Calculus on Manifolds
Solution of Exercise Problems

Yan Zeng

Version 1.0, last revised on 2000-01-10.

Abstract

This is a solution manual of selected exercise problems from *Calculus on manifolds: A modern approach to classical theorems of advanced calculus*, by Michael Spivak. If you would like to correct any typos/errors, please send email to zypublic@hotmail.com.

Contents

1 Functions on Euclidean Space **2**
1.1 Norm and Inner Product . . . 2
1.2 Subsets of Euclidean Space . . . 2
1.3 Functions and Continuity . . . 3

2 Differentiation **3**
2.1 Basic Definitions . . . 3
2.2 Basic Theorems . . . 3
2.3 Partial Derivatives . . . 4
2.4 Derivatives . . . 5
2.5 Inverse Functions . . . 5
2.6 Implicit Functions . . . 6

3 Integration **6**
3.1 Basic Definitions . . . 6
3.2 Measure Zero and Content Zero . . . 6
3.3 Integrable Functions . . . 6
3.4 Fubini's Theorem . . . 8
3.5 Partitions of Unity . . . 10
3.6 Change of Variable . . . 10

4 Integration on Chains **11**
4.1 Algebraic Preliminaries . . . 11
4.2 Fields and Forms . . . 13
4.3 Geometric Preliminaries . . . 15
4.4 The Fundamental Theorem of Calculus . . . 15

5 Integration on Manifolds **17**
5.1 Manifolds . . . 17
5.2 Fields and Forms on Manifolds . . . 19
5.3 Stokes' Theorem on Manifolds . . . 21
5.4 The Volume Element . . . 21
5.5 The Classical Theorems . . . 22

1 Functions on Euclidean Space

1.1 Norm and Inner Product

▶ 1-2.

Proof. Re-examine the proof of Theorem 1-1(2), we see equality holds if and only if $\sum_i x_i y_i = |\sum_i x_i y_i| = |x||y|$. The second equality requires x and y are linearly dependent. The first equality requires $\langle x, y\rangle \geq 0$. Combined, we conclude x and y are linearly dependent and point to the same direction. □

▶ 1-7.

Proof. (a) Theorem 1-2 (4) and (5) establish a one-to-one correspondence between norm and inner product.

(b) If $Tx = 0$, $|x| = |Tx| = 0$, which implies $x = 0$. So T is injective. This further implies T is surjective (Lax [1] page 15). So T is an isomorphism and T^{-1} is well-defined. It's clear that T^{-1} is also norm preserving and hence must enjoy the same properties. □

▶ 1-8.

Proof. Refer to [1] □

▶ 1-9.

Proof. Use matrix to prove T is norm preserving. This will imply T is inner product preserving and hence angle preserving. Also use matrix to check $\langle x, Tx\rangle = \cos\theta|x||Tx|$. □

▶ 1-12.

Proof. Lax [1] page 66, Corollary 4′. □

1.2 Subsets of Euclidean Space

▶ 1-17.

Proof. Step1: We divide the square $[0,1]\times[0,1]$ into four equal squares by connecting $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, $(0, \frac{1}{2})$ and $(1, \frac{1}{2})$. We place on point in each of the squares and make sure no two points are on the same horizontal or vertical line.

· · ·

Step n: We divide each of the squares obtained in Step (n-1) into four equal squares. We place one point in each of the newly obtained squares and make sure no two points of all the points placed so far are on the same horizontal or vertical line.

· · ·

We continue this procedure infinitely and denote by A the collection of all the points placed according to the above procedure. Then $\partial A = [0,1]\times[0,1]$ and A contains at most one point on each horizontal and each vertical line.

□

▶ 1-18.

Proof. Clearly $A \subset [0,1]$. For any $x \in [0,1] - A$ and any interval (a,b) with $x \in (a,b)$, (a,b) must contain a rational point of $[0,1]$. So $(a,b)\cap A \neq \emptyset$ and $(a,b)\cap A^c \neq \emptyset$. This implies $[0,1] - A \subset \partial A$. Since A is open, a boundary point of A cannot be in A. This implies $\partial A \subset [0,1] - A$. Combined, we get $[0,1] - A = \partial A$. □

1.3 Functions and Continuity

2 Differentiation

2.1 Basic Definitions

▶ 2-4.

Proof. (a) If $x = 0$, then $h(t) \equiv 0$; if $x \neq 0$, $h(t) = txg(\frac{x}{|x|})$ by the property $g(0,1) = g(1,0) = 0$ and $g(-x) = -g(x)$. Since h is a linear function, it is differentiable.

(b) $f(x_1, 0) = f(0, x_2) \equiv 0$ by the property $g(0,1) = g(1,0) = 0$ and $g(-x) = -g(x)$. If f is differentiable, $Df(0,0)$ would have to be $(0,0)$. This implies $|x| \cdot g(\frac{x}{|x|}) = f(x) = o(|x|)$ for $x \neq 0$. So g has to be identically 0. □

▶ 2-5.

Proof. Let $g(x_1, x_2) = x_1|x_2|$ with (x_1, x_2) on unit circle. Then for $(x, y) \neq 0$,

$$g\left(\frac{(x,y)}{|(x,y)|}\right) = \frac{x}{\sqrt{x^2+y^2}} \cdot \frac{|y|}{\sqrt{x^2+y^2}} = \frac{x|y|}{x^2+y^2}.$$

Therefore

$$f(x,y) = |(x,y)|g\left(\frac{(x,y)}{|(x,y)|}\right).$$

□

▶ 2-6.

Proof. $f(x, 0) = f(0, y) \equiv 0$. So $\frac{\partial f(0,0)}{\partial x} = \frac{\partial f(0,0)}{\partial y} = 0$. Assume f is differentiable at $(0,0)$, then

$$f(\Delta x, \Delta y) - f(0,0) = \frac{\partial f(0,0)}{\partial x}\Delta x + \frac{\partial f(0,0)}{\partial y}\Delta y + o(\rho) = o(\rho),$$

where $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. This implies $\sqrt{|\Delta x \Delta y|} = o(\sqrt{(\Delta x)^2 + (\Delta y)^2})$. Let $\Delta y = \Delta x$, we get $|\Delta x| = o(\sqrt{2}|\Delta x|)$, contradiction. □

▶ 2-8.

Proof. Note for any $h \in \mathbb{R}^1$ and $\lambda = (\lambda^1, \lambda^2) \in \mathbb{R}^2$, we have

$$\max_{i=1,2}\{|f^i(a+h) - f^i(a) - \lambda^i h|\} \leq |f(a+h) - f(a) - \lambda h| \leq \sqrt{2}\max_{i=1,2}\{|f^i(a+h) - f^i(a) - \lambda^i h|\}.$$

□

2.2 Basic Theorems

▶ 2-12.

Proof. (a)

$$\begin{aligned} f(h,k) &= f(h_1, \cdots, h_n, k_1, \cdots, k_m) \\ &= \sum_{i=1}^n f(0, \cdots, h_i, \cdots, 0, k_1, \cdots, k_m) \\ &= \sum_{i=1}^n \sum_{j=1}^m f(0, \cdots, h_i, \cdots, 0, 0, \cdots, k_j, \cdots, 0) \\ &= \sum_{i=1}^n \sum_{j=1}^m h_i k_j f(0, \cdots, 1, \cdots, 0, 0, \cdots, 1, \cdots, 0). \end{aligned}$$

So there exists $M > 0$, so that $|f(h,k)| \le M \sum_{i,j} |h_i k_j| \le M|(h,k)|^2$. This implies $\lim_{(h,k)\to 0} |f(h,k)|/|(h,k)| = 0$.

(b) $f(a+h, b+k) = f(a, b+k) + f(h, b+k) = f(a,b) + f(a,k) + f(h,b) + f(h,k)$. So

$$\lim_{|(h,k)|\to 0} \frac{|f(a+h,b+k) - f(a,b) - f(a,k) - f(h,b)|}{|(h,k)|} = \lim_{|(h,k)|\to 0} \frac{|f(h,k)|}{|(h,k)|} = 0.$$

This implies $Df(a,b)(x,y) = f(a,y) + f(x,b)$. □

▶ 2-14.

Proof. We note

$$\begin{aligned} & f(a_1+h_1, a_2+h_2, \cdots, a_k+h_k) \\ = \; & f(a_1, \cdots, a_k) + \sum_{i=1}^{k} f(a_1, \cdots, a_{i-1}, h_i, a_{i+1}, \cdots, a_k) \\ & + \sum_{j=2}^{k} \sum_{(x_1,\cdots,x_k) \text{ consists of } j \ h\text{'s and } k-j \ a\text{'s}} f(x_1, \cdots, x_k). \end{aligned}$$

Then use (a). □

▶ 2-15.

Proof. (a) Note det is multi-linear, so we can use Problem 2.14(b).

(b) Define $x_i(t) = (a_{i1}(t), a_{i2}(t), \cdots, a_{in}(t))$. Then $x_i'(t) = (a_{i1}'(t), \cdots, a_{in}'(t))$ by Theorem 2.3(3) and f can be seen as the composition $g \circ x$ with $x(t) = (x_1(t), \cdots, x_n(t))^T \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ and $g(x) = \det(x)$. Theorem 2-2 (chain rule) implies $f'(t) = Dg(x(t)) \circ Dx(t)$. Since $Dg(x)(y) = \sum_{i=1}^n \det \begin{pmatrix} x_1 \\ \cdots \\ y \\ \cdots \\ x_n \end{pmatrix}$ and $Dx = \begin{pmatrix} x_1'(t) \\ \cdots \\ x_n'(t) \end{pmatrix}$, we have

$$f'(t) = Dg(x(t))(Dx(t)) = \sum_{i=1}^n \det \begin{pmatrix} x_1(t) \\ \cdots \\ Dx(t)_i \\ \cdots \\ x_n(t) \end{pmatrix} = \sum_{i=1}^n \det \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \cdots & \cdots & \cdots \\ a_{i1}'(t) & \cdots & a_{in}'(t) \\ \cdots & \cdots & \cdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

(c) Let $A(t) = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \cdots & \cdots & \cdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}$, $s(t) = \begin{pmatrix} s_1(t) \\ \cdots \\ s_n(t) \end{pmatrix}$ and $b(t) = \begin{pmatrix} b_1(t) \\ \cdots \\ b_n(t) \end{pmatrix}$. Then $A(t)s(t) = b(t)$. So $s(t) = A^{-1}(t)b(t)$ is differentiable. Moreover, $A'(t)s(t) + A(t)s'(t) = b'(t)$. So $s'(t) = A^{-1}(t)(b'(t) - A'(t)A^{-1}(t)b(t)) = A^{-1}(t)b'(t) - A^{-1}(t)A'(t)A^{-1}(t)b(t)$. □

2.3 Partial Derivatives

▶ 2-23.

