{2}\Big(f + g - (f - g)^+ - (f - g)^-\Big), \\ \sup(f, g) &= \frac{1}{2}\Big(f + g + (f - g)^+ + (f - g)^-\Big) \end{aligned} \tag{4.3.5}$$

are integrable. □

Page 405 Proposition 4.1.23: "A set $X \subset \mathbb{R}^n$ should be "a bounded set $X \subset \mathbb{R}^n$".

Page 405 Proposition 4.1.24: The t on the right side of equation 4.1.64 should be $|t|$: $\text{vol}_n(tA) = |t|^n \text{vol}_n(A)$.

Page 406 In keeping with the correction in proposition 4.1.24, many of the t in the proof of that proposition should be $|t|$:

since $\text{vol}_n(tC) = |t|^n \text{vol}_n(C)$ for every cube, this gives

$$\begin{aligned} |t|^n \sum_{C \in \mathcal{D}_N, C \subset A} \text{vol}_n(C) &= \int \sum_{C \in \mathcal{D}_N, C \subset A} \mathbf{1}_{|t|C} \leq L(\mathbf{1}_{|t|A}) \\ &\leq U(\mathbf{1}_{|t|A}) \leq \int \sum_{C \in \mathcal{D}_N, C \cap A \neq \emptyset} \mathbf{1}_{|t|C} \\ &= |t|^n \sum_{C \in \mathcal{D}_N, C \cap A \neq \emptyset} \text{vol}_n(C). \end{aligned} \tag{4.1.67}$$

The outer terms have a common limit as $N \to \infty$, so $L(\mathbf{1}_{|t|A}) = U(\mathbf{1}_{|t|A})$, and

$$\operatorname{vol}_n(tA) = \int \mathbf{1}_{|t|A}(\mathbf{x})|d^n\mathbf{x}| = |t|^n \operatorname{vol}_n(A). \qquad \square \qquad 4.1.68$$

Page 418 Definition 4.2.16: "Let S_1 and S_2 be the sample spaces for two experiments. If $f : S_1 \to \mathbb{R}$ and $g : S_2 \to \mathbb{R} \dots$ " should be "Let S be the sample space for an experiment. If "$f : S \to \mathbb{R}$ and $g : S \to \mathbb{R} \dots$ ".

Page 424 Last 2 lines: "a cube intersecting the support" should be "a (great big) cube containing the support".

Page 425 Equation 4.3.4, second line: the second overbrace with $\geq \epsilon_0$ is inaccurate. It is the sum of the $\operatorname{vol}_n C$ that is $\geq \epsilon_0$.

Page 427 The comment following theorem 4.3.8, saying that the theorem says that integration is translation invariant, is misleading. It only says it for continuous functions. We have added a corollary in section 4.4:

Corollary 4.4.13 (Integration is translation invariant). *If a function $f : \mathbb{R}^n \to \mathbb{R}$ is integrable, then for any $\vec{\mathbf{v}} \in \mathbb{R}^n$, the function $\mathbf{x} \mapsto f(\mathbf{x} - \vec{\mathbf{v}})$ is integrable.*

It follows from theorem 4.4.8.

Page 427 Line immediately after equation 4.3.7: It is not true that the points $\mathbf{x}_N$ and $\mathbf{y}_N$ are necessarily both in the support of f, but at least one must be, since otherwise we would have $|f(\mathbf{x}_N) - f(\mathbf{y}_N)| = |0 - 0|$. Since both are in the same cube C_N, we know that $|\mathbf{x}_N - \mathbf{y}_N \leq 1/2^N$. However, we propose adding a new theorem and proving theorem 4.3.8 from it; see the section "notes and amplifications".

Page 437 Line 4: The statement " ... then at least one of $\mathbf{x}$ and $\mathbf{x} + \mathbf{a}$ is in the interior of a dyadic cube" is incorrect. For instance, $\mathbf{x} = \begin{bmatrix} 0 \\ -\sqrt{2} \end{bmatrix}$ is not in the interior of any dyadic cube, nor is $\mathbf{x} + \mathbf{a} = \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$. To remedy this, we have added two statements on page 432 and changed the proof pages 436–437; see the notes for those pages, under the heading "Notes and amplifications".

Page 480 Second margin note: "then $T^{-1}T(\vec{\mathbf{v}}) \in T^{-1}T(C)$" should be "then $T^{-1}(\vec{\mathbf{v}}) \in T^{-1}T(C)$".

Page 506 Line immediately after equation 4.11.49: "is finite" should be "is convergent".

