(MULTIDIMENSIONAL INTEGRAL) CALCULUS

CMI COURSE NOTES AUGUST-NOVEMBER, 2022

T.R. RAMADAS

Preliminaries

The main reference is Spivak, M.: *Calculus on Manifolds*. We wish to cover Chapters 3,4 and (if possible) 5. This book has errors, some of which are in the Errata at the end; you will find lists of errata online as well. Rudin, W.: *Principles of Analysis* ("Baby Rudin") is more carefully written, and extremely comprehensive. It is also notoriously terse and the definition of the multi-dimensional Riemann integral is unnatural. A more recent reference that I liked is C.C. Pugh: *Real Mathematical Analysis*.

You are supposed to know about the *real numbers*. They form a set, conventionally denoted $\mathbb{R}$. The set $\mathbb{R}$ has two important subsets: the integers (denoted $\mathbb{Z}$) and the rationals[1] (denoted $\mathbb{Q}$): we have $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

(1) The set $\mathbb{R}$ has the algebraic structure of a *field*, the rationals $\mathbb{Q}$ form a subfield; the smallest such containing the subset of integers.

(2) Given real numbers a, b we can compare them and only one of the following is true: $a > b$, $a < b$ or $a = b$.

(3) The real numbers form a metric space with the distance between a, b defined to be $|a - b|$. Any real number is a limit of rational numbers. A Cauchy sequence of real numbers has a limit.

An important topological fact: **a subset of real numbers is compact iff it is closed and bounded.**

A nonempty set S of real numbers is said to be *bounded above* if there is a real number ("an upper bound") M such that every member of S is less than or equal to M. In symbols, we write:

$$x \in S \implies x \leq M$$

Given a bounded-above subset S, it has a *least upper bound*, its *supremum*, denoted

$$\sup S$$

characterised by the properties

Date: November 20, 2022.

[1] I mean *the set of rational numbers*; this informal language will recur.

(1) $\sup S$ is an upper bound, and

(2) if M is any upper bound, $S \leq M$.

The notion of the greatest lower bound ($\inf S$) for a bounded-below set S is defined analogously. In particular, given a set S that is bounded above and below (or simply "bounded"), $\inf S$ and $\sup S$ are both defined.

If $a \leq b$ are real numbers, recall that $[a, b]$ denotes the closed interval consisting of real numbers x satisfying $a \leq x \leq b$. In "set notation" we write:

$$[a, b] \equiv \{x \in \mathbb{R} | a \leq x \leq b\}.$$

From now on functions are real-valued, unless otherwise specified.

1. Recap of Riemann integration in dimensions 1

We will be concerned, for the most part, with functions defined on closed bounded intervals. Such an interval is of the form:

$$I = [a, b] \equiv \{x \in \mathbb{R} | a \leq x \leq b\}$$

The *length* of the interval I is $\ell(I) \equiv b - a$; we will usually suppose that $a < b$ so that $\ell(I) > 0$. A function f is said to be *bounded* if the image $f(I) \equiv \{f(x) | x \in I\}$ is bounded. That is, if there are real numbers m and M such that $m \leq f(x) \leq M$ for all $x \in I$. If f bounded, then its restriction to any closed subinterval J of I is bounded, Set

$$M_J(f) \equiv \sup\{f(x) | x \in J\}$$
$$m_J(f) \equiv \inf\{f(x) | x \in J\}$$

If I is a closed and bounded interval and f is continuous then $f(I)$ is a closed bounded interval. In particular, a continuous real-valued function f on a closed bounded interval "attains its maximum and minimum values". That is, there exist points x_{min} and x_{max} in I (not necessarily unique) such that $f(x_{min}) \leq f(x) \leq f(x_{max})$ for all $x \in I$.

Let us recall the basics of the Riemann integral, following Rudin, whose notation also we mostly copy. In high-school calculus you have encountered possibly heuristic definitions of the integral of a continuous function f on an interval $[a, b]$, denoted

$$\int_a^b f(x)dx$$

You would have learned how to compute such integrals. You would have learned also that

(1) If $f \geq 0$, then $\int_a^b f(x)dx$ has an interpretation as the "area under the graph", where the latter is also heuristically defined.

(2) If we define the function F on $[a,b]$ by

$$F(y) = \int_a^y f(x)dx$$

then F is differentiable on (a,b), and $F'(y) = f(y)$.

The Riemann integral is a formalisation of the notion of an integral, and (together with its extension to higher dimensions) this equips us to give a meaning to and prove rigorously statements such as (1) and (2) above.

If $a < b$, by a *partition* P of the interval $I = [a,b]$ we mean a sequence $a = t_0 < t_1 < \cdots < t_k = b$ (for some $k \geq 1$ which will depend on the partition[2]). Set

$$\Delta t_i = t_i - t_{i-1}$$

We make the important

Remark 1.1. A partition P determines finitely many closed intervals

$$I_1, \ldots, I_i \equiv [t_{i-1}, t_i], \ldots, I_k$$

Note that $\ell(I_i) > 0$, $I = \cup_i I_i$ and any two of the intervals intersect at at most one point (Equivalently, $\sum_i \ell(I_i) = b - a$.) Conversely, any set of intervals as above determines a partition of I.

The remark is left as an exercise.

Given a partition P, and a (bounded) function f, define[3]

$$U(P,f) = \sum_i M_{I_i}(f)\ell(I_i)$$

$$L(P,f) = \sum_i m_{I_i}(f)\ell(I_i)$$

Note that $m_{[a,b]}(f)(b-a) \leq L(P,f) \leq U(P,f) \leq M_{[a,b]}(f)(b-a)$.

A partition R is said to be a *refinement* of a partition P if each interval corresponding to R is wholly contained in an interval corresponding to P. The proof of the following claim is left as an exercise: If R is a refinement of P, then:

$$L(P,f) \leq L(R,f)$$

$$U(R,f) \leq U(P,f)$$

(Informally speaking, "under refinement, the lower sum increases and the upper sum decreases".) This yields the following chain of inequalities (where the parts outside the underbraces follow from definitions):

$$m_{[a,b]}(f)(b-a) \leq \underbrace{L(P,f) \leq L(R,f)} \leq \underbrace{U(R,f) \leq U(P,f)} \leq M_{[a,b]}(f)(b-a)$$

[2]Warning: The term partition is used by Rudin, and differs slightly from the usual usage in set theory. We will not have occasion to use the set-theoretic version.

[3]The notation $U(P,f),\ L(P,f)$ is chosen following Rudin. Spivak, and many others, write $U(f,P),\ L(f,P)$.

Lemma 1.2. *Any two partitions have a common refinement.*

The proof (omitted) is not difficult but requires some book-keeping.

Exercise Let $I = [0, 10]$ and let $P = 0 < 1 < 3 < 5 < 7 < 9 < 10$, $Q = 0 < 2 < 4 < 6 < 8 < 10$. Write down a common refinement R.

An important consequence is

Lemma 1.3. *Given any two partitions P, Q of I and a bounded f, we have*

$$L(P, f) \leq U(Q, f) \ .$$

Proof. To see this let R be a common refinement; then

$$L(P, f) \leq L(R, f) \leq U(R, f) \leq U(Q, f)$$

□

Define the *upper Riemann integral* and the *lower Riemann integrals* respectively by

$$\overline{\int_a^b} f(x)dx = \inf_P U(P, f)$$
$$\underline{\int_a^b} f(x)dx = \sup_P L(P, f)$$

Start with Lemma 1.3, and take the infimum of all partitions Q while keeping the partition P fixed to get:

$$L(P, f) \leq \inf_Q U(Q, f) \underset{def}{=} \overline{\int_a^b} f(x)dx$$

Now taking the supremum over all P we get

$$\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx \tag{A}$$

This holds as long as f is bounded.

We could now define the Riemann integral, but we motivate the definition by stating and proving the

Theorem 1.4. *If f is continuous on $[a, b]$ then*

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$$

Proof. Let $\epsilon > 0$ is given. By uniform continuity of f (see below) there exists $\delta > 0$ depending on ϵ such that $|f(x) - f(y| < \epsilon$ whenever x, y are points in I with $|x - y| < \delta$. Let $P_\epsilon = \cup_i I_i$ be a partition of $[a, b]$ such that $\ell(I_i) < \delta$ for each i.