Proof. (b) $f(x) = x^2 1_{\{x>0,y>0\}} - x^2 1_{\{x>0,y<0\}}$. □

2.4 Derivatives

▶ 2-29.

Proof. (c)

$$D_x f(a) = \lim_{t\to 0}\frac{f(a+tx)-f(a)}{t} = \lim_{t\to 0}\frac{f(a+tx)-f(a)-Df(a)(tx)}{|tx|}\cdot\frac{|tx|}{t} + Df(a)(x).$$

Note $\frac{|tx|}{t}$ is bounded for fixed x, and $\lim_{t\to 0}|f(a+tx)-f(a)-Df(a)(tx)|/|tx| = 0$. So $D_x f(a) = Df(a)(x)$. □

▶ 2-30.

Proof. Note no matter $t > 0$ or $t < 0$, $\frac{f(tx)-f(0)}{t} = |x|g(\frac{x}{|x|})$. □

2.5 Inverse Functions

▶ 2-36.

Proof. By Theorem 2-11, for any $x \in A$, there exists an open set V_x containing x and an open set W_x containing $f(x)$ such that $f : V_x \to W_x$ has a continuous inverse $f^{-1} : W_x \to V_x$ which is differentiable. Then $f(A) = \cup_{x\in A} f(V_x) = \cup_{x\in A} W_x$ must be open. For any open subset B of A, $f(B) = (f^{-1})^{-1}(B)$. Since f^{-1} is continuous and B is open, $(f^{-1})^{-1}(B)$ must also be open.

Remark: For a proof without using Theorem 2-11, see [2] Theorem 8.2. □

▶ 2-37.

Proof. We only prove part (a). Part (b) can be similarly proved. Assume f is 1-1, then the rank of Df cannot be 1 in an open set. Indeed, if, for example, $D_1 f(x, y) \neq 0$ for all (x, y) in some open set A, consider $g : A \to \mathbb{R}^2$ defined by $g(x, y) = (f(x, y), y)$. Then $\det Dg \neq 0$ in A. By Inverse Function Theorem, g is 1-1 in an open subset B of A and $g(B)$ is open. So we can take two distinct points w_1 and w_2 in $g(B)$ with the same first coordinate. Suppose $w_1 = (f(x_1, y_1), y_1)$ and $w_2 = (f(x_2, y_2), y_2)$, then we must have $y_1 \neq y_2$ and $f(x_1, y_1) = f(x_2, y_2)$. This is contradictory with f being 1-1. So our claim must be true. (It is also a straightforward corollary of the rectification theorem, Theorem 2-13).

Consequently, for any given $(x, y) \in \mathbb{R}^2$ and any neighborhood of (x, y), there is at least one point (x', y') such that $D_1 f(x', y') = 0$. By the continuity of $D_1 f$, we conclude $D_1 f(x, y) = 0$. Similarly, we can prove $D_2 f(x, y) = 0$. Combined, we have $Df(x, y) = 0$ for any $(x, y) \in \mathbb{R}^2$. This implies f is a constant (Problem 2-22). So f cannot be 1-1 and we get contradiction again. Therefore, f cannot be 1-1.

Remark: The hint is basically Theorem 2.13 in the next section. □

▶ 2-38.

Proof. (a) Assume not, then there exists $x_1, x_2 \in \mathbb{R}$, such that $x_1 \neq x_2$ and $f(x_1) = f(x_2)$. WLOG, assume $x_1 < x_2$, since f is differentiable everywhere, f is continuous. So there exists $y, z \in [x_1, x_2]$ such that $f(y) = \max_{a\in[x_1,x_2]} f(a)$ and $f(z) = \min_{a\in[x_1,x_2]} f(a)$. Since $f'(a) \neq 0$ for all $a \in \mathbb{R}$, at least one of y, z is not equal to x_1 or x_2. WLOG, assume $y \notin \{x_1, x_2\}$. Then $y \in (x_1, x_2)$. By Fermat's theorem, $f'(y) = 0$. Contradiction.

(b) $f'(x, y) = \begin{pmatrix} e^x\cos y & -e^x\sin y \\ e^x\sin y & e^x\cos y\end{pmatrix}$. So $\det f'(x, y) = e^{2x} \neq 0$. But clearly, $f(x, y + 2n\pi) = f(x, y)$ $(n = \pm 1, \pm 2, \cdots)$. So f is not 1-1. □

▶ 2-39.

Proof. $f'(x) = \begin{cases} \frac{1}{2}, & \text{if } x = 0 \\ \frac{1}{2} + 2x\sin\frac{1}{x} - \cos\frac{1}{x}, & \text{if } x \neq 0 \end{cases}$. Note as $x \to 0$, $2x\sin\frac{1}{x} \to 0$ and $\cos\frac{1}{x}$ oscillates between -1 and 1. So in any neighborhood of 0, $f'(x)$ becomes 0 infinitely many times. In particular, in any neighborhood of 0, we can find a point x, so that f' has different signs on the two sides of x. This implies f is not 1-1 in a neighborhood of x. □

2.6 Implicit Functions

▶ 2-40.

Proof. Define $f : \mathbb{R}^1 \times \mathbb{R}^n \to \mathbb{R}^n$ as $f(t, s_1, \cdots, s_n) = (\sum_{j=1}^n a_{j1}(t)s_j - b_1(t), \cdots, \sum_{j=1}^n a_{jn}s_j - b_n(t))$. Then

$$(D_{1+j}f^i(t, s_1, \cdots, s_n))_{1\leq i,j\leq n} = \begin{pmatrix} a_{11}(t) & a_{21}(t) & \cdots & a_{n1}(t) \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix}$$

is non-singular. By Implicit Function Theorem, there is an open interval I containing t and an open set A containing $(s_1, \cdots, s_n)$ such that for each $\bar{t} \in I$ there is a unique $s(\bar{t}) \in A$ with $f(\bar{t}, s(\bar{t})) = 0$ and $s(\bar{t})$ differentiable. □

3 Integration

3.1 Basic Definitions

3.2 Measure Zero and Content Zero

▶ 3-9.

Proof. (b) The set of integers. □

▶ 3-10.

Proof. (a) By definition, $\forall \varepsilon > 0$, there is a finite cover $\{U_1, \cdots, U_n\}$ of A by *closed* rectangles such that $\sum_{i=1}^n v(U_i) < \varepsilon$. Clearly $\bar{C} \subset \cup_{i=1}^n U_i$. So $\partial C = \bar{C} \setminus C^0 \subset \cup_{i=1}^n U_i$ and hence the content of ∂C is 0.

(b) $C = [0, 1] \cap \mathbb{Q}$, then $\partial C = [0, 1]$. □

▶ 3-11.

Proof. Suppose ∂A has measure 0, then by the fact $\partial A = [0, 1] - A$ and $\sum_{i=1}^\infty (b_i - a_i) < 1$, we conclude $[0, 1] = \partial A \cup A$ has a measure less than 1. But $[0, 1]$ is compact, so $[0, 1]$ has a content less than 1, contradictory with Theorem 3-5. □

3.3 Integrable Functions

▶ 3-14.

Proof. Use Theorem 3-8. □

▶ 3-15.

Proof. By definition of content zero, C had content 0 implies $\bar{C}$ has content 0. So $\partial C \subset \bar{C}$ has content zero and C is Jordan-measurable. By the definition of content zero, C is bounded, hence is contained by some closed rectangle A. Finally, by Problem 3-8, C cannot contain any closed rectangle. So for any partition of A, if S is a closed rectangle contained by A, $\min_{x\in S} 1_C(x) = 0$. So $\int_A 1_C(x)dx = L(1_C, P) = 0$. □

▶ 3-16.

Proof. As in Problem 3-10(b), $C = [0, 1] \cap \mathbb{Q}$ works since $\partial C = [0, 1]$. □

▶ 3-17.

Proof. Since C has measure 0, C cannot contain any closed rectangle by Theorem 3-6 and Problem 3-8. So for any sub-rectangle S of A, 1_C obtains 0 on S. This implies $L(1_C, P) = 0$ for any partition P of A. By definition of integration, $\int_A 1_C = 0$. □

▶ 3-18.

Proof. It suffices to show $B_n = \{x : f(x) > 1/n\}$ has content 0. Indeed, for any $\varepsilon > 0$, we can find a partition P of A, such that

$$\varepsilon > U(f,P) = \sum_{S\in P} M_S(f) \geq \sum_{S\in P, S\cap B_n\neq\emptyset} M_S(f) > \sum_{S\in P, S\cap B_n\neq\emptyset} \frac{1}{n}\cdot v(S).$$

So the collection of sets in P that intersect with B_n is a finite cover of B_n and has total content at most $n\varepsilon$. Since ε is arbitrary, we conclude B_n has content 0. □

▶ 3-19.

Proof. Let $A = \{x : f(x) \neq 1_U(x)\}$. By Problem 1-8, $\partial U = [0,1] \setminus U$. So $\forall x \in \partial U \setminus A$, $f(x) = 1_U(x) = 0$. Fix x, by definition of boundary and the fact that U is open, any neighborhood of x must contain an interval $I \subset U$. Since A has measure 0, $I \setminus A \neq \emptyset$. Therefore, any neighborhood of x contains points of $U \setminus A$, on which f takes the value 1. So f is discontinuous at x. That is, the set of discontinuous points of f contains $\partial U \setminus A$. Problem 3-11 shows ∂U does not have measure 0. So $\partial U \setminus A$ does not have measure 0. By Theorem 3-8, f is not integrable on $[0,1]$. □

▶ 3-20.

Proof. Use Problem 3-12 and Theorem 3-8. □

▶ 3-21.

Proof. First of all, ∂C is bounded and closed, hence compact. By Theorem 3-6, A has measure 0 if and only if A has content 0. By Theorem 3-9, it suffices to show: ∂C has content 0 if and only if for every $\varepsilon > 0$, there is a partition P of A such that $\sum_{S\in\mathcal{S}_1} v(S) - \sum_{S\in\mathcal{S}_2} v(S) < \varepsilon$, where S_1 consists of all sub-rectangles intersection C and S_2 all sub-rectangles contained in C.

For necessity, we note ∂C has content 0, which implies $\forall\varepsilon > 0$, we can find finitely many closed rectangles $\{U_1, \cdots, U_n\}$ such that $v(U_1) + \cdots + v(U_n) < \varepsilon$. We can extend $\{U_1, \cdots, U_n\}$ to a partition P of A, then

$$\sum_{S\in\mathcal{S}_1} v(S) - \sum_{S\in\mathcal{S}_2} v(S) \leq \sum_{i=1}^{n} v(U_i) < \varepsilon.$$

For sufficiency, it is clear that the condition implies ∂C has content 0. □

▶ 3-22.