Page 517 Exercise 4.6 part b: This is not a probability density (because the integral diverges) so the problem doesn't make sense. We suggest replacing this exercise by the following:

What are the expectation, variance, and standard deviation, if they exist, of the random variable $f(x) = x$, for the following probability densities.

a. $\mu(x) = e^{-x}\mathbf{1}_{[0,\infty]}$ b. $\mu(x) = \dfrac{1}{(x+1)^2}\mathbf{1}_{[0,\infty)}$ c. $\mu(x) = \dfrac{2}{(x+1)^3}\mathbf{1}_{[0,\infty)}$

Page 519 Exercise 4.28 is identical to part b of exercise 4.6 and does not make sense. We suggest replacing it by:

a. Show that $\dfrac{1}{x^2}\mathbf{1}_{[1,\infty]}(x)$ is a probability density.

b. Show that the random variable $f(x) = x$ does not have an expectation (i.e., that the expectation is infinite).

c. Show that $\dfrac{2}{x^3}\mathbf{1}_{[1,\infty]}(x)$ is a probability density.

d. Show that the random variable $f(x) = x$ does have an expectation, and compute it. Show that it does not have a variance (again, the variance is infinite).

Page 557 Exercise 5.10: We should have said that the vectors $\mathbf{x}_1, \dots, \mathbf{x}_k$ are linearly independent. In part a, $d(\mathbf{x}, M)$ should be $d(\mathbf{x}_0, M)$.

Page 597 Equation 6.5.7 and the following line should be

$$\det[\vec{F}(\mathbf{x}),\ \vec{\mathbf{v}},\ \vec{\mathbf{w}}] = \vec{F}(\mathbf{x}) \cdot (\vec{\mathbf{v}} \times \vec{\mathbf{w}})$$

gives the signed volume of the 3-parallelogram spanned by the three vectors.

Page 608 Definition 6.6.1, 3rd line: $\mathbf{x} \in M$, not $\mathbf{x} \in X$.

Notes and amplifications

Page 11 Second paragraph of subsection "Is arcsin a function?": We are told that some readers were confused by

"The daughter of," as a function from people to women ... is not a function.

By the first "function", we meant a rule going from people to girls and women. By the second "function", we meant a function in the mathematical sense.

Page 16 Exercise 0.4.2: The word *relation* has a technical meaning: a relation R between X and Y is an arbitrary subset $R \subset X \times Y$. Thus a function $f : X \to Y$ is a special kind of relation, one that associates to any element of X exactly one element of Y (definition 0.4.1). More precisely (definition 0.4.3), it is a subset $\Gamma_f \subset X \times Y$ such that for every $x \in X$, there exists a unique $y \in Y$ with $(x, y) \in \Gamma_f$.

The rule "daughter of" with $X = Y$ the set of women is not a function since it is not everywhere defined: there exist women (x) with no corresponding daughter (y). But it is a relation.

A relation is 1–1 if for all y the set $\{y \mid (x, y) \in R\}$ has *at most* one element. It is onto if for all y the set $\{x \mid (x, y) \in R\}$ has *at least* one element. The relation "daughter of" is 1–1 and onto: every woman has one and only one biological mother.

There are other relations of interest: equivalence relations, order relations, and partial orders, for instance. They are not discussed in this book.

Page 16 Exercise 0.4.6: We should have written "to real nonnegative numbers" rather than "to real positive numbers". We meant positive (≥ 0) as opposed to strictly positive (> 0), but we intended in this book to stick to the terminology "nonnegative" for ≥ 0 and "nonpositive" for ≤ 0.

Page 16 Exercises 0.4.3–0.4.6 : *Transformation* is a synonym for function but it is most often used in the expression *linear transformation*, defined in section 1.3. We should have written *function* or *mapping* here.

Page 60 Equation 1.3.10: The line at angle θ and the line at angle $\theta + \pi$ are the same line, so the reflection matrices should also be the same. Note that $2(\theta + \pi) = 2\theta + 2\pi$ is the same angle as 2θ. That is why 2θ, not θ, appears in the matrix.

Page 60 2nd and 3rd lines from the bottom: To take the projection of a point onto a line, we draw a line from the point perpendicular to the line, and see where on the line it arrives. We have dropped the adjective "orthogonal" (right-angle) in "orthogonal projection"; if we needed to speak of a non-orthogonal projection, we would call it an "oblique projection".

Page 70 Corollary 1.4.4: By "line spanned by $\vec{\mathbf{x}}$" we mean the line formed of all multiples of $\vec{\mathbf{x}}$.

Page 72 Corollary 1.4.8: "Two vectors are orthogonal" means that they form a right angle.

Page 75 Proposition 1.4.14 and the caption to figure 1.4.9: When we speak of the parallelogram "spanned" by $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, we mean the parallelogram with edges $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$, as shown by the figure.