Fix an i. Let $x_{max,i} \in I_i$ be such that $f(x_{max,i}) = M_{I_i}(f)$ and $x_{min,i} \in I_i$ such that $f(x_{min,i}) = m_{I_i}(f)$. Then

$$M_{I_i}(f) - m_{I_i}(f) = f(x_{max,i}) - f(x_{min,i}) = |f(x_{max,i}) - f(x_{min,i})| < \epsilon$$

As a consequence

$$\begin{aligned} U(P_\epsilon, f) - L(P_\epsilon, f) &= \sum_i M_{I_i}(f)\ell(I_i) - \sum_i m_{I_i}(f)\ell(I_i) \\ &= \sum_i \{M_{I_i}(f) - m_{I_i}(f)\}\ell(I_i) \\ &< \epsilon \sum_i \ell(I_i) = \epsilon\ell(I) \end{aligned}$$

We conclude: *Given any $\epsilon > 0$, there is a partition P_ϵ of $[a, b]$ such that $U(P_\epsilon, f) < L(P_\epsilon, f) + \epsilon(b - a)$.*

It follows that

$$\begin{aligned} \overline{\int_a^b} f(x)dx &\underset{def}{=} \inf_P U(P, f) \leq U(P_\epsilon, f) < L(P_\epsilon, f) + \epsilon(b - a) \\ &\leq \sup_P L(P, f) + \epsilon(b - a) \underset{def}{=} \underline{\int_a^b} f(x)dx + \epsilon(b - a) \end{aligned}$$

from which we conclude that if f is continuous then

$$\overline{\int_a^b} f(x)dx < \underline{\int_a^b} f(x)dx + \epsilon(b - a) \tag{B}$$

with arbitrary $\epsilon > 0$.

Putting together the inequalities (A) and (B), we see that if f is continuous, then for arbitrary $\epsilon > 0$ we have

$$\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx < \underline{\int_a^b} f(x)dx + \epsilon(b - a) \tag{C}$$

This proves the Theorem. □

Continuity and Uniform Continuity

Before moving on, we pause to make sure you understand continuity and uniform continuity. Let us introduce some ad-hoc terminology to aid us. Given $\delta > 0$, and two real numbers x, y, let us agree that "x and y are δ-close" means $|x - y| < \delta$.

Consider a real-valued function f defined on $\mathbb{R}$. Recall that f is

- *continuous* if for every real number x and $\epsilon > 0$, there is a $\delta > 0$ (depending on x *and* ϵ) such that $f(y)$ and $f(x)$ are ϵ-close provided y and x are δ-close.

- *uniformly continuous* if for every $\epsilon > 0$, there is a $\delta > 0$ (depending *only on* ϵ) such that $f(y)$ and $f(x)$ are ϵ-close provided y and x are δ-close.

The function $f(x) = x$ is uniformly continuous on $\mathbb{R}$ but the function $f(x) = x^2$ is not.

Given a function on a closed/open/"clopen" interval $J \subset \mathbb{R}$, one can make the same definitions, remembering in the above definitions that x, y should belong to J for $f(x), f(y)$ to make sense. Important point: a function that is not uniformly continuous on an interval J' might become uniformly continuous when you restrict to a subinterval J.

You **should know** a proof of the following

Theorem 1.5. *Any continuous function f defined on a closed interval $I = [a, b]$ is uniformly continuous.*

Proof. Let $\epsilon > 0$ be given. For each $x \in I$, there is a $\delta_x > 0$ such that $f(y)$ is $\epsilon/2$-close to $f(x)$ provided y belongs to the interval $I \cap (x - \delta_x, x + \delta_x)$. Let $I_x = I \cap (x - \delta_x/2, x + \delta_x/2)$. Since I is compact, finitely many of the I_x "cover" I. That is to say, there are points $x_1, \ldots x_n$ in I such that

$$I = \cup_i I_{x_i}$$

Suppose now x, y are δ-close, where

$$\delta \equiv \frac{1}{2} min\{\delta_{x_1}, \ldots, \delta_{x_n}\}$$

Since the sets I_{x_i} cover I, x belongs to at least one of them, say I_{x_1}. Then $|y - x_1| < |y - x| + |x - x_1| < \delta_{x_1}$ so that $|f(y) - f(x_1)| < \epsilon/2$. This yields

$$|f(y) - f(x)| \leq |f(y) - f(x_1)| + |f(x_1) - f(x))| < \epsilon$$

□

Finally, we can make the important

Definition 1.6. Given a (bounded, real-valued) function f on a bounded closed interval $I = [a, b]$, we say that f is *Riemann integrable* if the upper and lower Riemann integrals of f are equal, in which we define the *Riemann integral* of f to be

$$\underbrace{\int_a^b f(x)dx}_{Notation} = \underline{\int_a^b} f(x)dx = \overline{\int_a^b} f(x)dx$$

Lemma 1.7. ("Cauchy criterion" for Riemann integrability). *Let f be a bounded function on an interval $I = [a, b]$. Then f is Riemann integrable iff for every $\epsilon > 0$ there exists a partition R of I such that*

$$U(R, f) - L(R, f) < \epsilon$$

Proof. We have

$$0 \leq \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \leq U(R, f) - L(R, f)$$

for any R, which shows that the if the condition of the Lemma holds, then

$$\overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx = 0$$

Conversely, suppose $\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx$, and let $\epsilon > 0$ be given. There exist partitions P, Q such that

$$U(P, f) - \overline{\int_a^b} f(x)dx < \frac{\epsilon}{2}$$
$$\underline{\int_a^b} f(x)dx - L(Q, f) < \frac{\epsilon}{2}$$

This yields

$$U(P, f) - L(Q, f) < \epsilon$$

Let R be a common refinement of P, Q. We have then:

$$U(R, f) - L(R, f) \leq U(P, f) - L(Q, f) < \epsilon$$

□

We have seen that continuous functions are Riemann-integrable. *The characterisation of Riemann-integrable functions is a subtle and (at this stage) theoretical business, which we defer.*

Here are some basic properties of Riemann integrable functions and the Riemann integral. **All proofs are left as exercises**; with help from Rudin if necessary.

Proposition 1.8. *Let f be a bounded real-valued function on I. If f is continuous except at finitely many points in I, then it is Riemann-integrable.*

Furthermore,

(1) The sum and products of Riemann integrable functions are Riemann integrable. Constant functions are Riemann-integrable. So the set of Riemann integrable functions is a real vector space.

(2) The map $f \mapsto \int_a^b f(x)dx$ is a linear map from the space of Riemann integrable functions to $\mathbb{R}$. In particular, if f, g are Riemann-integrable, and α, β real numbers, then

$$\int_a^b (\alpha f(x) + \beta g(x))dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx \ .$$

(3) If $g : \mathbb{R} \to \mathbb{R}$ is a continuous function and $f : I \to \mathbb{R}$ is Riemann integrable then $g \circ f$ is Riemann integrable.

(4) If f, g are integrable and $f(x) \leq g(x)\ \forall\ x$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.

(5) If f is integrable, so is $|f|$, and $|\int_a^b f(x)dx| \leq \int_a^b |f(x)|dx$.

(6) Given a real-valued function f, define the functions[4] $f_\pm$, both with values in $[0, \infty)$, by:

$$f_+(x) = max(f(x), 0) \qquad and \qquad f_-(x) = -min(f(x), 0)$$

so that $f = f_+ - f_-$. Then if f is integrable, so are $f_\pm$, and

$$\int_a^b f(x)dx = \int_a^b f_+(x)dx - \int_a^b f_-(x)dx$$

and

$$\int_a^b |f(x)|dx = \int_a^b f_+(x)dx + \int_a^b f_-(x)dx$$

Finally,

Theorem 1.9. (The indefinite integral is continuous.) *Let f be Riemann-integrable on $I = [a, b]$. Define the function F on I by*

$$F(y) = \int_a^y f(x)dx \ .$$

Then F is continuous. If f is continuous at $x_0 \in I$, then F is differentiable[5] *at x_0 and $F'(x_0) = f(x_0)$.*

and

Theorem 1.10. (The fundamental theorem of calculus.) *Let F is a differentiable function on $I = [a, b]$ such that $f \equiv F'$ is Riemann-integrable on I. Then*

$$\int_a^b f(x)dx = F(b) - F(a)$$

[4] respectively the *positive part* and *negative part* of f.