Proof. First note A is necessarily bounded by the definition of Jordan-measurable. Then use Problem 3-21:

$$\int_{A-C} 1 = \int_A 1 - \int_C 1 \leq \sum_{S\in\mathcal{S}_1} v(S) - \sum_{S\in\mathcal{S}_2} v(S) < \varepsilon.$$

□

3.4 Fubini's Theorem

▶ 3-23.

Proof. The key trick consists of two parts: for compact sets, content 0 is equivalent to measure 0; Problem 3-18.

First, we can assume without loss of generality that C is closed. Indeed, denote by $\bar{C}$ the closure of C. Then by definition of content 0, for any $\varepsilon > 0$, there exists a finite cover $\{U_1, \cdots, U_n\}$ of C by closed rectangles such that $\sum_{i=1}^n v(U_i) < \varepsilon$. Since U_i's are closed, they also consist of a cover of $\bar{C}$. In conclusion, $\bar{C}$ has content 0. Moreover, $A' = \{x \in A : \{y \in B : (x, y) \in C\}$ is not of content $0\}$ is contained by $A'' = \{x \in A :\ \{y \in B : (x, y) \in \bar{C}\}$ is not of content 0 $\}$. So for our problem, it suffices to consider $\bar{C}$ and prove A'' has measure 0. Note $C \subset A \times B$ is bounded, so if C is closed, then C is compact.

By Problem 3-15, C is Jordan-measurable and $\int_{A\times B} 1_C = 0$. By Fubini's Theorem, for $\mathcal{L}(x) = L\int_B 1_C$ and $\mathcal{U}(x) = U\int_B 1_C$, we have $\int_{A\times B} 1_C = \int_A \mathcal{U} = \int_A \mathcal{L}$. Again by Problem 3-15, $1_{A'}\mathcal{U} = \mathcal{U}$. Since C is compact and coordinate projection is continuous, for any $x \in A$, $\{y : (x, y) \in C\}$ is compact (Theorem 1-9). Therefore, on the set A', $\mathcal{U} > 0$ by Problem 3-18 and Theorem 3-6. On the other hand, $0 = \int_{A\times B} 1_C = \int_A \mathcal{U} = \int_A 1_{A'}\mathcal{U}$. By Problem 3-18 and the fact $\mathcal{U} > 0$ on A', we conclude A' has measure 0. □

▶ 3-24.

Proof. C is the union of countably many segments: $\{1\} \times [0, 1]$, $\{\frac{1}{2}\} \times [0, \frac{1}{2}]$, $\{\frac{1}{3}\} \times [0, \frac{1}{3}]$, $\{\frac{2}{3}\} \times [0, \frac{1}{3}]$, $\cdots$. So the rectangle $[0, 1] \times [0, \frac{1}{n}]$ will cover all of these segments, except finitely many. These finitely many segments can be covered by finitely many rectangles whose total volume is as small as we want. Therefore, C has content 0. Meanwhile $A' = \mathbb{Q} \cap [0, 1]$. So A' has measure 0 but is not of content 0. □

▶ 3-25.

Proof. Since the set is compact, by Theorem 3.6, we can either prove the claim for content or prove the claim for measure. When $n = 1$, Theorem 3-5 shows $[a_1, b_1] \times \cdots \times [a_n, b_n]$ is not a set of measure 0. Suppose for $n \le N$, $[a_1, b_1] \times \cdots \times [a_n, b_n]$ is not of measure 0, then

$$\int_{[a_1,b_1]\times\cdots\times[a_{N+1},b_{N+1}]} 1 = \int_{[a_{N+1},b_{N+1}]} \int_{[a_1,b_1]\times\cdots\times[a_N,b_N]} 1 > 0$$

by Problem 3-18 and assumption. Mathematical induction concludes the claim is true for any $n \ge 1$. □

▶ 3-26.

Proof. The boundary of A_f consists of four parts: $\Gamma_1 = \{a\} \times [0, f(a)]$, $\Gamma_2 = \{b\} \times [0, f(b)]$, $\Gamma_3 = [a, b] \times \{0\}$, and $\Gamma_4 = \{(x, y) : a \le x \le b, f(x-) \wedge f(x+) \le y \le f(x-) \vee f(x+)\}$. It suffices to show Γ_4 has measure 0. Indeed, by the integrability of f, $A = \{x \in [a, b] : f$ is discontinuous at $x\}$ has measure 0. So $\forall\varepsilon > 0$, there exists open cover $(U_n)_{n\ge1}$ of A, so that $\sum_{n=1}^\infty v(U_n) < \frac{\varepsilon}{4M}$, where $M = \sup_{a\le x\le b} |f(x)|$. Then $(U_n \times (-2M, 2M))_{n\ge1}$ is an open cover of $\{(x, y) : x \in A, f(x-) \wedge f(x+) \le y \le f(x-) \vee f(x+)\}$ and $\sum_{n=1}^\infty v(U_n \times (-2M, 2M)) = \sum_{n=1}^\infty v(U_n) \cdot 4M < \varepsilon$. The set $B = [a, b] - \cup_{n=1}^\infty U_n$ is a compact set that consist of continuous points of f. By uniform continuity of f on B, we can find an open cover of $\{(x, f(x)) : x \in B\}$ whose total volume is as small as we want. Combined, $\Gamma_4 \subset \{(x, f(x)) : x \in \cup_{n=1}^\infty U_n\} \cup \{(x, f(x)) : x \in B\}$ can be covered by open covers of arbitrarily small volume. So Γ_4 has measure 0. This shows A_f is Jordan-measurable. Its area is $\int_a^b f$ can be easily proved by Fubini's Theorem. □

▶ 3-27.

Proof. $\int_a^b \int_a^y f(x, y)dxdy = \int_a^b \int_a^b 1_{\{x\le y\}} f(x, y)dxdy$. $1_{\{x\le y\}}$ is clearly integrable. So $1_{\{x\le y\}} f(x, y)$ is integrable and Fubini's Theorem applies: $\int_a^b \int_a^b 1_{\{x\le y\}} f(x, y)dxdy = \int_a^b \int_a^b 1_{\{x\le y\}} f(x, y)dydx = \int_a^b \int_x^b f(x, y)dydx$. □

▶ 3-28.

Proof. Assume $D_{1,2}f(a) - D_{2,1}f(a) > 0$, then by continuity, there is a rectangle $A = [a_1, b_1] \times [a_2, b_2]$ containing a such that $D_{1,2}f - D_{2,1}f > 0$ on A. By Fubini's Theorem and Problem 3-18,

$$0 < \int_A D_{1,2}f - D_{2,1}f = \int_{a_1}^{b_1} [D_1 f(x, b_2) - D_1 f(x, a_2)] - \int_{a_2}^{b_2} [D_2 f(b_1, y) - D_1 f(a_1, y)] = 0.$$

Contradiction. □

► 3-29.

Proof. Let A be a Jordan-measurable set in the yz-plane. For simplicity and for sake of illustration, we assume A is in the half plane $\{(y, z) : y \geq 0\}$. Let Γ be the set of $\mathbb{R}^3$ obtained by revolving A about the z-axis. Then $(x, y, z) \in \Gamma$ if and only if $(\sqrt{x^2 + y^2}, z) \in A$. So

$$\int_{\mathbb{R}^3} \chi_\Gamma(x, y, z) dx dy dz = \int_{\mathbb{R}^3} \chi_A(\sqrt{x^2 + y^2}, z) dx dy dz.$$

Using the change of variable

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta, \end{cases}$$

we have

$$\int_{\mathbb{R}^3} \chi_\Gamma(x, y, z) dx dy dz = \int_{(\rho,\theta,z)\in[0,\infty)\times[0,2\pi)\times(-\infty,\infty)} \chi_A(\rho, z)\rho d\rho d\theta dz = 2\pi \int_{-\infty}^{\infty} \int_0^{\infty} \chi_A(\rho, z)\rho d\rho dz.$$

□

► 3-32.

Proof. To prove $F'(y) = \int_a^b D_2 f(x, y) dx$, we only need to show for any $c \leq y_1 \leq y_2 \leq d$, $F(y_2) - F(y_1) = \int_{y_1}^{y_2} \int_a^b D_2 f(x, y) dx dy$. Since $D_2 f$ is continuous, it is integrable. So Fubini's Theorem implies

$$\int_{y_1}^{y_2} \int_a^b D_2 f(x, y) dx dy = \int_a^b \int_{y_1}^{y_2} D_2 f(x, y) dy dx = \int_a^b [f(x, y_2) - f(x, y_1)] dx = F(y_2) - F(y_1).$$

In particular, our proof shows Leibnitz's rule holds as far as $D_2 f(x, y)$ is integrable on $[a, b] \times [c, d]$. □

► 3-34.

Proof. By Leibnitz's rule: $D_1 f(x, y) = g_1(x, 0) + \int_0^y D_1 g_2(x, t) dt = g_1(x, 0) + \int_0^y D_2 g_1(x, t) dt = g_1(x, y)$. □

► 3-35.

Proof. The key difficulty is to show any linear transformation is the composition of linear transformations of the type considered in (a). Recall the three linear transformations correspond to three basic manipulations of matrices that reduce an invertible matrix to an identity matrix. So through the linear transformations of the type considered in (a), any identity matrix can be transformed into any invertible matrix. □

3.5 Partitions of Unity

▶ 3-37.

Proof. (a) Since $f \geq 0$, $\int_\varepsilon^{1-\varepsilon} f$ is monotone increasing as $\varepsilon \to 0$. So $\lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f$ exists if and only if $\lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f$ is bounded. Let Φ be a partition of unity, subordinate to an admissible cover $\mathcal{O}$ of $(0,1)$. Because $[\varepsilon, 1-\varepsilon]$ is a compact set, only finitely many $\phi \in \Phi$ are not 0 on $[\varepsilon, 1-\varepsilon]$. So $\int_\varepsilon^{1-\varepsilon} f = \int_\varepsilon^{1-\varepsilon} \sum_{\phi\in\Phi} \phi \cdot f = \sum_{\phi\in\Phi}\int_\varepsilon^{1-\varepsilon} \phi \cdot f \leq \sum_{\phi\in\Phi}\int_{(0,1)} \phi\cdot f$, and $\lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f \leq \sum_{\phi\in\Phi}\int_{(0,1)}\phi\cdot f$. On the other hand, for any finitely many ϕ in Φ, say $\phi_1, \phi_2, \cdots, \phi_n$, there exists $\varepsilon > 0$ such that all of ϕ_i's ($i = 1, \cdots, n$) are 0 outside $[\varepsilon, 1-\varepsilon]$. So

$$\sum_{i=1}^n \int_{(0,1)} \phi \cdot f = \sum_{i=1}^n \int_\varepsilon^{1-\varepsilon} \phi_i \cdot f = \int_\varepsilon^{1-\varepsilon} \sum_{i=1}^n \phi_i \cdot f \leq \int_\varepsilon^{1-\varepsilon} f \leq \lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f.$$

Let $n \to \infty$, we get $\sum_{\phi\in\Phi}\int_0^1 \phi\cdot f \leq \lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f$. So by definition,

$$\textstyle\int_{(0,1)} f \text{ exists} \Longleftrightarrow \displaystyle\sum_{\phi\in\Phi}\int_{(0,1)}\phi\cdot f < \infty \Longleftrightarrow \lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f < \infty \Longleftrightarrow \textstyle\lim_{\varepsilon\to 0}\int_\varepsilon^{1-\varepsilon} f \text{ exists.}$$

(b) This problem is about the distinction between absolute convergence and conditional convergence of the series $\sum_{n=1}^\infty \frac{(-1)^n}{n}$. Look up any standard textbook in analysis. □

▶ 3-38.