Page 75 David Meyer proposes a different proof of part 1 of proposition 1.4.14, in which one first proves that $\det(RA) = \det(A)$, where R is rotation by angle θ and $A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$:

$$\overbrace{\det\left(\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}\right)}^{\det(RA)} = \det\begin{bmatrix} a_1\cos\theta - a_2\sin\theta & b_1\cos\theta - b_2\sin\theta \\ a_1\sin\theta + a_2\cos\theta & b_1\sin\theta + b_2\cos\theta \end{bmatrix}$$
$$= a_1 b_2(\cos^2\theta + \sin^2\theta) - b_1 a_2(\sin^2\theta + \cos^2\theta)$$
$$= a_1 b_2 - b_1 a_2 = \underbrace{\det\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}}_{\det(A)}$$

(Since the determinant of the rotation matrix is 1, this is a special case of the formula $\det AB = \det A \det B$, which we prove for $n \times n$ matrices in theorem 4.8.4.) Rotate the parallelogram by angle t so that the edge $\vec{\mathbf{a}}$ lies on the x-axis, becoming the vector $\begin{bmatrix} |\vec{\mathbf{a}}| \\ 0 \end{bmatrix}$, where $|\vec{\mathbf{a}}|$ is the length of $\vec{\mathbf{a}}$, and the original edge $\vec{\mathbf{b}}$ becomes whatever the rotation makes it, which we will denote by $\begin{bmatrix} b_1' \\ b_2' \end{bmatrix}$:

$$RA = \begin{bmatrix} |\vec{\mathbf{a}}| & b_1' \\ 0 & b_2' \end{bmatrix}.$$

Then the area of the parallelogram is

$$\underbrace{|\vec{\mathbf{a}}|}_{\text{base}}\ \underbrace{|b_2'|}_{\text{height}} = |\det(RA)| = |\det(A)|.$$

Moreover, since $\begin{bmatrix} b_1' \\ b_2' \end{bmatrix} = \begin{bmatrix} |\vec{\mathbf{b}}|\cos\theta \\ |\vec{\mathbf{b}}|\sin\theta \end{bmatrix}$, where θ is the angle between $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ this gives the formula for the area of a parallelogram:

$$\text{Area} = |\vec{\mathbf{a}}|\,|\vec{\mathbf{b}}|\,|\sin\theta|.$$

PAGE 105 Exercise 1.5.24 deserves a star.

PAGE 286 Mid-page, the statement "$\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$" is not really correct, and we have decided to change definition 3.1.1 of a graph to read

Definition 3.1.1 (Graph). The *graph* $\Gamma(\mathbf{f})$ of a function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ is the set of points in $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \in \mathbb{R}^{n+m}$ such that $\mathbf{f}(\mathbf{x}) = \mathbf{y}$.

It is convenient to denote a point in the graph of a function $\mathbf{f} : \mathbb{R}^k \to \mathbb{R}^{n-k}$ as $\begin{pmatrix} \mathbf{x} \\ \mathbf{f}(\mathbf{x}) \end{pmatrix}$, with $\mathbf{x} \in \mathbb{R}^k$ and $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^{n-k}$. But when this notation is used to describe a manifold, it is misleading, since there is no reason to suppose that the k "active" variables come first, or even that the k active variables at one point of the manifold are the same as the k active variables at another point. How then might we describe a point in the graph of a function $\mathbf{f}:\mathbb{R}^k \to \mathbb{R}^{n-k}$? Here is a accurate if heavy-handed approach, using the "concrete to abstract" linear transformation Φ of definition 2.6.14. Set

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_{n-k} \end{pmatrix}$$

with $\mathbf{f}(\mathbf{a}) = \mathbf{b}$. Suppose that the variables of the domain d of $\mathbf{f}$ are $i_1, \dots, i_k$ and that those of the codomain c are $j_1, \dots, j_{n-k}$, and define

$$\Phi_{\mathrm{d}}(\mathbf{a}) = a_1 \vec{\mathbf{e}}_{i_1} + \cdots + a_k \vec{\mathbf{e}}_{i_k}$$

$$\Phi_{\mathrm{c}}(\mathbf{f}(\mathbf{a})) = b_1 \vec{\mathbf{e}}_{j_1} + \cdots + b_{n-k} \vec{\mathbf{e}}_{j_{n-k}},$$

with all the $\vec{\mathbf{e}}$ in $\mathbb{R}^n$. Then a point $\mathbf{z}$ in the manifold M (i.e., in the graph of $\mathbf{f}$) can be written

$$\mathbf{z} = \Phi_{\mathrm{d}}(\mathbf{a}) + \Phi_{\mathrm{c}}(\mathbf{f}(\mathbf{a})).$$

PAGE 307 Definition 3.2.1: We had thought of $\mathbf{f}$ as a element of a k-dimensional subspace of $\mathbb{R}^n$ and of $[\mathbf{Df}]$ as an element of an $(n-k)$-dimensional subspace of $\mathbb{R}^n$, so that it makes sense to add them, but on further reflection, we decided that what we are really adding is $\Phi_{\mathrm{d}}(\mathbf{a})$ and $\Phi_{\mathrm{c}}(\mathbf{f}(\mathbf{a}))$; see the note for page 286. Thus we propose replacing definition 3.2.1 by

Definition 3.2.1 (Tangent space to a manifold). Let $M \subset \mathbb{R}^n$ be a k-dimensional manifold. If $\mathbf{z} \in M$ is the point

$$\mathbf{z} = \Phi_{\mathrm{d}}(\mathbf{a}) + \Phi_{\mathrm{c}}(\mathbf{f}(\mathbf{a})),$$

with $\mathbf{a} \in \mathbb{R}^k$ and $\mathbf{f}(\mathbf{a}) \in \mathbb{R}^{n-k}$, then the *tangent space* to M at $\mathbf{z}$, denoted $T_{\mathbf{z}}M$, is the graph of the linear transformation $\mathbf{Df}(\mathbf{a})$.