[5] What does this mean if x_0 happens to be one of the end-points? If, say $x_0 = b$, the meaning is this: F extends to a function $\tilde{F}$ on $[a, b')$ for some $b' > b$, with $\tilde{F}$ differentiable at b, and $\tilde{F}'(b) = f(b)$.

We close with some worked-out examples where we compute Riemann integrals from first principles.. Let $I = [a, b]$, with $a < b$.

(1) Let $f(x) = 1$ Given a partition P given by a sequence $a = t_0 < t_1 < \cdots < t_k = b$, we have

$$U(P, f) = \sum_{i=1}^{k} (t_i - t_{i-1}) = b - a$$

$$L(P, f) = \sum_{i=1}^{k} (t_i - t_{i-1}) = b - a$$

So that $\sup_P L(P, f) = \inf_P U(P, f) = b - a = \int_a^b 1 \, dx$.

(2) Let $f(x) = x$. Given a partition P as above, we have

$$U(P, f) = \sum_{i=1}^{k} t_i (t_i - t_{i-1})$$

$$L(P, f) = \sum_{i=1}^{k} t_{i-1} (t_i - t_{i-1})$$

so that

$$U(P, f) + L(P, f) = \sum_{i=1}^{k} (t_i + t_{i-1})(t_i - t_{i-1}) = \sum_{i=1}^{k} t_i^2 - t_{i-1}^2 = b^2 - a^2$$

$$U(P, f) - L(P, f) = \sum_{i=1}^{k} (t_i - t_{i-1})(t_i - t_{i-1}) = \sum_{i=1}^{k} (t_i - t_{i-1})^2$$

Adding the two equations we get

$$2U(P, f) = b^2 - a^2 + \sum_{i=1}^{k} (t_i - t_{i-1})^2$$

which gives

$$2\overline{\int_a^b} f(x)dx = b^2 - a^2 + \inf_P \{\sum_{i=1}^{k} (t_i - t_{i-1})^2\}$$

We claim that $\inf_P \{\sum_{i=1}^{k} (t_i - t_{i-1})^2\} = 0$. To see this, note that this infimum is non-negative. If the t_i are evenly spaced at intervals of length $(b - a)/k$, then $\sum_{i=1}^{k} (t_i - t_{i-1})^2 = (b - a)^2/k$; now let $k \to \infty$.

Exercise Finish the proof that f is Riemann integrable, so that

$$\int_a^b f(x)dx = \frac{b^2 - a^2}{2}$$

(3) Let c be such that $a < c < b$ and let H_c (H for Heaviside) be defined by:

$$H_c(x) = \begin{cases} 0 & if \quad x \leq c \\ 1 & if \quad x > c \end{cases}$$

Given a partition P given by a sequence $a = t_0 < t_1 < \cdots < t_{i_0-1} \leq c < t_{i_0} \cdots < t_k = b$,

$$U(P, H_c) = \sum_{i=1,\ t_i > c}^{k} (t_i - t_{i-1}) = b - t_{i_0-1}$$

$$L(P, H_c) = \sum_{i=1,\ t_{i-1} > c}^{k} (t_i - t_{i-1}) = b - t_{i_0}$$

so that $\sup_P L(P, H_c) = \inf_P U(P, H_c) = b - c = \int_a^b H_c(x)\ dx$.

(4) The same computation shows that for $x \in I$, the function R_c defined by

$$R_c(y) = \int_a^y H_c(x)dx$$

is the "ramp function"

$$R_c(x) = \begin{cases} 0 & if \quad x \leq c \\ x - c & if \quad x > c \end{cases}$$

Note that R_c is continuous on I and differentiable except at $x = c$, the point of discontinuity of H_c; away from $x = c$ we have $R_c'(x) = H_c(x)$. This illustrates Theorem 1.9.

(5) In contrast, if $f(x) = x$ as in example 2, we have $F'(x) = f(x)$ where

$$F(x) = \frac{x^2}{2} + u$$

for an arbitrary constant u, and

$$\int_a^y f(x)dx = F(y) - F(a)$$

for any $a \leq y \leq b$. This illustrates the fundamental theorem of calculus.

Exercises on the (one-dimensional) Riemann Integral

We will let $I = [a, b]$, with $a < b$. By a function we will mean real-valued function defined on I.

(1) Let $x_0 \in I$, and let f_{x_0} be the function defined by:

$$f_{x_0}(x) = \begin{cases} 0 & if \quad x \neq x_0 \\ 1 & if \quad x = x_0 \end{cases}$$

Prove that f_{x_0} is Riemann-integrable and $\int_a^b f_{x_0}(x)dx = 0$.

(2) Let f be a continuous function such that $f \geq 0$, and suppose $\int_a^b f(x)dx = 0$. Prove that $f \equiv 0$. (Just to clarify: $f \geq 0$ means $f(x) \geq 0$ for all x and $f \equiv 0$ means "f is identicaly zero", that is, $f(x) = 0$ for all x.)

(3) Let $\mathcal{I}$ be the function

$$\mathcal{I}(x) = \begin{cases} 1 & if \quad x \text{ is rational} \\ -1 & if \quad x \text{ is irrational} \end{cases}$$

Is $\mathcal{I}$ Riemann integrable?

(4) Prove or disprove: if f^2 is Riemann integrable, so is f.

(5) Give an example of Riemann integrable f such that F, defined as in the statement of Theorem 1.9, is differentiable but $F'(x_0) \neq f(x_0)$ for some x_0.

(6) If P is the function defined by

$$P(x) = a_1 + 2a_2x \cdots + (n+1)a_{n+1}x^n$$

write down an expression for

$$\int_a^b P(x)dx$$

Justify your answer.

(7) The techniques of high-school integration work with Riemann integration. For example, if f, g are continuously differentiable on I, then

$$\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)dx$$

(A function f is *continuously differentiable* on I if it extends to an open interval $(a_2, b_2) \supset [a, b]$ as a differentiable function with continuous derivative.)

(8) Let $a \leq a_1 < b_1 \leq b$, and set $I_1 = [a_1, b_1]$. If f is a Riemann integrable function on I, prove that its restriction f_1 to I_1 is Riemann integrable and

$$\int_{a_1}^{b_1} f_1(x)dx = \int_a^b f(x)\chi_{I_1}(x)dx$$

Explanation of notation: Given a set X and a subset E, the *indicator function* of E (also called its *characteristic function*) is the real-valued function χ_E on X defined by

$$\chi_E(x) = \begin{cases} 1 & if \quad x \in E \\ 0 & if \quad x \notin E \end{cases}$$

Darboux/Riemann integrability The definition we have given of the Riemann integral is due to Darboux. (We follow Rudin, who does not even mention Darboux.) You might have encountered a definition of the Riemann integral in terms of *tagged partitions*. See

`http://www.math.toronto.edu/karshon/courses/mat157/notes.1901-Riemann_integral.pdf`

for a brief explanation of the relationship between the two definitions. Note that there our construction is called the Darboux integral, and the term Riemann integral is applied to the definition in terms of tagged partitions.

1.1. **Integration on open sets, all of $\mathbb{R}$.** We now need to ask: what about integrals over open or unbounded intervals, of the form (a, b), (a, ∞), $(-\infty, \infty)$? For example, we have the beautiful formula:

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

How do we give a rigorous meaning to this, and then actually prove it?

We postpone discussion of this question.