Proof. We leave out the details and only describe the main idea. For each n, we can find a number $a_n \in (n, n+1)$, such that $\int_{A_n\cap(-\infty,a_n]} f = \frac{1}{2}\frac{(-1)^n}{n}$ and $\int_{A_n\cap[a_n,\infty)} f = \frac{(-1)^n}{2n}$. For a given small number $\varepsilon > 0$, the partition of unity in $(n, n+1)$ can be chosen in two ways: for Φ, there are $\phi_1 \in \Phi$ and $\phi_2 \in \Phi$ such that $\phi_1 \equiv 1$ on $(a_{n-1}, a_n - \varepsilon)$ and $\phi_2 \equiv 1$ on $(a_n + \varepsilon, a_{n+1})$, and they overlap a little bit around a_n; for Ψ, there are $\psi_1 \in \Psi$ and $\psi_2 \in \Psi$ such that $\psi_1 \equiv 1$ on $(n, n+2-\varepsilon)$ and $\psi_2 \equiv 1$ on $(n+2+\varepsilon, n+4)$, and they overlap a little bit around $n+2$. Then $\sum_{\phi\in\Phi}\int\phi\cdot f \approx \sum_{n=-\infty}^\infty \frac{1}{2}[\frac{(-1)^n}{n} + \frac{(-1)^{n+1}}{n+1}] = \frac{1}{2}\sum_{n=-\infty}^\infty \frac{(-1)^n}{n(n+1)}$ is absolutely convergent, and $\sum_{\psi\in\Psi}\int\psi\cdot f \approx \sum_{n=-\infty}^\infty[\frac{(-1)^n}{n} + \frac{(-1)^{n+1}}{n+1}] = \sum_{n=-\infty}^\infty \frac{(-1)^n}{n(n+1)}$ is absolutely convergent. Clearly they converge to different values. □

3.6 Change of Variable

▶ 3-39.

Proof. Since g is continuous, the set $B = \{x \in A : \det g'(x) = 0\}$ is a closed set. Apply Theorem 3-13 to open set $A' = A \setminus B$, we have $\int_{g(A')} f = \int_{A'}(f\circ g)|\det g'| = \int_A (f\circ g)|\det g'|$. By Sard's Theorem, $g(B)$ has measure 0. So $\int_{g(A)} f = \int_{g(A')\cup g(B)} f = \int_{g(A')} f$. So we still have Theorem 3-13 without the assumption $\det g'(x) = 0$.

Remark: By Theorem 2-11 (Inver Function Theorem), if we stick to the condition $\det g'(x) \neq 0$ in Theorem 3-13, then g is an open mapping. So $g(A)$ is open and talking about integration on $g(A)$ is meaningful by the definition of integrability in the extended sense. But without $\det g'(x) \neq 0$, we cannot guarantee g is still an open mapping and it may not be meaningful, rigorously speaking, to talk about integration on $g(A)$ – we have to check $\partial g(A)$ has measure 0.

However, if B is bounded, it is compact and hence $g(B)$ is compact. So $\partial g(B) \subset g(B)$ has measure 0, and it is meaningful to talk about integration on $\partial g(B)$. □

4 Integration on Chains

4.1 Algebraic Preliminaries

▶ 4-1.

Proof. (a) By Theorem 4-4(3) and induction, when $i_1, \cdots, i_k$ are distinct,

$$\begin{aligned}
\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(e_{i_1}, \cdots, e_{i_k}) &= k! Alt(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k})(e_{i_1}, \cdots, e_{i_k}) \\
&= k! \frac{1}{k!} \sum_{\sigma \in S_k} \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}(e_{\sigma(i_1)}, \cdots, e_{\sigma(i_k)}) \cdot sgn\sigma \\
&= \sum_{\sigma \in S_k} \varphi_{i_1}(e_{\sigma(i_1)}) \cdots \varphi_{i_k}(e_{\sigma(i_k)}) \cdot sgn\sigma \\
&= 1.
\end{aligned}$$

If the factor $\frac{(k+l)!}{k!l!}$ did not appear in the definition of $\wedge$, the right hand side would be $\frac{1}{k!}$.

(b) Suppose $v_l = \sum_{j=1}^n a_{lj} e_j$ $(1 \le l \le k)$, then

$$\begin{aligned}
\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(v_1, \cdots, v_k) &= \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(\sum_{j=1}^n a_{1j}e_j, \cdots, \sum_{j=1}^n a_{kj}e_j) \\
&= \sum_{j=1}^n \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(a_{1j}e_j, \sum_{j=1}^n a_{2j}e_j, \cdots, \sum_{j=1}^n a_{kj}e_j) \\
&= \sum_{j=1}^k \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(a_{1i_j}e_{i_j}, \sum_{j=1}^n a_{2j}e_j, \cdots, \sum_{j=1}^n a_{kj}e_j) \\
&= \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(\sum_{j=1}^k a_{1i_j}e_{i_j}, \sum_{j=1}^n a_{2j}e_j, \cdots, \sum_{j=1}^n a_{kj}e_j) \\
&= \cdots \\
&= \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(\sum_{j=1}^k a_{1i_j}e_{i_j}, \sum_{j=1}^k a_{2i_j}e_{i_j}, \cdots, \sum_{j=1}^k a_{ki_j}e_{i_j}) \\
&= \det(a_{li_j})_{1 \le l \le k, 1 \le j \le k} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(e_{i_1}, \cdots, e_{i_k}) \\
&= \det(a_{li_j})_{1 \le l \le k, 1 \le j \le k}.
\end{aligned}$$

Remark: Combined with Theorem 4-6, this tells us how to calculate the result of a differential form acting on vectors. □

▶ 4-2.

Proof. $\Lambda^n(V)$ has dimension 1. So f^* as a linear mapping from $\Lambda^n(V)$ to $\Lambda^n(V)$ must be a multiplication by some constant c. Let $e_1, \cdots, e_n$ be a basis of V and $\varphi_1, \cdots, \varphi_n$ the dual basis. Then

$$f^*(\varphi_1 \wedge \cdots \wedge \varphi_n)(e_1, \cdots, e_n) = \varphi_1 \wedge \cdots \wedge \varphi_n(f(e_1), \cdots, f(e_n)) = \det(f) \cdot \varphi_1 \wedge \cdots \wedge \varphi_n(e_1, \cdots, e_n)$$

by Theorem 4-6. By multi-linearity and alternating property, $f^*(\varphi_1 \wedge \varphi_n)(x_1, \cdots, x_n) = \det(f)\varphi_1 \wedge \cdots \wedge \varphi_n(x_1, \cdots, x_n)$ is true for any $x_1, \cdots, x_n \in V^n$. □

▶ 4-3.

Proof. Let $v_1, \cdots, v_n$ be an orthonormal basis and $w_i = \sum_{j=1}^n a_{ij} v_j$. By Theorem 4-6, $\omega(w_1, \cdots, w_n) = \det(a_{ij})\omega(v_1, \cdots, v_n) = \det(a_{ij})$. Meanwhile, $g_{ij} = T(w_i, w_j) = T(\sum_{k=1}^n a_{ik}v_k, \sum_{k=1}^n a_{jk}v_k) = \sum_{k=1}^n a_{ik}a_{jk}$. If we denote the matrix (g_{ij}) by G and (a_{ij}) by A, we have $G = AA^T$. So $\det G = (\det A)^2$, i.e.

$$|\omega(w_1, \cdots, w_n)| = \sqrt{\det(g_{ij})}.$$

□

▶ 4-4.

Proof. We first show $f(e_1), \cdots, f(e_n)$ is an orthonormal basis of V. Indeed, $T(f(e_i), f(e_j)) = f^*T(e_i, e_j) = \langle e_i, e_j\rangle = \delta_{ij}$. From this, we can easily show $(f(e_i))_{i=1}^n$ is linearly independent. Since $\dim V = n$, $(f(e_i))_{i=1}^n$ is an orthonormal basis. Since $[f(e_1), \cdots, f(e_n)] = \mu$, $f^*\omega(e_1, \cdots, e_n) = \omega(f(e_1), \cdots, f(e_n)) = \omega(e_1, \cdots, e_n) = 1$. By the multi-linearity and alternating property, $f^*\omega = \det$. □

▶ 4-5.

Proof. Suppose $c^i(0) = \sum_{j=1}^n a_{ij}(t)c^j(t)$ $(0 \le t \le 1)$. If we denote the matrix $(a_{ij}(t))$ by $A(t)$, we have $(c^1(0), \cdots, c^n(0))^T = A(t)(c^1(t), \cdots, c^n(t))^T$. So $\det[(c^1(0), \cdots, c^n(0))^T] = \det A(t) \det[(c^1(t), \cdots, c^n(t))^T]$. Since $A(t)$ is a continuous function of t, so is $\det A(t)$. Because $A(t)$ is non-singular for any $t \in [0,1]$, $\det A(t)$ does not change sign. This implies $[c^1(0), \cdots, c^n(0)] = [c^1(t), \cdots, c^n(t)]$, $\forall t \in [0,1]$. □

▶ 4-6.

Proof. (b) Denote $v_1 \times \cdots \times v_{n-1}$ by z. We first show $v_1, \cdots, v_{n-1}, z$ are linearly independent. Indeed, assume not, then $\langle z, z\rangle = \det \begin{pmatrix} v_1 \\ v_2 \\ \cdots \\ v_{n-1} \\ z \end{pmatrix} = 0$. So $z = 0$. On the other hand, by the linear independence of $v_1, \cdots, v_{n-1}$, there exists $v_n \in \mathbb{R}^n$ such that $v_1, \cdots, v_n$ are linearly independent. As a result, $\langle v_n, z\rangle = \det \begin{pmatrix} v_1 \\ \cdots \\ v_n \end{pmatrix} \neq 0$. Contradiction. So $v_1, \cdots, v_{n-1}, z$ is a basis of $\mathbb{R}^n$. Moreover, for any $\omega \in \Lambda^n(\mathbb{R}^n)$, $\omega(v_1, \cdots, v_{n-1}, z) = \det \begin{pmatrix} v_1 \\ \cdots \\ v_{n-1} \\ z \end{pmatrix} \omega(e_1, \cdots, e_n) = \langle z, z\rangle \omega(e_1, \cdots, e_n)$. So $[v_1, \cdots, v_{n-1}, z]$ is the usual orientation. □

▶ 4-7.