The tangent space is thus the set of points

$$\Phi_{\mathrm{d}}(\dot{\mathbf{a}}) + \Phi_{\mathrm{c}}([\mathbf{Df}(\mathbf{a})]\dot{\mathbf{a}})$$

where $\dot{\mathbf{a}}$ denotes an increment to $\mathbf{a}$.

PAGE 348 The last paragraph is shaky, since V is just a subspace of $\mathbb{R}^n$ and proposition 3.5.15 is stated and proved for Q defined on all of $\mathbb{R}^n$. The following substitute avoids this problem:

If $l > 0$, there exists $\vec{\mathbf{h}}$ with $Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) < 0$. Then

$$f(\mathbf{a} + t\vec{\mathbf{h}}) - f(\mathbf{a}) = Q_{f,\mathbf{a}}(t\vec{\mathbf{h}}) + r(t\vec{\mathbf{h}}) = t^2 Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) + r(t\vec{\mathbf{h}}).$$

Thus

$$\frac{f(\mathbf{a} + t\vec{\mathbf{h}}) - f(\mathbf{a})}{t^2} = Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) + \frac{r(t\vec{\mathbf{h}})}{t^2},$$

and since $\lim_{t\to 0} \frac{r(t\vec{\mathbf{h}})}{t^2} = 0$, we have

$$f(\mathbf{a} + t\vec{\mathbf{h}}) < f(\mathbf{a})$$

for $|t| > 0$ sufficiently small.

PAGE 349 For the same reason as for the proof of theorem 3.6.9, the proof of theorem 3.6.11 is shaky. Here is a fix:

Write

$$f(\mathbf{a} + \vec{\mathbf{h}}) = f(\mathbf{a}) + Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) + r(\vec{\mathbf{h}}) \quad \text{with} \quad \lim_{\vec{\mathbf{h}}\to\vec{\mathbf{0}}} \frac{r(\vec{\mathbf{h}})}{|\vec{\mathbf{h}}|^2} = 0,$$

as in equations 3.6.11 and 3.6.12.

By definition 3.6.10, there exist vectors $\vec{\mathbf{h}}$ and $\vec{\mathbf{k}}$ such that $Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) > 0$ and $Q_{f,\mathbf{a}}(\vec{\mathbf{k}}) < 0$. Then

$$\frac{f(\mathbf{a} + t\vec{\mathbf{h}}) - f(\mathbf{a})}{t^2} = \frac{t^2 Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) + r(t\vec{\mathbf{h}})}{t^2} = Q_{f,\mathbf{a}}(\vec{\mathbf{h}}) + \frac{r(t\vec{\mathbf{h}})}{t^2}$$

is strictly positive for $t\mathbf{e}0$ sufficiently small, and

$$\frac{f(\mathbf{a} + t\vec{\mathbf{k}}) - f(\mathbf{a})}{t^2} = \frac{t^2 Q_{f,\mathbf{a}}(\vec{\mathbf{k}}) + r(t\vec{\mathbf{h}})}{t^2} = Q_{f,\mathbf{a}}(\vec{\mathbf{k}}) + \frac{r(t\vec{\mathbf{k}})}{t^2}$$

is negative for $t \neq 0$ sufficiently small.

PAGE 377 First paragraph: These are not actually the best possible coordinates; we could use the spectral theorem to get rid of the quadratic terms in XY; see equation 5.4.4. We will see in chapter 5 that this can simplify computations.

PAGE 379 We have changed some signs in proposition 3.8.10, to make them compatible with thinking of the unit normal as pointing upwards. The signs in the textbook are not wrong, but correspond to thinking of the chosen normal as being the downward pointing normal.

Proposition 3.8.10 (Putting a surface into "best" coordinates). *Let $U \subset \mathbb{R}^2$ be open, $f : U \to \mathbb{R}$ a C^2 function, and S the graph of f. Let the Taylor polynomial of f at the origin be*

$$z = f\begin{pmatrix} x \\ y \end{pmatrix} = a_1 x + a_2 y + \frac{1}{2}\Big(a_{2,0}x^2 + 2a_{1,1}xy + a_{0,2}y^2\Big) + \cdots . \qquad 3.8.34$$

Set $c = \sqrt{a_1^2 + a_2^2}$. If $c \neq 0$, i.e., if S is not in "best" coordinates, then S is in best coordinates with respect to the coordinates X, Y, Z corresponding to the orthonormal basis

$$\begin{bmatrix} +\frac{a_2}{c} \\ -\frac{a_1}{c} \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{a_1}{c\sqrt{1+c^2}} \\ \frac{a_2}{c\sqrt{1+c^2}} \\ \frac{c}{\sqrt{1+c^2}} \end{bmatrix}, \quad \begin{bmatrix} \frac{-a_1}{\sqrt{1+c^2}} \\ \frac{-a_2}{\sqrt{1+c^2}} \\ \frac{+1}{\sqrt{1+c^2}} \end{bmatrix}. \qquad 3.8.35$$