2. Multidimensional integrals: definition, Jordan-measurable sets

2.1. **Integration over closed rectangles.** To begin with, we will define upper and lower Riemann integrals of bounded real-valued functions over closed rectangles, which will be the analogues of closed bounded intervals in the one-dimensional theory. By a closed rectangle R in $\mathbb{R}^n$ we will mean a product

$$R = I_1 \times I_2 \times \cdots \times I_n$$

where each I_i is a bounded closed interval: $I_i = [a_i, b_i]$, where a_i, b_i are real numbers such that $a_i \leq b_i$. The *volume* of R is defined to be

$$vol(R) = \prod_{i=1}^{n} (b_i - a_i)$$

Usually we assume that R has nonzero volume, in other words, that $a_i < b_i \ \forall \ i$. Of course, if $n = 1$ a more appropriate term for volume is "length" and if $n = 2$, "area".

For simplicity let us consider integration in $\mathbb{R}^2$. It will be clear how to extend the theory to higher dimensions. Consider a rectangle R is of the form $I_1 \times I_2$ where $a_l < b_l, \ l = 1, 2$ and $I_l = [a_l, b_l] \subset \mathbb{R}$. The *area* of the rectangle R is

$$area(R) = (b_2 - a_2)(b_1 - a_1)$$

By a partition P of R we shall mean partitions of each of the "factor intervals" I_l

$$P_1 : a_1 = t_{1,0} < t_{1,1} < \ldots t_{1,k_1 - 1} < t_{1,k_1} = b_1$$
$$P_2 : a_2 = t_{2,0} < t_{2,1} < \ldots t_{2,k_2 - 1} < t_{2,k_2} = b_2$$

(We may use the notation $P = P_1 \times P_2$.) Given such a partition, we can write R as a union of rectangles $R_{i_1,i_2} = [t_{1,i_1-1}, t_{1,i_1}] \times [t_{2,i_2-1}, t_{2,i_2}]$ which are disjoint except possibly for some common horizontal/vertical sides and/or corners.

Given a bounded real-valued function f defined on R, we define its upper and lower Riemann sums w.r.to P by

$$U(P, f) \equiv \sum_{i_1, i_2} \Big(\sup_{x \in R_{i_1,i_2}} f(x) \Big) \times area(R_{i_1,i_2})$$
$$L(P, f) \equiv \sum_{i_1, i_2} \Big(\inf_{x \in R_{i_1,i_2}} f(x) \Big) \times area(R_{i_1,i_2})$$

A refinement $P' = P_1' \times P_2'$ consists of refinements $P_1', \ P_2'$ of $P_1, \ P_2$ respectively. Clearly two partitions of R have a common refinement. (Easy proof.) As a consequence:

$$L(P, f) \leq U(Q, f)$$

for *any two* partitions P, Q. We define upper and lower Riemann integrals as before:

$$\overline{\int_R} f(x)dx = \inf_P U(P,f)$$
$$\underline{\int_R} f(x)dx = \sup_P L(P,f)$$

We say that f is *Riemann integrable* if $\overline{\int_R} f(x)dx = \underline{\int_R} f(x)dx$, in which we case we use the notation

$$\int_R f(x)dx = \overline{\int_R} f(x)dx = \underline{\int_R} f(x)dx$$

Remark 2.1. There is not much standardisation in the notation for multi-dimensional integrals. Note that above x is a "dummy variable" that stands for a "dummy point" in $\mathbb{R}^2$.

(1) If the coordinates are denoted (x_1, x_2), we use the notation

$$\int_R f(x)dx = \int_R f(x_1, x_2)dx_1 dx_2$$

(2) If the coordinates are denoted (x^1, x^2), we use the notation

$$\int_R f(x)dx = \int_R f(x^1, x^2)dx^1 dx^2$$

(This notation is sometimes favoured by differential geometers, and is used by Spivak.)

(3) Rudin uses the notation

$$\int_R f(x)dx = \int_R f(\vec{x})d\vec{x}$$

where the point in $\mathbb{R}^2$ is denoted as a vector $\vec{x}$.

(4) We will see that under suitable conditions (for example, if f is continuous on R), the integral over R can be written as an "iterated integral" $\int_{a_2}^{b_2}(\int_{a_1}^{b_1} f(x_1, x_2)dx_1)dx_2$, and (implicitly assuming that this is the case), one writes

$$\int_R f(x)dx = \int_{a_2}^{b_2}\int_{a_1}^{b_1} f(x_1, x_2)dx_1 dx_2$$

This notation conflates a definition and theorem; I would avoid it.

(5) Sometimes we omit the dummy variable altogether, and write simply

$$\int_R f(x)dx = \int_R f$$

As in one dimension,

(1) Any two partitions of a closed rectangle have a common refinement.

(2) ("Cauchy criterion") A bounded f on a closed rectangle is integrable iff for any $\epsilon > 0$ there exists a partition P such that $U(P, f) - L(P, f) < \epsilon$.

(3) A continuous function (on a closed rectangle) is Riemann integrable.

2.2. **Integration on (all of) $\mathbb{R}^n$ of functions vanishing outside a compact set.** We need to extend the definition of Riemann integrals to suitable classes of (real-valued) functions on $\mathbb{R}^n$. Unlike in the case of Lebesgue integral, different authors take slightly different (but equivalent) approaches to this. We will choose one particular route.

Given a real-valued function f on $\mathbb{R}^n$ and $A \subset \mathbb{R}^n$, let us agree that "f vanishes outside A" means $f(x) = 0$ if $x \notin A$.

Lemma 2.2. *Let f be a bounded real-valued function on $\mathbb{R}^n$ which vanishes outside a closed rectangle R and is integrable on R. Then*

(a) given any other closed rectangle $R' \supset R$ then f is integrable on R' and

$$\int_R f|_R = \int_{R'} f|_{R'}$$

(b) given any other closed rectangle R'' such that f vanishes outside R'', f is integrable on R'' and

$$\int_R f|_R = \int_{R''} f|_{R''}$$

Proof. Given a partition P of R, we can extend it to a partition P' of R'. We have then

$$L(P, f|_R) = L(P', f|'_R), \text{ and}$$
$$U(P, f|_R) = U(P', f|_{R'})$$

Now use the Cauchy criterion to conclude (a). The proof of (b) is left as an exercise. □

We can now extend the definition of Riemann integrability as follows:

Definition 2.3. If f be a bounded real-valued function on $\mathbb{R}^n$ which vanishes outside some closed rectangle R, we say that f *is integrable on* $\mathbb{R}^n$ if f is integrable on R. We set

$$\int_{\mathbb{R}^n} f \underset{def}{=} \int_R f|_R$$

Definition 2.4. Given a real-valued continuous function f (on $\mathbb{R}^n$, a rectangle, and more generally whenever the definition makes sense) define its

support to be the closure of the subset of points where f is non-vanishing. In other words,

$$\text{support of } f \underset{Notation}{=} supp(f) \underset{Definition}{=} \overline{\{x | f(x) \neq 0\}} \ .$$

We say that f *has compact support* if $supp(f)$ is compact.

Note that a continuous function with compact support is bounded.

Exercise In the following f is a continuous real-valued function.

(1) $f|_{supp(f)^c} = 0$.

(2) If U is open and $f|_U = 0$, then $U \subset supp(f)^c$. So $supp(f)^c$, the complement of $supp(f)$, is the largest open set on which f restricts to zero.

(3) If f is a continuous function defined on an open subset U of $\mathbb{R}^n$, and f has compact support, the extension $\tilde{f}$ of f to $\mathbb{R}^n$ defined by

$$\tilde{f}(x) = \begin{cases} f(x) & if\ x \in U,\ and \\ 0 & if\ x \notin U, \end{cases}$$

is continuous, and $supp(\tilde{f}) = supp(f)$. Note that the support of f is defined w.r.to U (closure taken in U) and the support of $\tilde{f}$ w.r. to $\mathbb{R}^n$.

If f is a continuous function on $\mathbb{R}^n$ with compact support, we set

$$\int_{\mathbb{R}^n} f = \int_R f|_R$$

where R is any rectangle that contains $supp(f)$.