Proof. For any non-zero $\omega \in \Lambda^n(V)$, we can always find a basis $v_1, \cdots, v_n$ so that $\omega(v_1, \cdots, v_n) = 1$. Define a bilinear functional T on $V \times V$ by designating $T(v_i, v_j) = \delta_{ij}$. Then T can be extended to $V \times V$ and is an inner product. Under T, $v_1, \cdots, v_n$ is an orthonormal basis. Suppose $v_1', \cdots, v_n'$ is another basis of V, which is orthonormal under T and $[v_1', \cdots, v_n'] = [v_1, \cdots, v_n]$. Then $\omega(v_1', \cdots, v_n') = \det(a_{ij})\omega(v_1, \cdots, v_n) = \det(a_{ij})$, where $A = (a_{ij})$ is the representation matrix of $v_1', \cdots, v_n'$ under $v_1, \cdots, v_n$. Since $\delta_{ij} = T(v_i', v_j') = T(\sum_{k=1}^n a_{ik}v_k, \sum_{k=1}^n a_{jk}v_k) = \sum_{k=1}^n a_{ik}a_{jk}$, $A = (a_{ij})$ is an orthonormal matrix. Hence $|\det A| = 1$, which implies $\det A = 1$. □

▶ 4-8.

Proof. $v_1 \times \cdots \times v_{n-1}$ is the unique element $z \in V$, such that for any $w \in V$, $T(w, z) = \omega(v_1, \cdots, v_{n-1}, w) = \det \begin{pmatrix} v_1 \\ \cdots \\ v_{n-1} \\ w \end{pmatrix} \omega(e_1, \cdots, e_n) = \det \begin{pmatrix} v_1 \\ \cdots \\ v_{n-1} \\ w \end{pmatrix}$. □

▶ 4-9.

Proof. (a) We only prove $e_1 \times e_1 = 0$ and $e_1 \times e_3 = -e_2$. Suppose $e_1 \times e_1 = z$, then for any $w \in V$, $\langle w, z\rangle = \det \begin{pmatrix} e_1 \\ e_1 \\ w \end{pmatrix} = 0$. So $z = 0$. Suppose $e_1 \times e_3 = a_1e_1 + a_2e_2 + a_3e_3$, then for any $w = b_1e_1 + b_2e_2 + b_3e_3$, $\langle w, z\rangle = a_1b_1 + a_2b_2 + a_3b_3 = \det \begin{pmatrix} e_1 \\ e_3 \\ b_1e_1 + b_2e_2 + b_3e_3 \end{pmatrix} = b_2 \det \begin{pmatrix} e_1 \\ e_3 \\ e_2 \end{pmatrix} = -b_2$. Since b_1, b_2, b_3 are arbitrary, $a_1 = a_3 = 0$ and $a_2 = -1$. That is $e_1 \times e_3 = -e_2$.

(b) Suppose $z = (z^1, z^2, z^3)$, then

$$\langle z, v \times w\rangle = \det \begin{pmatrix} v^1 & v^2 & v^3 \\ w^1 & w^2 & w^3 \\ z^1 & z^2 & z^3 \end{pmatrix} = z^1(v^2w^3 - v^3w^2) + z^2(w^1v^3 - v^1w^3) + z^3(v^1w^2 - w^1v^2).$$

Since z is arbitrary, we conclude $v \times w = (v^2w^3 - v^3w^2, w^1v^3 - v^1w^3, v^1w^2 - w^1v^2)$.

(c) Note $\sin^2\theta = 1 - \cos^2\theta = 1 - \frac{\langle v,w\rangle^2}{|v||w|^2}$. So $|v|^2|w|^2\sin^2\theta = |v|^2|w|^2 - \langle v, w\rangle^2 = |v \times w|^2$, by part (b). $\langle v \times w, v\rangle = \langle v \times w, w\rangle = 0$ is clear by definition of "$\times$".

(d) $\langle v, w \times z\rangle = \langle w, z \times v\rangle = \langle z, v \times w\rangle = \det \begin{pmatrix} v \\ w \\ z \end{pmatrix} = \det \begin{pmatrix} z \\ v \\ w \end{pmatrix} = \det \begin{pmatrix} w \\ z \\ v \end{pmatrix}$. For rest of this part of problem, use (b).

(e) Use (c). □

▶ 4.10.

Proof. Let $V = \text{span}\{w_1, w_{n-1}\}$ and define $\varphi \in \Lambda^{n-1}(V)$ by $\varphi(x_1, \cdots, x_{n-1}) = \det \begin{pmatrix} x_1 \\ \cdots \\ x_{n-1} \\ \frac{w_1 \times \cdots \times w_{n-1}}{|w_1 \times \cdots \times w_{n-1}|} \end{pmatrix}$, $\forall x_1, \cdots, x_{n-1} \in V$. By definition of $w_1 \times w_2 \times \cdots \times w_{n-1}$, $\varphi(w_1, \cdots, w_{n-1}) = |w_1 \times \cdots \times w_{n-1}|$. According to Problem 4-5, it suffices to show φ is the volume element of V determined by $\langle , \rangle$ and some orientation μ. Indeed, $w_1 \times \cdots \times w_{n-1}$ is perpendicular to V, so for any orthonormal basis v_1, $\cdots$, v_{n-1} of V, v_1, $\cdots$, v_{n-1}, $\frac{w_1 \times \cdots \times w_{n-1}}{|w_1 \times \cdots \times w_{n-1}|}$ is an orthonormal basis of $\mathbb{R}^n$. This implies

$$\varphi(v_1, v_{n-1}) = \det \begin{pmatrix} v_1 \\ \cdots \\ v_{n-1} \\ \frac{w_1 \times \cdots \times w_{n-1}}{|w_1 \times \cdots \times w_{n-1}|} \end{pmatrix} = \pm 1.$$

We can choose an orthonormal basis v_1^*, $\cdots$, v_{n-1}^* in such a way that $\varphi(v_1^*, \cdots, v_{n-1}^*) = 1$. Denote $[v_1^*, \cdots, v_{n-1}^*]$ by μ, then by Theorem 4-6, φ is the volume element determined by $\langle , \rangle$ and μ. □

▶ 4-11.

Proof. The property $T(x, f(y)) = T(f(x), y)$ ($\forall x, y \in V$) holds if and only if $T(v_i, f(v_j)) = T(f(v_i), v_j)$ ($1 \le i, j \le n$). Since $f(v_j) = \sum_{k=1}^n a_{jk}v_k$ and $f(v_i) = \sum_{k=1}^n a_{ik}v_k$, $T(v_i, f(v_j)) = \sum_{k=1}^n a_{jk}T(v_i, v_k) = a_{ji}$ and $T(f(v_i), v_j) = \sum_{k=1}^n a_{ik}T(v_k, v_j) = a_{ij}$. So we have $a_{ji} = a_{ij}$. □

4.2 Fields and Forms

▶ 4-13.

Proof. (a) For any $v_p \in \mathbb{R}^n_p$,

$$(g \circ f)_*(v_p) = (D(g \circ f)(p)(v))_{g \circ f(p)} = (Dg(f(p)) \circ Df(p)(v))_{g \circ f(p)} = g_*[(Df(p)(v))_{f(p)}] = g_* \circ f_*(v_p).$$

For any $\omega \in \Lambda^k(\mathbb{R}^p_{g\circ f(p)})$,

$$(g\circ f)^*(\omega) = \omega((g\circ f)_*) = \omega(g_* \circ f_*) = g^*\omega(f_*) = f^* \circ g^*\omega.$$

So $(g\circ f)^* = f^* \circ g^*$.

(b) Apply Theorem 4-10(2) with $k = l = 0$. □

► 4-14.

Proof. Let $p = c(t)$, then $f_*(v) = Df(p)(v) = (\sum_{i=1}^n D_i f'(p)v^i, \cdots, \sum_{i=1}^n D_i f^m(p)v^i)$. Meanwhile, the tangent vector to $f\circ c$ at t is

$$\begin{aligned}((f'\circ c)'(t), \cdots, (f^m \circ c)'(t)) &= (\sum_{i=1}^n D_i f'(p)(c^i)'(t), \cdots, \sum_{i=1}^n D_i f^m(p)(c^i)'(t)) \\ &= (\sum_{i=1}^n D_i f'(p)v^i, \cdots, \sum_{i=1}^n D_i f^m(p)v^i).\end{aligned}$$

So the tangent vector to $f\circ c$ at t is $f_*(v)$. □

► 4-18.

Proof. Let $g(t) = p + tv$, then

$$\begin{aligned}D_v f(p) &= \lim_{t\to 0} \frac{f(p+tv) - f(p)}{t} \\ &= \lim_{t\to 0} \frac{f\circ g(t) - f\circ g(0)}{t} \\ &= Df(g(0))Dg(0) \\ &= (D_1 f(p), \cdots, D_n f(p)) \begin{pmatrix} v^1 \\ \cdots \\ v^n \end{pmatrix} \\ &= \sum_{i=1}^n D_i f(p) v^i \\ &= \langle v, w\rangle.\end{aligned}$$

After normalizing v to unit vector and by Cauchy-Schwartz inequality, $D_v f(p)| \le |w| = D_{\frac{w}{|w|}} f(p)$. This shows the gradient of f at p is the direction in which f is changing fastest at p. □

► 4-20.

Proof. We only describe the intuition. We define a continuous transformation that pastes together the segments AB and CD, where $A = (0,0)$, $B = (1,0)$, $C = (1,1)$, and $D = (0,1)$. The resulting image is a ring. This solves the case $n = 1$.

For general $n \ge 2$, divide the unit cube vertically into strips of equal size, and identify all strips with one while getting the orientation of edges "right". □

► 4-21.

Proof. $\int_{c\circ p} \omega = \int_I^k (c\circ p)^*\omega = \int_{I^k} p^* \circ c^*\omega$. Suppose $c^*\omega = h dx^1 \wedge \cdots \wedge dx^n$. So by change of variable formula and Theorem 4-9,

$$\int_{I^k} p^* \circ c^*\omega = \int_{I^k} h\circ p \;\mathrm{det} p' = \int_{I^k} h\circ p\; |\mathrm{det} p'| = \int_{I^k} h = \int_{I^k} c^*\omega = \int_c \omega.$$

□

4.3 Geometric Preliminaries

▶ 4-24.