With respect to these coordinates, S is the graph of Z as a function F of X and Y:

$$F\begin{pmatrix} X \\ Y \end{pmatrix} = \frac{1}{2}\Big(A_{2,0}X^2 + 2A_{1,1}XY + A_{0,2}Y^2\Big) + \cdots , \qquad 3.8.36$$

which starts with quadratic terms, where

$$\begin{aligned} A_{2,0} &= \frac{1}{c^2\sqrt{1+c^2}}\left(a_{2,0}a_2^2 - 2a_{1,1}a_1a_2 + a_{0,2}a_1^2\right) \\ A_{1,1} &= \frac{1}{c^2(1+c^2)}\left(a_1a_2(a_{2,0} - a_{0,2}) + a_{1,1}(a_2^2 - a_1^2)\right) \\ A_{0,2} &= \frac{1}{c^2(1+c^2)^{3/2}}\left(a_{2,0}a_1^2 + 2a_{1,1}a_1a_2 + a_{0,2}a_2^2\right). \end{aligned} \qquad 3.8.37$$

Page 380 Since we changed signs in proposition 3.8.10, to make them compatible with thinking of the unit normal as pointing upwards, we also change the sign of the mean (scalar curvature) in part 2 of proposition 3.8.11:

2. The mean curvature of S at the origin is

$$H(\mathbf{0}) = \frac{1}{2(1+c^2)^{3/2}}\left(a_{2,0}(1+a_2^2) - 2a_1a_2a_{1,1} + a_{0,2}(1+a_1^2)\right). \qquad 3.8.39$$

Page 388 Exercise 3.8.1: One might begin this exercise by computing the curvature of a circle of radius r. The equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ is the equation of an ellipse. Exercise 3.8.2: The equation $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ is the equation of a hyperbola. Exercise 3.8.4: One might begin this exercise by computing the Gaussian curvature of a sphere of radius r.

PAGE 396 The following definition of "support" is perhaps more standard:

Definition 4.1.2 (Support of a function: Supp(f)). The *support* Supp(f) of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the closure of the set

$$\{\, \mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \neq 0 \,\}. \qquad 4.1.9$$

With this definition, the comment page 427 that "compact support" and "bounded support" mean the same thing is correct.

PAGE 399 In definition 4.1.12, we should have specified that f is bounded with bounded support.

PAGE 401 The converse of part 4 of proposition 4.1.14 is false. Consider the function

$$f(x) = \begin{cases} 1 & \text{if } x \in [0,1] \cap \mathbb{Q} \\ -1 & \text{if } x \in [0,1] \cap (\mathbb{R} - \mathbb{Q}) \\ 0 & \text{otherwise} \end{cases}$$

The function $|f| = \mathbf{1}_{[0,1]}$ is integrable, but f is not.

PAGE 427 We propose adding a new theorem, to be called 4.3.8b (to keep subsequent numbering unchanged), and proving theorem 4.3.8 from it:

Theorem 4.3.6. *Any continuous function on $\mathbb{R}^n$ with bounded support is integrable.*

Theorem 4.3.6 follows almost immediately from theorem 4.3.7.

Theorem 4.3.7. *Let $X \subset \mathbb{R}^n$ be compact. A continuous function $f : X \to \mathbb{R}$ is uniformly continuous.*

Proof. Uniform continuity says:

$$(\forall \epsilon > 0)(\exists \delta > 0)\Big(|\mathbf{x} - \mathbf{y}| < \delta \implies |f(\mathbf{x}) - f(\mathbf{y})| < \epsilon\Big).$$

By contradiction, suppose f is not uniformly continuous. Then there exist $\epsilon > 0$ and sequences $i \mapsto \mathbf{x}_i$, $i \mapsto y_i$ such that

$$\lim_{i \to \infty} |\mathbf{x}_i - \mathbf{y}_i| = 0, \quad \text{but for all } i \text{ we have} \quad |f(\mathbf{x}_i) - f(\mathbf{y}_i)| \geq \epsilon.$$

Since X is compact, we can extract a subsequence $j \mapsto \mathbf{x}_{i_j}$ that converges to some point $\mathbf{a} \in X$. Since $\lim_{i \to \infty} |\mathbf{x}_i - \mathbf{y}_i| = 0$, the sequence $j \mapsto \mathbf{y}_{i_j}$ also converges to $\mathbf{a}$.