Remark 2.5. You might wonder why the notion of support was defined only for continuous functions, whereas the definition would make sense without the assumption of continuity. The reason is that in the context of the Lebesgue integral there exists a more appropriate notion of *essential support* of a Lebesgue-measurable function.

As in one dimension,

(1) The sum, products, and scalar multiples of Riemann integrable functions are Riemann integrable. The set of Riemann integrable functions is a real vector space.

(2) The map $f \mapsto \int_{\mathbb{R}^n} f$ is a linear map from the space of Riemann integrable functions to $\mathbb{R}$. In particular, if f, g are Riemann-integrable and α, β real numbers, then

$$\int_{\mathbb{R}^n} (\alpha f + \beta g) = \alpha \int_{\mathbb{R}^n} f dx + \beta \int_{\mathbb{R}^n} g$$

(3) If $g : \mathbb{R} \to \mathbb{R}$ is a continuous function and f is Riemann integrable then $g \circ f$ is Riemann integrable.

(4) If f, g are integrable and $f(x) \leq g(x)\ \forall\ x$, then $\int_{\mathbb{R}^n} f \leq \int_{\mathbb{R}^n} g$.

(5) If f is integrable, so is $|f|$, and $|\int_{\mathbb{R}^n} f| \leq \int_{\mathbb{R}^n} |f|$.

(6) If f is integrable, so are $f_\pm$ (defined earlier in the one-dimensional context), and

$$\int_{\mathbb{R}^n} f(x)dx = \int_{\mathbb{R}^n} f_+(x)dx - \int_{\mathbb{R}^n} f_-(x)dx$$

and

$$\int_{\mathbb{R}^n} |f(x)|dx = \int_{\mathbb{R}^n} f_+(x)dx + \int_{\mathbb{R}^n} f_-(x)dx$$

3. Integration over more general bounded sets; Jordan-measurable sets

.

Recall first that given $S \subset \mathbb{R}^n$, the indicator function $\chi_S : \mathbb{R}^n \to \{0, 1\}$ is defined by:

$$\chi_S(x) = \begin{cases} 1 & if \quad x \in S \\ 0 & if \quad x \notin S \end{cases}$$

The "correct" condition to impose on S so that one has a good theory of integration over S, and in particular so that one can define a sensible notion of its area (in arbitrary dimension, we would call this volume) is the following:

 # S is bounded and χ_S is Riemann-integrable on $\mathbb{R}^n$.

Note that this means that S is contained in some closed rectangle R and χ_S is integrable on R, and that this implies that χ_S is integrable on any other closed rectangle R' containing S.

Given such an S, we define its *volume* to be

$$vol(S) \underset{def}{=} \int_{\mathbb{R}^n} \chi_S$$

One can check that if $R = I_1 \times I_2 \times \cdots \times I_n$, then

$$vol(R) = \prod_{i=1}^{n} |I_i|$$

where $|I_i|$ is the length of the interval I_i.

Given an integrable function f on $\mathbb{R}^n$, we define its integral over S to be

$$\int_S f \underset{def}{=} \int_{\mathbb{R}^n} f\chi_S$$

This makes sense since the function $f\chi_S$ is a product of integrable functions and hence integrable. Note that the integral over S only depends on the values of f at points of S in the sense that if $f|_S$ is the zero function $f\chi_S$ vanishes identically on $\mathbb{R}^n$ and hence has zero integral.

If S is as above and g is a bounded real-valued function on S, let $\tilde{g}$ denote the function defined on $\mathbb{R}^n$ by

$$\tilde{g}(x) = \begin{cases} g(x), & \text{if } x \in S, \text{ and} \\ 0 & \text{if } x \notin S \ . \end{cases}$$

Sometimes $\tilde{g}$ is called the *extension by zero* of g. Note that even if g is continuous on S, $\tilde{g}$ need not be continuous on $\mathbb{R}^n$. We say that g *is integrable on* S if $\tilde{g}$ is integrable on $\mathbb{R}^n$, and set

$$\int_S g = \int_{\mathbb{R}^n} \tilde{g} = (\text{according to our earlier definition}) \int_R \tilde{g}|_R$$

where R is any closed rectangle containing S. Note that

(1) if $g = f|_S$, with f integrable on $\mathbb{R}^n$, then $\tilde{g} = f\chi_S$, so integrable functions on S are precisely restrictions of integrable functions on $\mathbb{R}^n$, and

(2) if $S = R$, a closed rectangle, this is consistent with our initial definitions.

Exercise: If R is a closed rectangle, show that its interior R^0 satisfies # by checking that χ_{R^0} is Riemann-integrable on R. Show that

$$\int_{\mathbb{R}^n} \chi_{R^0} = vol(R)$$

We make a number of important remarks, which you are asked to verify.

(1) Note that all the above definitions make sense in arbitrary dimensions, including dimension 1.

(2) It could happen that S has zero volume and all integrals over S vanish. Intuitively, this happens if S is a "lower-dimensional" set. Examples are: a finite set of points (in $\mathbb{R}^n$ for any $n \geq 1$), a line segment in $\mathbb{R}^2$ (say, $\{(x_1, 0) | 0 \leq x_1 \leq 1\}$ in $\mathbb{R}^2$).

(3) If S, S' satisfy # so do $S \cup S'$, $S \cap S'$, and $S \setminus S'$. (Use the identities: $\chi_{S \cup S'} = \chi_S + \chi_{S'} - \chi_S \chi_{S'}$, and $\chi_{S \setminus S'} = \chi_S - \chi_S \chi_{S'}$.)

(4) Integrals are additive under (finite) unions of sets. In particular volumes are additive.

3.1. **Characterisation of sets satisfying #.** Given a subset T of $\mathbb{R}^n$ we say that T has *measure zero* if for every $\epsilon > 0$, there are *countably many* rectangles $R_1, R_2, \ldots$, such that $T \subset \cup_i R_i$ and $\sum_i volume(R_i) < \epsilon$. The set T has *content zero* if there are *finitely many* rectangles $R_1, R_2, \ldots, R_N$ such that $T \subset \cup_i R_i$ and $\sum_i volume(R_i) < \epsilon$.

Note that we have **not** defined the notion of the measure or content of a set, only the notion of a set of measure or content zero.

Proposition 3.1. *If T is compact and has measure zero, it has content zero.*

Proof. Consider the case $n = 2$. Suppose T is compact and of measure zero. Let $\epsilon > 0$ be given and let $R'_1, R'_2, \ldots$ be a countable family of rectangles such that $T \subset \cup_i R'_i$ and $\sum_i area(R'_i) < \epsilon/4$. Let R_i be a rectangle of twice the size (all sides are doubled) such that $R'_i \subset R^0_i$, where R^0_i is the interior of R_i (see below for the definition). Then $T \subset \cup_i R^0_i$, and since the R^0_i are open and T is compact, there exist finitely many (which by relabelling, we can take to be $R^0_1, \ldots, R^0_n$ such that $T \subset R^0_1 \cup \cdots \cup R^0_n \subset R_1 \cup \cdots \cup R_n$. On the other hand

$$area(R_1) + \cdots + area(R_n) = 4(area(R'_1) + \cdots + area(R'_n)) \leq 4 \sum_i area(R'_i) < \epsilon$$

In other dimensions $\neq 2$ the factor 4 has to be replaced. □

Here is a characterisation of integrable functions on a rectangle, due to Lebesgue. This is Theorem 3-8 of Spivak, and you can read the proof there. (*You will not be expected to recall the proof.*)

Theorem 3.2. *Let R be a rectangle, and f a bounded real-valued function defined on R. Then f is Riemann integrable iff "the set of discontinuities of f", i.e., $\{x \in R | f$ is not continuous at $x\}$ has measure zero.*

With this theorem in hand we can characterise subsets satisfying #.