Proof. We fix a line L that goes through 0 and partition $[0,1]$ so that each $c([t_{i-1},t_i])$ $(1 \le i \le m)$ is contained on one side of L. Let $p(t) = \{(x,y) : x^2+y^2=1\} \cap Oc(t)$ where O denotes the origin $(0,0)$. Then $c([t_{i-1},t_i])$ and $p([t_{i-1},t_i])$ are contained in the same side of L. Define a singular 2-cube $c_i \subset \mathbb{R}^2 - 0$ as

$$c_i(t,s) = tc(t_i s + t_{i-1}(t-s)) + (1-t)p(t_i s + t_{i-1}(1-s)).$$

Then

$$\partial c_i = [tc(t_{i-1}) + (1-t)p(t_{i-1})] + c(t_i s + t_{i-1}(1-s)) - [tc(t_i) + (1-t)p(t_i)] - p(t_i s + t_{i-1}(1-s)).$$

Define a singular 2-cube $c^2 \subset \mathbb{R}^2 - 0$ by

$$c^2(t,s) = c_i(t, \frac{s-t_{i-1}}{t_i - t_{i-1}}), \text{ if } t_{i-1} \le s < t_i.$$

Then it's easy to see $\partial c^2 = c - c_{1,n}$ for some integer n.

Remark: Basically, the idea is to use homotopy to construct singular 2-cube in $\mathbb{R}^2 - 0$ and paste them together through boundaries. After cancelation, ∂c^2 becomes $c - c_{1,n}$. □

4.4 The Fundamental Theorem of Calculus

▶ 4-26.

Proof.

$$\begin{aligned}
\int_{c_{R,n}} d\theta &= \int_{[0,1]} c_{R,n}^* \left(\frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy \right) \\
&= \int_{[0,1]} \frac{-R\sin 2\pi nt}{R^2} (R\cos 2\pi nt)' dt + \frac{R\cos 2\pi nt}{R^2} (R\sin 2\pi nt)' dt \\
&= \int_{[0,1]} (\sin 2\pi nt)^2 2\pi n + (\cos 2\pi nt)^2 2\pi n \\
&= 2\pi n.
\end{aligned}$$

By Stokes' Theorem, for any 2-chain c in $\mathbb{R}^2 - 0$, $\int_{\partial c} d\theta = \int_c d(d\theta) = 0$. This shows $c_{R,n} \neq \partial c$. □

▶ 4-27.

Proof. By $c - c_1 = \partial c^2$ and Stokes' Theorem, $\int_c d\theta - \int_{c_{1,n}} d\theta = \int_{\partial c^2} d\theta = 0$. So $\int_c d\theta = \int_{c_{1,n}} d\theta = 2\pi n$. This shows n is unique $(= \int_c d\theta / 2\pi)$. □

▶ 4-31.

Proof. Suppose for any $\omega \neq 0$, we can find a chain $c(\omega)$ such that $\int_c \omega \neq 0$. Then if $d^2\omega \neq 0$, we can find a chain c such that $0 \neq \int_c d^2\omega = \int_{\partial c} d\omega = \int_{\partial^2 c} \omega = \int_0 \omega = 0$. Contradiction. To construct $c(\omega)$, suppose $\omega = f dx_1 \wedge \cdots dx_n$, then $f \not\equiv 0$. So we can find a point x_0 such that $f(x_0) \neq 0$. Let $c(\omega)$ be a sufficiently small cube centered at x_0. Then $\int_{c(\omega)} \omega \neq 0$. □

▶ 4-32.

Proof. (a) $c(s,t) = (1-t)c_1(s)+tc_2(s)$, then $c(s,0) = c_1(s)$, $c(s,1) = c_2(s)$, $c(0,t) = c_1(0)$, and $c(1,t) = c_1(1)$. Let $c_3 = c_1(t)$ and $c_4 = c_1(0)$, then $\partial c = c_1 - c_2 + c_3 - c_4$. This implies

$$\int_{\partial c} \omega = \int_{c_1} \omega - \int_{c_2} \omega + \int_{c_3} \omega - \int_{c_4} \omega.$$

If we suppose $\omega = u(x,y)dx + v(x,y)dy$, then $c_3^*(\omega) = u \circ c_3(c_3'(t))'dt + v \circ c_3(c_3^2(t))'dt = 0$. Similarly, $c_4^*(\omega) = 0$. So $\int_{\partial c} \omega = \int_{c_1} \omega - \int_{c_2} \omega$. There are two ways to show $\int_{\partial c} \omega = 0$. First, if ω is well-defined in c, then Stokes' Theorem gives $\int_{\partial c} \omega = \int_c d\omega = 0$, provided ω is closed. Second, if ω is not well-defined everywhere in c, but is exact, say $\omega = d\omega_1$, then Stokes' Theorem gives

$$\int_{\partial c} \omega = \int_{\partial c} d\omega_1 = \int_{\partial^2 c} \omega_1 = 0.$$

Problem 4-21 gives the counter example.

(b) Suppose $\omega = u(x,y)dx + v(x,y)dy$. Choose a point (x_0, y_0) in the domain of ω, and define $\omega_1 = \int_{x_0}^x u(\xi, y_0)d\xi + \int_{y_0}^y v(x,\eta)d\eta$. This is the integration of ω along first $[x_0, x]$ and then $[y_0, y]$. So $d\omega_1 = u(x,y_0)dx + v(x,y)dy + \int_{y_0}^y \frac{\partial}{\partial x} v(x,\eta)d\eta \cdot dx$, which implies $\frac{\partial w_1}{\partial y} = v(x,y)$. By condition, we can also write ω_1 as $\omega_1 = \int_{y_0}^y v(x_0,\eta)d\eta + \int_{x_0}^x u(\xi,y)d\xi$. This is the integration of ω along first $[y_0,y]$ and then $[x_0,x]$. So $d\omega_1$ can also be written as $d\omega_1 = v(x_0,y)dy + u(x,y)dx + \int_{x_0}^x \frac{\partial u}{\partial y}(\xi,y)d\xi \cdot dy$, which implies $\frac{\partial w_1}{\partial x} = u(x,y)$. Combined, we conclude $d\omega_1 = \omega$.

□

▶ 4-33.

Proof. (a) We only show $f(z) = \bar{z}$ is not analytic.

$$\frac{f(z) - f(z_0)}{z - z_0} = \frac{\bar{z} - \bar{z}_0}{z - z_0} = \frac{(x - x_0) - i(y - y_0)}{(x - x_0) + i(y - y_0)}.$$

If $z \to z_0$ along the line $x = x_0$, the limit is -1; if $z \to z_0$ along the line $y = y_0$, the limit is 1. So $\lim_{z \to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ does not exist.

(c) Let $z = x + yi$, then $T(z) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$. If $z_0 = x_0 + y_0 i$ is a complex number and $T(z) = z_0 z = (x_0 + y_0 i)(x + yi) = (x_0 x - y_0 y) + (x_0 y + y_0 x)i$, we must have $ax + by = x_0 x - y_0 y$, $cx + dy = x_0 y + y_0 x$. Since x and y are arbitrary, this implies $a = x_0$, $b = -y_0$, $c = y_0$ and $d = x_0$, or equivalently, $a = d$, $b = -c$. By part(b) Cauchy-Riemann equations:

$$Df(z_0) = \begin{pmatrix} D_x u(z_0) & D_y u(z_0) \\ D_x v(z_0) & D_y v(z_0) \end{pmatrix} = \begin{pmatrix} D_x u(z_0) & D_y u(z_0) \\ -D_y u(z_0) & D_x u(z_0) \end{pmatrix}.$$

$Df(z_0)$ is a multiplication by the complex number $D_x u(z_0) - iD_y u(z_0)$.

(d)

$$\begin{aligned} d(fdz) &= d[(u+iv)(dx+idy)] \\ &= d[udx - vdy + i(vdx + udy)] \\ &= \frac{\partial u}{\partial y}dy \wedge dx - \frac{\partial v}{\partial x}dx \wedge dy + i[\frac{\partial v}{\partial y}dy \wedge dx + \frac{\partial u}{\partial x}dx \wedge dy] \\ &= -[\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}]dx \wedge dy + i(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y})dx \wedge dy. \end{aligned}$$

So $d(fdz) = 0$ if and only if $\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}$, $\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$.

(e) By part(d) and Stokes' Theorem, $\int_c fdz = \int_{\partial c'} fdz = \int_{\partial c'} d(fdz) = 0$.

(f)

$$gdz = \frac{1}{x+yi}(dx+idy) = \frac{x-yi}{x^2+y^2}(dx+idy) = \frac{xdx+ydy}{x^2+y^2} + \frac{xdy-ydx}{x^2+y^2}i = \frac{1}{2}d\log(x^2+y^2) + id\theta.$$

Let $h = \frac{1}{2}\log(x^2+y^2) : \mathbb{C}-0 \to \mathbb{R}$. Then $gdz = id\theta + dh$. So by Problem 4-26,

$$\int_{c_{R,n}} \frac{dz}{z} = \int_{c_{R,n}} id\theta + dh = 2\pi ni + \int_{[0,1]} c_{R,n}^*(dh).$$

Since $c_{R,n}^*(dh) = dc_{R,n}^*(h) = d\frac{1}{2}\log R^2 = 0$. So $\int_{c_{R,n}} \frac{dz}{z} = 2\pi ni$.

(g) By Problem 4-23, Stokes' Theorem and part(d), for some singular 2-cube c, $c_{R_1,n} - c_{R_2,n} = \partial c$ and

$$\int_{c_{R_1,n}-c_{R_2,n}} \frac{f(z)}{z}dz = \int_{\partial c} \frac{f(z)}{z}dz = \int_c d\left(\frac{f(z)}{z}dz\right) = 0.$$

According to part(f), $\forall\varepsilon > 0$, when R is sufficiently small

$$\left|\int_{c_{R,n}} \frac{f(z)}{z}dz - 2\pi inf(0)\right| = \left|\int_{c_{R,n}} \left(\frac{f(z)}{z} - \frac{f(0)}{z}\right)dz\right| \le \varepsilon\left|\int_{c_{R,n}} \frac{dz}{z}\right| = \varepsilon 2\pi n.$$

So $\lim_{R\to 0}\int_{c_{R,n}} \frac{f(z)}{z}dz = 2\pi nif(0)$. By the first part of (g), $\int_{c_{R,n}} \frac{f(z)}{z}dz \equiv 2\pi nf(0)i$. By Problem 4-24 (note we can further require $c^2 \subset \mathbb{C}-0$), $\int_{c-c_{1,n}} \frac{f(z)}{z}dz = \int_{\partial c^2} \frac{f(z)}{z}dz = \int_{c^2} d\left(\frac{f(z)}{z}dz\right) = 0$. So $\int_c \frac{f(z)}{z}dz = \int_{c_{1,n}} \frac{f(z)}{z}dz = 2\pi nif(0)$, i.e. $nf(0) = \frac{1}{2\pi i}\int_c \frac{f(z)}{z}dz = 0$. □

5 Integration on Manifolds

5.1 Manifolds

▶ 5-1.