By hypothesis, f is continuous *at* $\mathbf{a}$, so there exists $\delta > 0$ such that

$$|\mathbf{x} - \mathbf{a}| < \delta \implies |f(\mathbf{x}) - f(\mathbf{a})| < \frac{\epsilon}{3}. \qquad 4.3.7$$

Further, there exists J such that

$$j \geq J \implies |\mathbf{x}_{i_j} - \mathbf{a}| < \delta \quad \text{and} \quad |\mathbf{x}_{i_j} - \mathbf{a}| < \delta.$$

Thus for $j \geq J$ we have

$$|f(\mathbf{x}_{i_j}) - f(\mathbf{y}_{i_j})| \leq |f(\mathbf{x}_{i_j}) - f(\mathbf{a})| + |f(\mathbf{a}) - f(\mathbf{y}_{i_j})| \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon.$$

This is a contradiction. □

Proof of theorem 4.3.6. Now let us prove theorem 4.3.6 from theorem 4.3.7. Let f be continuous with bounded support. Since the support is compact, f is uniformly continuous. Choose ϵ, and use theorem 4.3.7 to find $\delta > 0$ such that

$$|\mathbf{x} - \mathbf{y}| < \delta \implies |f(\mathbf{x}) - f(\mathbf{y})| < \epsilon. \qquad 4.3.8$$

For all N such that $\sqrt{n}/2^N < \delta$, any two points of a cube of $\mathcal{D}_N(\mathbb{R}^n)$ are at most distance δ apart. Thus if $C \in \mathcal{D}_N(\mathbb{R}^n)$, then

$$|f(\mathbf{x}) - f(\mathbf{y})| < \epsilon. \qquad 4.3.9$$

This proves the theorem: f satisfies a much stronger requirement than theorem 4.3.1 requires. Theorem 4.3.1 only requires that the oscillation be greater than ϵ on a set of cubes of total volume $< \epsilon$, wheras in this case for sufficiently large N there are *no* cubes of $\mathcal{D}_N(\mathbb{R}^n)$ with oscillation $\geq \epsilon$. □

PAGE 399 In definition 4.1.12, we should have specified that f is bounded with bounded support.

PAGE 426 The title for proposition 4.3.6 should specify "graph of integrable function".

PAGE 427 Corollary 4.3.9 should be replaced by the following two corollaries:

Corollary 4.3.8 is not just a special case of proposition 4.3.4 because although we could define f on all of $\mathbb{R}^n$ by having it be 0 outside X, we are not requiring that such an extension of f be integrable, and it may not be.

Corollary 4.3.8. *Let $X \subset \mathbb{R}^n$ be compact and let $f : X \to \mathbb{R}$ be continuous. Then the graph $\Gamma_f \subset \mathbb{R}^{n+1}$ has volume 0.*

Proof. Since X is compact, it is bounded, and there is a number A such that the number of cubes $C \in \mathcal{D}_N(\mathbb{R}^n)$ such that $X \cap C \neq \phi$ is at most $A2^{nN}$. Choose $\epsilon > 0$, and use theorem 4.3.7 to find $\delta > 0$ such that if $\mathbf{x}_1, \mathbf{x}_2 \in X$,

$$|\mathbf{x}_1 - \mathbf{x}_2| < \delta \implies |f(\mathbf{x}_1) - f(\mathbf{x}_2| < \epsilon.$$

Further choose N such that $\sqrt{n}/2^N < \delta$, so that for any $C \in \mathcal{D}_N(\mathbb{R}^n)$ and any $\mathbf{x}_1, \mathbf{x}_2 \in C$, we have $|\mathbf{x}_1 - \mathbf{x}_2| < \delta$.

For any $C \in \mathcal{D}_N(\mathbb{R}^n)$ such that $C \cap X \neq \phi$, at most $2^N\epsilon + 1$ cubes of $\mathcal{D}_N(\mathbb{R}^{n+1})$ intersect Γ_f, hence Γ_f is covered by at most $A2^{nN}(2^N\epsilon + 1)$ cubes with total volume

$$\frac{1}{2^{(n+1)N}} A2^{nN}(2^N\epsilon + 1),$$

which can be made arbitrarily small by taking ϵ sufficiently small. □

Corollary 4.3.9. *Let $U \subset \mathbb{R}^n$ be open and let $f : U \to \mathbb{R}^n$ be a continuous function. Then any compact part Y of the graph of f has $(n+1)$-dimensional volume 0.*

Proof. The projection X of Y into $\mathbb{R}^n$ is compact, and the restriction g of f to X satisfies the hypotheses of corollary 4.3.8. □

PAGE 429 The paragraph "polynomials are integrable" is flawed. The statement that a polynomial is integrable over a set of finite volume does not depend on corollary 4.4.10. Corollary 4.4.11 should therefore be moved to section 4.3, where it can be proved using theorem 4.3.10:

Corollary 4.3.13: The polynomial p is of course not integrable, since it does not have bounded support.

Corollary 4.3.13. *Any polynomial function p can be integrated over any set A of finite volume; that is, $p \cdot \mathbf{1}_A$ is integrable.*

Proof. The function $p \cdot \mathbf{1}_A$ meets the conditions of theorem 4.3.10: it is bounded with bounded support and is continuous except on the boundary of A, which has volume 0. □

PAGE 431 Theorem 4.4.4: "Union" should be "countable union", as indicated by the notation $X_1 \cup X_2 \cup \dots$. Theorem 4.4.4 is of course not true for arbitrary unions: any set is the union of its points, which all have measure 0. Thus measure theory depends on distinguishing between countable and uncountable infinities, and could only come after Cantor's work. Indeed Riemann intergration, which doesn't depend on Cantor's work, came before, but Lebesgue itegration, which does, comes after.