We will need to recall some preliminary definitions[6]. Given a metric space X and a subset E, the *interior* E^0 consists of interior points of E, where in turn, a point $x \in E$ is an *interior point* if some open ball centred on x is entirely contained in E. The closure $\bar{E}$ consists of limit points of E, where in turn, a point $x \in X$ is an *limit point of* E if some every open ball centred on x intersects E. The interior E^0 is the union of all open sets contained in E ("the biggest open set contained in E"). The closure $\bar{E}$ is closed (this needs proof) and is the intersection of all closed sets containing E ("the smallest closed set containing E"). The boundary of E, denoted ∂E, is defined by

$$\partial E \equiv \bar{E} \setminus E^0$$

Note that ∂E is closed. If E is closed, so that $\bar{E} = E$, we have $\partial E \equiv E \setminus E^0$. We will use the following fact, which we leave as an exercise. If E^c denotes the complement (in X) of a subset E, we have

$$\partial E = \partial(E^c)$$

Note that the definitions depend on E as well as the "ambient" set X.

[6]For these matters, see `http://math.stanford.edu/~conrad/diffgeomPage/handouts/closure.pdf`. Brian Conrad has excellent and careful notes on Mathematics at all levels; a wonderful resource!

Remark 3.3. This is regarding the definition of limit points of E using sequences. There seems to be some variation in terminology, but you can check that with with our definition, $x \in X$ is a limit point of E iff there exists a sequence $x_n \in E$ such that $x_n \to x$. Note that there is no requirement that the elements of the sequence are distinct. For example, if $x \in E$, the sequence $x_n = x$ tends to x so every point of E is a limit point.

Theorem 3.4. *Let $S \subset \mathbb{R}^n$ be a bounded set. Then S satisfies # iff it is Jordan-measurable, i.e., the boundary ∂S has content zero.*

Proof. Let $R \supset S$ be a rectangle such that $\overline{S} \subset R^0$. (That is the closure of S is contained in the interior of R.) We have the following chain of two-way implications.

(1) By definition, S satisfies # iff χ_S is Riemann-integrable, where χ_S is the characteristic function of S relative to R.

(2) χ_S is Riemann-integrable iff its set of discontinuity has measure zero. (This is by the previous Theorem.)

(3) The set of discontinuities of χ_S is the boundary of S relative to R. On the other hand, since $\overline{S} \subset R^0$, ∂S is also the boundary ∂S of S relative to $\mathbb{R}^n$. Since ∂S is closed and bounded, it is compact so it has measure zero iff it has content zero.

□

Remark 3.5. Given an arbitrary function f the set of discontinuities need not be closed (equivalently, the set of points where f is continuous need not be open.). On the other hand, given a set S, the set of discontinuities of χ_S is closed since it is the boundary of S.

We can use Lebesgue's theorem to prove:

Proposition 3.6. *If $S \subset \mathbb{R}^n$ satisfies # and f a bounded continuous real-valued function on S, then f is Riemann-integrable on S.*

Proof. Let $\tilde{f}$ be the extension by zero to $\mathbb{R}^n$. Then $\tilde{f}$ is identically zero on the open set $\overline{S}^c$, and continuous on the interior S^0. Therefore the set of discontinuities of $\tilde{f}$ is contained in ∂S, which has content zero. Thus $\tilde{f}$ is integrable on $\mathbb{R}^n$, which is the definition of integrability of f on S. □

The proof of the next Proposition is left as an exercise:

Proposition 3.7. *If $S \subset \mathbb{R}^n$ satisfies #, so do its interior S^0, and its closure $\overline{S}$. If f is integrable on $\mathbb{R}^n$, we have*

$$\int_S f(x)dx = \int_{\overline{S}} f(x)dx = \int_{S^0} f(x)dx$$

3.2. **Examples.** Let us work out some examples.We will mostly consider (compact) sets S which are connected or a finite disjoint union of such sets. (In one dimension, such an S is a disjoint union of finitely many finite closed intervals, possibly singletons.) The sets of interest will have relatively "nice" boundaries, which will be easy to determine, and we will be able to check that these boundaries have zero content, thus ensuring that S is Jordan-measurable.

(1) Let s_0 be a real number and consider the singleton set $S \equiv \{s_0\} \subset \mathbb{R}$. The set S is compact. It has empty interior, so $S = \partial S$. Given any $\epsilon > 0$, the closed interval $[s_0 - \epsilon/3, s_0 + \epsilon/3]$ (a one-dimensional analogue of "rectangle") contains S and has $length = "volume" = \frac{2}{3}\epsilon < \epsilon$. So ∂S has content zero. So given any closed interval I containing s_0, the indicator function of S relative to I is integrable. This can, of course, be checked directly from the definition. We have

$$length(S) = \int_I \chi_S(x)dx = 0$$

which is as it should be.

(2) Let us check that with our definitions area is additive under finite disjoint unions. Let $S_1, S_2, \ldots S_n$ be disjoint sets satisfying #, and let $S = \cup_i S_i$. Then $\chi_S = \sum_i \chi_{S_i}$ (because the sets are disjoint) so χ_S is Riemann-integrable (being a finite sum of Riemann integrable functions. This shows that S satisfies #. Finally,

$$area(S) \underset{def}{=} \int_{\mathbb{R}^n} \chi_S(x)dx = \int_{\mathbb{R}^n} \sum_i \chi_{S_i}(x)dx = \sum_i \int_{\mathbb{R}^n} \chi_{S_i}(x) = \sum_i area(S_i)$$

(3) It will be more useful to split up integrals over S into integrals over subsets which are "almost disjoint". We will not formulate a general theorem, since already the load of abstractions is threatening to sink our ship. Here is an illustrative case. You will have to repeat the arguments quickly in problems when necessary.

Let S be the "football pitch": the union of three compact subsets: a rectangle R_1, and two closed half-discs D_1, D_2 at opposite ends so that diameter segments d_1, d_2 are opposite sides of R_1. One can check that the discs satisfy # (for example, you can think of them as rotated versions of areas under graphs, see the next subsection), as also the diameters d_1, d_2. Now S is no longer a disjoint union, but it is still true that it satisfies #, and in fact

$$area(S) = area(D_1) + area(R_1) + area(D_2)$$

Here is the argument. Let R be any rectangle containing S. Then

$$\chi_S = \chi_{R_1} + \chi_{D_1} + \chi_{D_2} - \chi_{d_1} - \chi_{d_2}$$

from which we conclude that S satisfies #. Further,

$$area(S) \underset{def}{=} \int_{\mathbb{R}^n} \chi_S(x)dx = area(R_1) + area(D_1) + area(D_1)$$

since

$$\int_{\mathbb{R}^n} \chi_{d_i}(x)dx = 0,\ i = 1, 2\ .$$

3.3. **Exercise: Integral as area under the graph.** Let $f : I \equiv [a, b] \to \mathbb{R}$ be a continuous function, and assume that $f(x) \geq 0, x \in I$. Let $S_f \subset \mathbb{R}^2$ be the "area under the graph":

$$S_f = \{(x, y) \in \mathbb{R}^2 | x \in I,\ 0 \leq y \leq f(x)\}$$

Determine the boundary ∂S_f, and show that its content is 0, thus verifying that S_f satisfies #. Compute the area of S_f from first principles and show that

$$area(S_f) = \int_a^b f(x)dx$$

One will later reprove this using iterated integrals.

3.4. **Jordan-measurable sets, locally integrable functions.** A bounded $S \subset \mathbb{R}^n$ whose boundary ∂S has measure zero is called (in Spivak, and many other references) *Jordan-measurable*. We have seen that these are precisely the subsets satisfying the condition #.

Let us agree that a real-valued function f on $\mathbb{R}^n$ is *locally integrable* if its restriction to any rectangle $R \subset \mathbb{R}^n$ is (bounded and) integrable on R. [7]

Prove the following:

Theorem 3.8. *f is locally integrable iff its set of discontinuity has Jordan-measure zero. Given a (not necessarily bounded) subset S, χ_S is locally integrable iff ∂S has measure zero.*

In earlier versions of the notes and possibly in class I talked of sets of "Jordan-measure zero." Mr. Jordan's name now only occurs in the phrase "Jordan-measurabe".