Proof. We note that ∂M in this problem is interpreted as "boundary of a manifold", not "boundary of a set".

For any $x \in \partial M$, there is an open set U containing x, an open set $V \subset \mathbb{R}^n$, and a diffeomorphism $h : U \to V$ such that

$$h(U\cap\partial M) = h(U\cap M\cap\partial M) = V\cap(\mathbb{H}^k\times\{0\})\cap\{y_k = 0\} = \{y \in V : y^k = y^{k+1} = \cdots = y^n = 0\}.$$

By definition, ∂M is a $(k-1)$-dimensional manifold. Similarly, for any $x \in M - \partial M$, there is an open set U containing x, an open set $V \subset \mathbb{R}^n$, and a diffeomorphism $h : U \to V$ such that

$$h(U\cap(M-\partial M)) = h(U\cap M\cap(M-\partial M)) = V\cap(\mathbb{H}^k\times\{0\})\cap\{y_k > 0\} = \{y \in V : y^k > 0, y^{k+1} = \cdots = y^n = 0\}.$$

Define $V' = V\cap\{y^k > 0\}$ and $U' = h^{-1}(V')$. Then h remains a diffeomorphism from U' to V' and $h(U'\cap(M-\partial M)) = \{y \in V' : y^{k+1} = \cdots = y^n = 0\}$. This shows $M-\partial M$ is a k-dimensional manifold. □

▶ 5-2.

Proof. See, for example, [2] §23, Example 3. □

▶ 5-3. (a)

Proof. We first clarify that "boundary A" means set boundary. We show any point in boundary A belongs to one of the two types of points as illustrated by the problem's counter example.

Indeed, for any point x in boundary A, by the definition of boundary A being an $(n-1)$-dimensional manifold, we can find an open set U of $\mathbb{R}^n$ containing x, and a diffeomorphism $h : U \to V$ ($V := h(U)$) such that $h(U \cap \text{boundary } A) = \{y \in V : y^n = 0\}$. Without loss of generality, we assume U is connected. Then V is also a connected open set. Since $V - V \cap \{y^n = 0\}$ is the disjoint union of two connected open sets, $U - \text{boundary } A = h^{-1}(V - V \cap \{y^n = 0\})$ is also the disjoint union of two connected open sets, say U_1 and U_2.

Since x is a (set) boundary point of A, any neighborhood that contains x must contains points of A. So at least one of U_1 and U_2 must contain points of A. Because A is open, after proper shrinking, we can assume that the U_i that contains points of A is entirely contained by A, $i = 1, 2$.

Then there are two cases to analyze. Case one, both U_1 and U_2 contain points of A. This implies $x \in U \subset [U_1 \cup U_2 \cup \text{boundary } A] \subset N$. That is, x is an interior point of N and hence satisfies condition (M). Case two, only one of U_i's contains points of A. Without loss of generality, we assume U_1 contains points of A but U_2 does not. Then x does not satisfy condition (M) but satisfies condition (M'). Combining these two cases, we can conclude any point in boundary A satisfies either condition (M) or condition (M').

It is clear that any point in A satisfies condition (M). So, we can conclude N is an n-dimensional manifold-with-boundary. □

► 5-4.

Proof. For any $x \in M$, there is an open set U containing x, an open set $V \subset \mathbb{R}^n$, and a diffeomorphism $h : U \to V$ such that

$$h(U \cap M) = V \cap (\{\mathbb{R}^k \times \{0\}\}) = \{y \in V : y^{k+1} = \cdots = y^n = 0\}.$$

Let $p : V \to \mathbb{R}^{n-k}$ be the projection onto the last $(n-k)$ coordinates. Define $g = p \circ h$. Then $g : U \to \mathbb{R}^{n-k}$ is differentiable and $g^{-1}(0) = h^{-1} \circ p^{-1}(0) = h^{-1}(\{y \in V : y^{k+1} = \cdots = y^n = 0\}) = U \cap M$. Since $Dg = Dp \circ Dh$, rank of $Dg(y)$ = rank of $Dp(y) = n - k$ when $g(y) = 0$.

Remark: It's a partial converse because g is only defined locally. □

► 5-5.

Proof. Let X be a k-dimensional subspace of $\mathbb{R}^n$. Choose an orthonormal basis $a_1, \cdots, a_k$ and extend it to be an orthonormal basis $a_1, \cdots, a_k, a_{k+1}, \cdots, a_n$ of $\mathbb{R}^n$. Define h to be the orthogonal transformation that maps X to $\{x \in \mathbb{R}^n : x_{k+1} = \cdots = x_n = 0\}$. Then condition (M) is satisfied by h. □

► 5-6.

Proof. Consider $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ defined by $g(x, y) = y - f(x)$. Suppose f is differentiable, then g is differentiable and $Dg(x, y) = \begin{bmatrix} D_x f(x) & I_m \end{bmatrix}$. So Dg has rank m. By Theorem 5-1, $g^{-1}(0) = \{(x, y) : y = f(x)\}$ is an n-dimensional manifold. Conversely, if $\{(x, y) : y = f(x)\}$ is a manifold, by Theorem 5-2 $f(x)$ is necessarily differentiable. Therefore the graph of f is an n-dimensional manifold if and only if f is differentiable. □

► 5-8. (a)

Proof. Use the fact that in $\mathbb{R}^n$, any open covering of a set has a countable sub-covering, i.e. the second countability axiom for a separable metric space. □

(b)

Proof. For any $x \in$ boundary M (set boundary), $x \in M$ by the closedness of M. Clearly x cannot satisfy condition (M), so it must satisfy condition (M'). This implies $x \in \partial M$ (manifold boundary). Conversely it is always true, by definition, that $\partial M \subset$ boundary M. So boundary $M = \partial M$.

The counter example is already given in Problem 5-3(a). □

(c)

Proof. By part (b), boundary M agrees with ∂M. By part (a) and Problem 5-1, ∂M has measure zero. Then we are done. (Compactness is used for boundedness of M, which is inherent in the definition of Jordan-measurability.) □

5.2 Fields and Forms on Manifolds

▶ 5-9.

Proof. We suppose f is a coordinate system around x and $f(a) = x$. Let $\mathcal{X}$ be the collection of curves in M with $c(0) = x$. We need to show $f_*(\mathbb{R}^k_a) = \text{span}\{c'(0) : c \in \mathcal{X}\}$. Indeed, let $\bar{\mathcal{X}}$ be the collection of curves in $\mathbb{R}^k$ with $\bar{c}(0) = a$, then f establishes a one-to-one correspondence between $\bar{\mathcal{X}}$ and $\mathcal{X}$ by $f : \bar{c} \to f(\bar{c})$. By Problem 4-14, the tangent vector to $c = f(\bar{c})$ at 0 is $f_*(\bar{c}'(0))$. So $f_*(\mathbb{R}^k_a) \supset \text{span}\{c'(0) : c \in \mathcal{X}\}$. For "$c$", note $f_*(e_i) = c_i'(0)$, where $c_i(t) = f(a + te_i)$. □

▶ 5-10.

Proof. We first define an orientation on M. $\forall x \in M$, suppose $f \in \mathcal{C}$ is a coordinate system around x, such that $f(a) = x$ for some $a \in \mathbb{R}^k$. Define $\mu_x = [f_*((e_1)_a), \cdots, f_*((e_k)_a)]$. We need to check such an orientation μ_x is independent of the choice of f. Indeed, if $g \in \mathcal{C}$ is another coordinate patch around x with $g(b) = x$, by the fact $\det(f^{-1} \circ g)' > 0$,

$$[(f^{-1} \circ g)_*((e_1)_b), \cdots, (f^{-1} \circ g)_*((e_k)_b)] = [(e_1)_a, \cdots, (e_k)_a].$$

Apply f_* to both sides, we have

$$[g_*((e_1)_b), \cdots, g_*((e_k)_b)] = [f_*((e_1)_a), \cdots, f_*((e_k)_a)].$$

Second, by the very definition of μ, $\forall f \in \mathcal{C}$, f is clearly orientation-preserving. Moreover, the requirement that any element of $\mathcal{C}$ is orientation-preserving dictates that the orientation has to be defined in this way.

Finally, we show μ is consistent. Suppose $h : W \to \mathbb{R}^n$ is a coordinate patch and $a, b \in W$. Let $x = h(a)$ and $y = h(b)$. If $[h_*((e_1)_a), \cdots, h_*((e_k)_a)] = \mu_{h(a)}$, there exists some $f \in \mathcal{C}$, such that $[h_*((e_1)_a), \cdots, h_*((e_k)_a)] = \mu_x = [f_*((e_1)_\alpha), \cdots, f_*((e_k)_\alpha)]$ $(f(\alpha) = x = h(a))$. This implies $\det(f^{-1} \circ h)'(a) > 0$ (see the argument on page 118-119). Since both f and h are non-singular, $\det(f^{-1} \circ h)'$ does not change sign throughout its domain. In particular, if $h(b) = y$ falls within the range of f, we must have $\det(f^{-1} \circ h)'(b) > 0$, which implies $[h_*((e_1)_b), \cdots, h_*((e_k)_b)] = [f_*((e_1)_\beta), \cdots, f_*((e_k)_\beta)] = \mu_y$ $(h(b) = f(\beta) = y)$. If $h(b) = y$ is not within the range of f, we use a sequence of elements from $\mathcal{C}$ to "connect" x and y, provided M is connected. □

▶ 5-16. Let $g : A \to \mathbf{R}^p$ be as in Theorem 5-1. If $f : \mathbf{R}^n \to \mathbf{R}$ is differentiable and the maximum (or minimum) of f on $g^{-1}(0)$ occurs at a, show that there are $\lambda_1, \cdots, \lambda_p \in \mathbf{R}$, such that

$$(1) D_j f(a) = \sum_{i=1}^{p} \lambda_i D_j g^i(a), j = 1, \cdots, n.$$

Hint: This equation can be written $df(a) = \sum_{i=1}^{n} \lambda_i dg^i(a)$ and is obvious if $g(x) = (x^{n-p+1}, \cdots, x^n)$.