PAGE 432 We have added the following statements after the proof of theorem 4.4.4:

Corollary 4.4.4 b. *Let $B_R(\mathbf{0})$ be the ball of radius R centered at $\mathbf{0}$. If for all R the subset $X \subset \mathbb{R}^n$ satisfies $\mathrm{vol}_n(X \cap B_R(\mathbf{0})) = 0$, then X has measure 0.*

Proof. Since $X = \cup_{m=1}^{\infty} (X \cap B_m(\mathbf{0}))$, it is a countable union of sets of volume 0, hence measure 0.

Proposition 4.4.4 c. *Any subspace of $\mathbb{R}^n$ of dimension $k < n$ has measure 0. Any translate of a subspace of $\mathbb{R}^n$ of dimension $k < n$ has measure 0.*

Proof. The first statement follows from proposition 4.3.5 and corollary 4.4.4 b; the second from proposition 4.1.22. □

PAGES 436–437 We suggest replacing the end of the proof, starting 2 lines after equation 4.4.7, with the following:

Choose $\delta > 0$ and apply equation 4.4.7, denoting by $\mathcal{C}_{N_1}$ the finite collection of cubes $C \in \mathcal{D}_{N_1}(\mathbb{R}^n)$ with $\operatorname{osc}_C f \geq \epsilon_1 = \delta/4$. These cubes have total volume less than $\delta/4$. Now let $\mathcal{C}_{N_2}$ be the finite collection of cubes with $\operatorname{osc}_C f \geq \epsilon_2 = \delta/8$; these cubes have total volume less than $\delta/8$. Continue with $\epsilon_3 = \delta/16, \ldots$.

Finally, consider the infinite sequence of open boxes $B_1, B_2, \ldots$ obtained by listing first the interiors of the cubes in $\mathcal{C}_{N_1}$, then those of the cubes in $\mathcal{C}_{N_2}$, etc.

This almost solves our problem: the sums of the volumes of the boxes in our sequence is at most $\delta/4 + \delta/8 + \cdots = \delta/2$. The problem is that discontinuities on the boundary of dyadic cubes may go undetected by oscillation on dyadic cubes: the value of the function over one cube could be 0, and the value over an adjacent cube could be 1; in each case the oscillation over the cube would be 0, but the function would be discontinuous at points on the border between the two cubes.

This is easily dealt with: the union B of all the boundaries of all dyadic cubes has measure 0. To see this, denote by $\delta\mathcal{D}_N(\mathbb{R}^n)$ the union of the boundaries of the dyadic cubes of $\mathcal{D}_N$. Then

1. For each N, the boundary $\delta\mathcal{D}_N(\mathbb{R}^n)$ is a countable union of translates of subspaces of dimension $n-1$, hence has measure 0 by theorem 4.4.4 and proposition 4.4.4 c.
2. The set $B = \cup_{N=1}^{\infty} \delta\mathcal{D}_N(\mathbb{R}^n)$ has measure 0, since it is a countable union of sets of measure 0. □ theorem 4.4.8.

PAGE 438 Exercise 4.4.3: These statements are proved in the text. This exercise could be replaced by:

Prove that any compact subset of $\mathbb{R}^n$ of measure 0 has volume 0. *Hint*: Use the Heine-Borel theorem (theorem A3.3 in appendix A.3).

PAGE 443 2nd line of example 4.5.7: "clearly" could be replaced by "by proposition 4.1.24".

PGAE 462 At the end of definition 4.7.3, "the diameter of P, denoted $\operatorname{diam}(P)$, is the supremum of the distance between points $\mathbf{x}, \mathbf{y} \in P$," not "the maximum distance". Sentence following the definition: "For any bounded function f with bounded support, we can define an upper sum $U_{\mathcal{P}_N}(f)$ and a lower sum $L_{\mathcal{P}_N}(f)$ with respect to any paving ... "

PAGE 659 In equation 6.11.34, the partial derivatives in the gradient are computed only with respect to $\mathbf{x}$.

Minor typos, spelling, etc.

PAGE 18 Last line of first paragraph: $[x]_3$ should be $[x]_{-3}$:
"for the number in equation 0.5.1, $[x]_{-3} = 0.350$; it is not 0.349."

PAGE 75 Caption of figure 1.4.9: To be consistent with our notation, $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ should be $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ should be $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

PAGE 142 In the second line of equation 1.8.11, two open parentheses are not closed. It would be simplest to delete them, changing $(g(\mathbf{a}+\vec{\mathbf{h}}) + f(\mathbf{a})(g(\mathbf{a}+\vec{\mathbf{h}})$ to $g(\mathbf{a}+\vec{\mathbf{h}}) + f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}})$. Similarly, in the third line of the first margin note, change $f(\mathbf{a})(g(\mathbf{a}+\vec{\mathbf{h}})$ to $f(\mathbf{a})g(\mathbf{a}+\vec{\mathbf{h}})$.