4. Iterated integrals; "Fubini" Theorem

We will first state and prove Fubini almost exactly as done in Spivak. (For a change, we will work in dimension $\mathbb{R}^m \times \mathbb{R}^n$.) This is in more generality than needed in "practical applications". But we do this because the generality actually helps with understanding the hypothesis and proof of the "practical" version of the theorem that follows.

Note that Fubini proved his theorem in the context of Lebesgue integrals. Hence the quotation marks.

[7] Equivalently (why?), for any $x_0 \in \mathbb{R}^n$, there is a rectangle R_{x_0} containing x_0 in its interior such that f is (bounded and) integrable on R_{x_0}.

Theorem 4.1. *Let $A \subset \mathbb{R}^m$ and $B \subset \mathbb{R}^n$ be rectangles, and let R be the product rectangle $A \times B$. Let f be a (real-valued, bounded) Riemann integrable function on R. For $x \in A$, define*[8] *$fr_x : B \to \mathbb{R}$ by*

$$fr_x(y) = f(x, y)$$

Let $\mathcal{L}$ and $\mathcal{U}$ be the functions defined on the rectangle A by

$$\mathcal{L}(x) \equiv \underline{\int_B} fr_x(y)dy$$

$$\mathcal{U}(x) \equiv \overline{\int_B} fr_x(y)dy$$

Then $\mathcal{L}$ and $\mathcal{U}$ are Riemann integrable on A and

$$\int_R f(u)du = \int_A \mathcal{L}(x)dx = \int_A \mathcal{U}(x)dx$$

Remark on Notation: The "dummy" variable u denotes a point of $\mathbb{R}^m \times \mathbb{R}^n$. It is more suggestive to substitute u with the dummy variables $x \in A$ and $y \in B$ and write

$$\int_{A\times B} f(x, y)dxdy = \int_A \mathcal{L}(x)dx = \int_A \mathcal{U}(x)dx$$

You should not let all this talk of dummy variables confuse you. Understand clearly the definitions of the functions $\mathcal{L}$,$\mathcal{U}$ (defined on the rectangle A) and the equation (with all dummy variables dropped):

$$\int_R f = \int_A \mathcal{L} = \int_A \mathcal{U}$$

where each of the integrals is defined as a Riemann integral over the appropriate rectangle.

Proof. Start with partitions P_A, P_B of the rectangles A, B respectively. Let $P = P_A \times P_B$ be the corresponding partition of R. (Note that according to our definitions, this is how a general partition of R looks.) A general sub-rectangle of R is of the form $S_A \times S_B$ where S_A belongs to the partition P_A and S_B belongs to the partition P_B. Then

$$L(P, f) = \sum_S m_S(f)vol(S) = \sum_{S_A} vol(S_A) \sum_{S_B} m_{S_A\times S_B}(f)vol(S_B) \ .$$

On the other hand, if $x \in S_A$, we have

$$m_{S_B}(fr_x) = \inf_{y\in S_B} fr_x(y) = \inf_{y\in S_B} f(x, y) \geq \inf_{x'\in S_A, y\in S_B} f(x', y) = m_{S_A\times S_B}(f)$$

from which we extract

$$m_{S_B}(fr_x) \geq_a m_{S_A\times S_B}(f) \text{ provided } x \in S_A \ .$$

[8]The notation reflects the fact that fr_x is the function on B got by "freezing" the first variable at x. Note that fr_x is certainly bounded (as a function on B), but need not be Riemann integrable. But we can consider the lower and upper Riemann integrals. These are the functions $\mathcal{L}$ and $\mathcal{U}$ of the variable x.

Next,

$$\mathcal{L}(x) \geq_b L(P_B, fr_x) =_c \sum_{S_B} m_{S_B}(fr_x) vol(S_B)$$
$$\geq_d \sum_{S_B} m_{S_A \times S_B}(f) vol(S_B) \quad \textit{provided } x \in S_A \ .$$

Explanation: the inequality (b) holds because $\mathcal{L}(x)$ (the lower Riemann integral of fr_x) is by definition the supremum of $L(P_B, fr_x)$ over partitions P_B. The equality (c) is the definition of $L(P_B, fr_x)$. The inequality (d) uses (a). From the last chain of inequalities we extract:

$$\mathcal{L}(x) \geq \sum_{S_B} m_{S_A \times S_B}(f) vol(S_B) \textit{ provided } x \in S_A \ .$$

It follows that

$$m_{S_A}(\mathcal{L}) \geq \sum_{S_B} m_{S_A \times S_B}(f) vol(S_B)$$

Gathering inequalities, we get

$$\sum_{S_A} m_{S_A}(\mathcal{L}) vol(S_A) = L(P_A, \mathcal{L}) \geq_A L(P, f)$$

Similarly,

$$U(P_A, \mathcal{U}) \leq_B U(P, f)$$

Consider first the inequality (A): taking the supremum over all P gives

$$\underline{\int_A} \mathcal{L} \geq \int_R f$$

Similarly, taking infimum over all P yields

$$\overline{\int_A} \mathcal{U} \leq \int_R f$$

Now, since $\mathcal{L} \leq \mathcal{U}$, we have $\underline{\int_A} \mathcal{L} \leq \underline{\int_A} \mathcal{U}$. On the other hand, we have $\underline{\int_A} \mathcal{U} \leq \overline{\int_A} \mathcal{U}$. Putting it all together, we get

$$\int_R f \leq \underline{\int_A} \mathcal{L} \leq \underline{\int_A} \mathcal{U} \leq \overline{\int_A} \mathcal{U} \leq \int_R f$$

From this we can conclude that $\mathcal{U}$ is integrable and $\int_A \mathcal{U} = \int_R f$. Similarly we conclude that $\mathcal{L}$ is integrable and $\int_A \mathcal{L} = \int_R f$. □

Corollary 4.2. *Under the conditions of the Theorem (in particular f is integrable on $A \times B$), assume further that for $x \in A$, the functions fr_x are Riemann integrable on B. Then*

$$\int_A \{ \int_B f(x,y) dy \} dx = \int_{A \times B} f(x,y) dx dy$$

That is, the Riemann Integral over the product rectangle is the "iterated integral" on the left got by integrating the functions fr_x over B and then integrating the resulting function over A.

Corollary 4.3. *Under the conditions of the Theorem (in particular f is integrable on $A \times B$), assume further that for $x \in A$, the functions fr_x are Riemann integrable on B and that the (similarly defined, for $y \in B$) functions fr^y are Riemann integrable over A. Then*

$$\int_A \{\int_B f(x,y)dy\}dx = \int_{A\times B} f(x,y)dxdy = \int_B \{\int_A f(x,y)dx\}dy$$

In particular the two iterated integrals are equal:

$$\int_A \{\int_B f(x,y)dy\}dx = \int_B \{\int_A f(x,y)dx\}dy$$

In particular, the corollary applies if f is continuous on R.

As an application let us return to the exercise §3.3: Let $f : I \equiv [a,b] \to \mathbb{R}$ be a Riemann integrable[9] function, and assume that $f(x) \geq 0, x \in I$. Let $S_f \subset \mathbb{R}^2$ be the "area under the graph":

$$S_f = \{(x,y) \in \mathbb{R}^2 | x \in I,\ 0 \leq y \leq f(x)\}$$

Show that the content of ∂S_f is 0, thus verifying that S_f satisfies #. Compute the area of S from first principles and show that

$$area(S_f) = \int_a^b f(x)dx$$

Solution (f continuous): Let us consider the case when f is continuous. Fix a positive real number M such that $\sup_{x\in I} f(x) < M$. Let R be the rectangle $I \times [0, M]$.

Claim, with proof left as exercise: S_f is closed and the interior of S_f is $\{\{(x,y) \in \mathbb{R}^2 | x \in (a,b),\ 0 < y < f(x)\}$, so that ∂S_f is the (not disjoint!) union of $T = \{(x, f(x)) \in \mathbb{R}^2 | x \in I\}$ and three (closed) straight line segments.

It suffices to show that T has content zero to conclude that the same holds for ∂S_f.