The maximum of f on $g^{-1}(0)$ is sometimes called the maximum of f subject to the **constraints** $g^i = 0$. One can attempt to find a by solving the system of equations (1). In particular, if $g : A \to \mathbf{R}$, we must solve $n + 1$ equations

$$D_j f(a) = \lambda D_j g(a),$$
$$g(a) = 0,$$

in $n + 1$ unknowns $a^1, \cdots, a^n, \lambda$, which is often very simple if we leave the equation $g(a) = 0$ for last. This is **Lagrange's method**, and the useful but irrelevant λ is called a **Lagrangian multiplier**. The following problem gives a nice theoretical use for Lagrangian multipliers.

Proof. First, (1) can be re-written as $Df = (\lambda_1, \cdots, \lambda_p)Dg$, where Df and Dg are the Jacobian matrices of f and g, respectively. If g is the projection: $g(x) = (x^{n-p+1}, \cdots, x^n)$, then $Dg(a) = (0_{p\times(n-p)}, I_{p\times p})$ and $g^{-1}(0) = \{x \in \mathbb{R}^n : x^{n-p+1} = \cdots = x^n = 0\}$. So for any $x \in g^{-1}(0)$, $f(x) = f(x^1, \cdots, x^{n-p}, 0, \cdots, 0)$. When f achieves its maximum (or minimum) on $g^{-1}(0)$ at a, we must have $D_1 f(a) = \cdots = D_{n-p}f(a) = 0$. Define $\lambda_i = D_{n-p+i}f(a)$ $(1 \le i \le p)$, then

$$\begin{aligned} Df(a) &= (0, \cdots, 0, D_{n-p+1}f(a), \cdots, D_n f(a)) \\ &= (0, \cdots, 0, \lambda_1, \cdots, \lambda_p) \\ &= (\lambda_1, \cdots, \lambda_p)(0_{p\times(n-p)}, I_{p\times p}) \\ &= (\lambda_1, \cdots, \lambda_p)Dg(a). \end{aligned}$$

For general g, by Theorem 2-13, there is an open set $U \subset \mathbb{R}^n$ containing a and a differentiable function $h : U \to \mathbb{R}^n$ with differentiable inverse such that $g \circ h(x^1, \cdots, x^n) = (x^{n-p+1}, \cdots, x^n)$. Suppose $h(x_0) = a$, then $f \circ h$ achieves its maximum (or minimum) on $(g \circ h)^{-1}(0)$ at x_0. By previous argument, there exists λ_1, $\cdots$, λ_p such that

$$D(f \circ h)(x_0) = (\lambda_1, \cdots, \lambda_p)D(g \circ h)(x_0).$$

By Theorem 2-2 (chain rule), $D(f \circ h)(x_0) = Df(a)Dh(x_0)$ and $D(g \circ h)(x_0) = Dg(a)Dh(x_0)$. Since $Dh(x_0)$ is invertible, we have $Df(a) = (\lambda_1, \cdots, \lambda_p)Dg(a)$. □

► 5-17. (a) Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be self-adjoint with matrix $A = (a_{ij})$, so that $a_{ij} = a_{ji}$. If $f(x) = \langle Tx, x\rangle = \sum a_{ij}x^i x^j$, show that $D_k f(x) = 2\sum_{j=1}^n a_{kj}x^j$. By considering the maximum of $\langle Tx, x\rangle$ on S^{n-1} show that there is $x \in S^{n-1}$ and $\lambda \in \mathbf{R}$ with $Tx = \lambda x$.

(b) If $V = \{y \in \mathbf{R}^n : \langle x, y\rangle = 0\}$, show that $T(V) \subset V$ and $T : V \to V$ is self-adjoint.

(c) Show that T has a basis of eigenvectors.

(a)

Proof. Define $g(x) = |x|^2 - 1$. Then g is a differentiable function with $g^{-1}(0) = S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. So $g'(x) = 2x$ has rank 1 whenever $g(x) = 0$. S^{n-1} is a compact set, so $f(x) = \langle Tx, x\rangle$ achieves its maximum and minimum on S^{n-1}, say at $x_{\max}$ and $x_{\min}$, respectively. By Problem 5-16, there exists $\lambda \in \mathbb{R}$, such that $Df(x_{\min}) = \lambda Dg(x_{\min})$ and $Df(x_{\max}) = \lambda Dg(x_{\max})$. Note $D_k f(x) = \sum_{j=1}^n a_{kj}x^j + \sum_{i=1}^n a_{ik}x^i = 2\sum_{j=1}^n a_{kj}x^j = 2(x^1, \cdots, x^n)\begin{pmatrix} a_{k1} \\ \cdots \\ a_{kn} \end{pmatrix}$, we can re-write the above equation as

$$(x_0^1, \cdots, x_0^n)\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} = \lambda x_0,$$

where $x_0 = x_{\min}$ or $x_{\max}$. This shows $x_{\min}$ and $x_{\max}$ are eigenvectors of T on S^{n-1}. Since $x_0 \in \{x : \|x\| = 1\}$, $\lambda = \langle Ax_0, x_0\rangle$. This is the so-called Maximum Principle. See *Principle of Applied Mathematics*, by Robert Keener, page 17. □

(b)

Proof. $\forall y \in V$, $\angle x, Ty\rangle = \langle Tx, y\rangle = \lambda\langle x, y\rangle = 0$. So $T(V) \subset V$. Since the corresponding matrix A of T is symmetric, $T : V \to V$ is self-adjoint. □

(c)

Proof. Part(a) shows T has at least one eigenvector x_1. Part(b) shows T remains a self-adjoint linear operator on the $(n-1)$-dimensional space $\{x_1\}^\perp$. So we can find a second eigenvector $x_2 \in \{x_1\}^\perp$. Then $V' = \{y \in \mathbb{R}^n : \langle x_1, y\rangle = \langle x_2, y\rangle = 0\} = \{x_1, x_2\}^\perp = \{x_1\}^\perp \cap \{x_2\}^\perp$ again satisfies the properties enjoyed by V in part(b). So we can find a third eigenvector $x_3 \in V'$. Continue this procedure, we can find a basis consisting of eigenvectors of T. □

5.3 Stokes' Theorem on Manifolds

5.4 The Volume Element

▶ 5-23.

Proof. Suppose $c^*(ds) = \alpha(t)dt$. By definition,

$$\begin{aligned} \alpha(t) &= \alpha(t)dt(1) \\ &= c^*(ds)(1) \\ &= ds(c_*(1)) \\ &= ds(((c^1)', \cdots, (c^n)')) \\ &= \sqrt{\sum_{i=1}^n [(c^i)']^2} ds \left(\frac{((c^1)', \cdots, (c^n)')}{\sqrt{\sum_{i=1}^n [(c^i)']^2}} \right) \\ &= \sqrt{\sum_{i=1}^n [(c^i)']^2}, \end{aligned}$$

where the last "=" is by the definition of length element and the fact that c is orientation-preserving. □

▶ 5-24.

Proof. $dV(e_1, \cdots, e_n) = 1 = dx^1 \wedge \cdots \wedge dx^n(e_1, \cdots, e_n)$. □

▶ 5-25.

Proof. Suppose M is an oriented $(n-1)$-dimensional manifold in $\mathbb{R}^n$. $\forall x \in M$ and let $n(x)$ be the outward unit normal to M. We define $\omega \in \Lambda^{n-1}(M_x)$ by

$$\omega(v^1, \cdots, v^{n-1}) = \det \begin{pmatrix} n(x) \\ v^1 \\ \cdots \\ v^{n-1} \end{pmatrix}, \ \forall v^1, \cdots, v^{n-1} \in M_x.$$

So if $v^1, \cdots, v^{n-1}$ is an orthonormal basis of M_x with $[n(x), v^1, \cdots, v^{n-1}]$ equal to the usual orientation of $\mathbb{R}^n_x$, $\omega(v^1, \cdots, v^{n-1}) = 1$. This shows $dA = \omega$.

On the other hand, M_x is a subspace of $\mathbb{R}^n_x$. So $\Lambda^{n-1}(M_x)$ is a subspace of $\Lambda^{n-1}(\mathbb{R}^n_x)$. This implies dA can be represented as $dA = \alpha_1 dx^2 \wedge \cdots \wedge dx^n + (-1)\alpha_2 dx^1 \wedge dx^3 \wedge \cdots \wedge dx^n + \cdots + (-1)^{n+1}\alpha_n dx^1 \wedge \cdots \wedge dx^{n-1}$.

So $\forall v^1, \cdots, v^{n-1} \in M_x$, by Problem 4-1(b), $dA(v^1, \cdots, v^{n-1}) = \det \begin{pmatrix} \alpha \\ v^1 \\ \cdots \\ v^{n-1} \end{pmatrix}$. So $\alpha_i(-1)^{1+i} = n_i$. □

▶ 5-26.

Proof. (a) Define $c : [a, b] \times [0, 2\pi] \to \mathbb{R}^3$ by $c(x, \theta) = (x, f(x)\cos\theta, f(x)\sin\theta)$. So

$$Dc = \begin{pmatrix} 1 & 0 \\ f'(x)\cos\theta & -f(x)\sin\theta \\ f'(x)\sin\theta & f(x)\cos\theta \end{pmatrix}.$$

Using the notation in the text, we have $E = 1 + (f'(x))^2$, $G = f^2(x)$, $F = 0$. so the area element $dA = \sqrt{1 + (f'(x))^2} f(x) dx \wedge d\theta$ and $\int_S dA = \int_a^b 2\pi f(x)\sqrt{1 + (f'(x))^2} dx$.

(b) Let $a = -1$, $b = 1$, $f(x) = \sqrt{1-x^2}$, then

$$Area(S^2) = \int_{-1}^{1} 2\pi\sqrt{1-x^2}\frac{1}{\sqrt{1-x^2}}dx = 4\pi.$$

□

► 5-27.

Proof. Denote by dV_1 and dV_2 the volume elements of M and $T(M)$, respectively. The problem is reduced to proving $dV_1 = T^*(dV_2)$. Indeed, $\forall x \in M$ and suppose $(v_1, \cdots, v_n)$ is an orthonormal basis of M_x with $[v_1, \cdots, v_n] = \mu_x$. Then $(T(v_1), T(v_2), \cdots, T(v_n))$ is an orthonormal basis of $T(M)_{T(x)}$ as well. To see $[T(v_1), T(v_2), \cdots, T(v_n)] = T(\mu_x)$, note if $v = Aw$, then $T(v) = TAw = TAT^{-1}(Tw)$. Combined, by definition of volume element, we have $T^*(dV_2)(v_1, \cdots, v_n) = dV_2(T(v_1), \cdots, T(v_n)) = 1$. By the uniqueness of volume element on M, we conclude $T^*(dV_2) = dV_1$. □

5.5 The Classical Theorems

References

[1] Peter Lax. *Linear algebra.* John Wiley & Sons, Inc., New York, 1997. 2

[2] J. Munkres. *Analysis on manifolds*, Westview Press, 1997. 5, 17