PAGE 197 To be consistent with the notation used in sections 2.2 and 2.3, $[\widetilde{A}, \widetilde{\mathbf{0}}]$ should be $[\widetilde{A}\,|\,\widetilde{\mathbf{0}}]$, in two places: in equation 2.5.5 and two lines before that equation.

PAGE 222 Exercise 2.6.9: In the last line of part (a),

$$\{\underline{\mathbf{v}}\}_1, \dots, \underline{\mathbf{v}}_n$$

should be $\underline{\mathbf{v}}_1, \dots, \underline{\mathbf{v}}_n$.

PAGE 337 In the second margin note, "Exercise 3.5.1 asks you to derive two alternative decompositions" should be "Exercise 3.5.1 asks you to derive an alternative decomposition."

PAGE 363 Caption to figure 3.7.10: The transpose of the matrix $[\mathbf{DF}(\mathbf{p})]$ is $(m+n) \times m$, not $(m+n) \times n$.

PAGE 370 Exercise 3.7.12 duplicates theorem 3.7.15.

PAGE 388 In exercises 3.8.1, 3.8.2, 3.8.4, and 3.8.6 we used the same letters to denote variables and a particular point. In exercises 3.8.1 and 3.8.2, we should have asked for the curvature at $\begin{pmatrix} u \\ v \end{pmatrix}$, not $\begin{pmatrix} x \\ y \end{pmatrix}$. In exercises 3.8.4 and 3.8.6, we should have asked for the curvature at $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$.

PAGE 431 First line: "arbitrarily small", not "arbitrary small".

PAGE 434 Proof of theorem 4.4.6: in the last line of the first paragraph, $\text{osc}(f)$ should be $\text{osc}_C(f)$.

PAGE 442 Example 4.5.4, 1st line: "preceding", not "proceeding".

PAGE 478 Exercise 4.8.16: To be consistent in our notation, $\chi_A(A) = 0$ should be $\chi_A(A) = [0]$.

PAGE 478 Exercise 4.8.18, part b: We should have said that A is invertible.

PAGE 519 Exercise 4.30: In the 2d line, $\{\, \mathbf{x} \mid f(\mathbf{x}) \neq g)\mathbf{x})\,\}$ should be $\{\, \mathbf{x} \mid f(\mathbf{x}) \neq g(\mathbf{x})\,\}$

PAGE 557 Exercise 5.10 is lacking two commas: $\vec{\mathbf{x}}_0, \vec{\mathbf{x}}_1, \dots \vec{\mathbf{x}}_k$ should be $\vec{\mathbf{x}}_0, \vec{\mathbf{x}}_1, \dots, \vec{\mathbf{x}}_k$, and $\vec{\mathbf{x}}_1, \dots \vec{\mathbf{x}}_k$ should be $\vec{\mathbf{x}}_1, \dots, \vec{\mathbf{x}}_k$.

PAGE 633 Caption of figure 6.8.3: "Clockwise", not "counterclockwise".

PAGE 694 First equation in equation A4.3: $\vec{\mathbf{0}}$ on the right, not $\mathbf{v0}$.

PAGE 727 2 lines before equation A14.6: "$F\begin{pmatrix} x \\ y \end{pmatrix} = 0$" should be "$\mathbf{F}\begin{pmatrix} x \\ y \end{pmatrix} = 0$".

PAGE 780 We have rewritten exercise A23.2 to avoid any reference to the boundary of a manifold:

A23.2 Prove theorem 5.2.8. Outline: Let M be a manifold. For every $\mathbf{x} \in M$, let $r(\mathbf{x})$ be the supremum of the radii ρ of balls centered at $\mathbf{x}$ such that $M \cap B_\rho(\mathbf{x})$ is the graph of a C' function expessing some variables as functions of others.

a. Show that $r : M \to \mathbb{R}$ is continuous.

b. Show that for any ϵ, the set

$$M_\epsilon = \{\, \mathbf{x} \in M \mid r(\mathbf{x}) \geq \epsilon \quad \text{and} \quad |\mathbf{x}| \leq 2\epsilon \,\}$$

is compact, and that $M = \bigcup_{\epsilon > 0} M_\epsilon$.

c. Show that there exists a sequence of balls $B_{\rho_i}(\mathbf{x}_i)$ such that

$$M \subset \bigcup_i B_{\rho_i}(\mathbf{x}_i) \quad \text{with} \quad \rho_i = \frac{1}{2} r(\mathbf{x}_i).$$

d. Show that if you can parametrize

$$M_i \stackrel{\text{def}}{=} M \bigcup_{i=l}^{N} B_{\rho_i}(\mathbf{x}_i),$$

you can parametrize M_{i+1}.

e. Finish.