Let $\epsilon > 0$ be given. By uniform continuity of f (on the bounded closed interval I), there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon/2$ provided $|x - y| < \delta/2$. Divide I into N equal intervals I_j such that $(b-a) < N\delta$. For each of these sub-intervals I_j, let x_j be the midpoint, and R_j be the rectangle centred on $(x_j, f(x_j))$:

$$R_j = \{(x,y) | x \in I_j, |y - f(x_j)| \leq \epsilon/2\}$$

By construction $T \subset \cup_j R_j$ and

$$\sum_j area(R_j) = \epsilon \sum_j \ell(I_j) = (b-a)\epsilon$$

Since ϵ is arbitrarily small, it follows that T has content zero.

[9]Earlier we took f to be continuous; with f assumed merely intergrable this is Problem 3-6 of Spivak.

We now apply Fubini to the Riemann integrable function χ_{S_f} defined on R. For $x \in I$, the function $y \mapsto \chi_{S_f}(x,y)$ is the characteristic function of the set $[0, f(x)]$ relative to the interval $[0, M]$. This is an integrable function of y (with integral equal to $\ell([0, f(x)])$) and we can conclude

$$Area(S_f) = \int_R \chi_{S_f} = \int_I \{ \int_{[0,M]} \chi_{S_f}(x,y) dy \} dx$$
$$= \int_a^b \ell([0, f(x)]) dx = \int_a^b f(x) dx$$

Exercise: Prove this assuming f is bounded and continuous except at one point. More precisely, suppose that $a < c < b$, and f_1 is continuous on $[a, c]$, f_2 is continuous on $[c, b]$, with $f_1(c) \neq f_2(c)$, and f is defined by

$$f(x) = \begin{cases} f_1(x) & if\ x \in [a, c] \\ f_2(x) & if\ x \in (c, b] \end{cases}$$

so that "f is continuous from the left at c." Then try the general case of measurable f.

5. Revision of multivariable differential calculus

We begin considering *one* real-valued function of $n \geq 1$ variables.

For simplicity of notation, we often treat the case $n = 2$. Let $U \subset \mathbb{R}^2$ be an open set. Differential calculus is "local" so we can take U to be connected, and even a ball (=disc when $n = 2$); for the same reason we need not assume that f is bounded on U.

[It is natural to start setting up integration of bounded functions on *bounded closed* sets. They have a finite volume (when a volume is defined, which is why we require *Jordan-measurability* as well) and continuous functions are bounded and therefore integrable. Later we move on to integration on open, possibly unbounded sets.

For differentiation, we start by considering function on *open* sets because otherwise we would have to worry about what it means to differentiate a function at a point on the boundary. On the other hand, it does not bother us if the function in question is not bounded.]

5.1. **Directional derivative; derivative.** Let x, y denote the coordinates[10]. We think of these as real-valued functions on U. Let $\vec{p} = (x_0, y_0)$ be a point in U. Recall that the partial derivatives of f at $\vec{p}$ are defined as follows:

(1) the partial derivative w.r.to x:

$$f_x(\vec{p}) \underset{notation}{=} \frac{\partial f}{\partial x}(\vec{p}) = \lim_{t \to 0} \frac{f(x_0 + t, y_0) - f(x_0, y_0)}{t}$$

[10]For general n it is more convenient to use the notation $x_1, \ldots, x_n$, or as Spivak does, $x^1, \ldots, x^n$ for the coordinates.

(2) the partial derivative w.r.to y:

$$f_y(\vec{p}) \underset{notation}{=} \frac{\partial f}{\partial y}(\vec{p}) = \lim_{t\to 0} \frac{f(x_0, y_0 + t) - f(x_0, y_0)}{t}$$

whenever the limit exists.

Let $\vec{e}_x = (1,0)$ and $\vec{e}_y = (0,1)$ denote the standard basis vectors. We can rewrite the above definitions as follows:

$$f_x(\vec{p}) = \lim_{t\to 0} \frac{f(\vec{p} + t\vec{e}_x) - f(\vec{p})}{t}$$
$$f_y(\vec{p}) = \lim_{t\to 0} \frac{f(\vec{p} + t\vec{e}_y) - f(\vec{p})}{t}$$

More generally, given a vector $\vec{u} \in \mathbb{R}^2$, the *directional derivative of f along u at $\vec{p}$* is the limit

$$\lim_{t\to 0} \frac{f(\vec{p} + t\vec{u}) - f(\vec{p})}{t}$$

if it exists. We need a notation that takes into account the point $\vec{p}$ (which at the moment we are keeping fixed). We adapt Spivak's notation and set

$$Df(\vec{p})[\vec{u}] = \lim_{t\to 0} \frac{f(\vec{p} + t\vec{u}) - f(\vec{p})}{t}$$

We will call this the "the directional derivative of f at $\vec{p}$ along the vector $\vec{u}$". Note that

(1) if $\vec{u} = 0$ the expression vanishes, and

(2) if $\lambda \in \mathbb{R}$ is a nonzero scalar, then

$$Df(\vec{p})[\lambda\vec{u}] = \lim_{t\to 0} \frac{f(\vec{p} + t\lambda\vec{u}) - f(\vec{p})}{t} = \lambda \lim_{t\to 0} \frac{f(\vec{p} + t\lambda\vec{u}) - f(\vec{p})}{\lambda t} = \lambda Df(\vec{p})[\vec{u}]$$

Let us see if we can express $Df(\vec{p})[\vec{u}]$ in terms of $f_x(\vec{p})$ and $f_y(\vec{p})$. Write the vector $\vec{u}$ in terms of its components: $\vec{u} = u_x\vec{e}_x + u_y\vec{e}_y$. Then

$$\begin{aligned} Df(\vec{p})[\vec{u}] &= \lim_{t\to 0} \frac{f(\vec{p} + tu_x\vec{e}_x + tu_y\vec{e}_y) - f(\vec{p})}{t} \\ &= \lim_{t\to 0} \frac{f(\vec{p} + tu_x\vec{e}_x + tu_y\vec{e}_y) - f(\vec{p} + tu_y\vec{e}_y)}{t} + \lim_{t\to 0} \frac{f(\vec{p} + tu_y\vec{e}_y) - f(\vec{p})}{t} \ . \end{aligned}$$

Now, the second term equals $f_y(\vec{p})u_y$. As for the second term, we wave hands as follows:

$$\begin{aligned} \lim_{t\to 0} \frac{f(\vec{p} + tu_x\vec{e}_x + tu_y\vec{e}_y) - f(\vec{p} + tu_y\vec{e}_y)}{t} &= \lim_{s\to 0}\lim_{t\to 0} \frac{f(\vec{p} + tu_x\vec{e}_x + su_y\vec{e}_y) - f(\vec{p} + su_y\vec{e}_y)}{t} \\ &= \lim_{s\to 0} f_x(\vec{p} + su_y\vec{e}_y)u_x \\ &= f_x(\vec{p})u_x \end{aligned}$$

To make this work some assumptions have to be made – the last step above, for example requires f_x to be continuous at p. See below.

We can now make the following

Definition 5.1. the function f *is differentiable at* $\vec{p}$ if the following holds: there is a linear map[11] $Df(\vec{p}) : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\lim_{\vec{u}\to 0} \frac{|f(\vec{p}+\vec{u}) - f(\vec{p}) - Df(\vec{p})[\vec{u}]|}{||\vec{u}||} = 0$$

where $||\vec{u}|| = \sqrt{u_x^2 + u_y^2}$.

If the derivative exists, then

(1) the directional derivative $Df(\vec{p})[\vec{u}]$ exists along all vectors $\vec{u}$ (in particular the partial derivatives the partial derivatives $f_x(\vec{p}),\ f_y(\vec{p})$ exist) *and is a linear map* $\mathbb{R}^2 \to \mathbb{R}$. That is,

$$Df(\vec{p})[\lambda_1\vec{u}_1 + \lambda_2\vec{u}_2] = \lambda_1 Df(\vec{p})[\vec{u}_1] + \lambda_2 Df(\vec{p})[\vec{u}_2]$$

(2) the directional derivative $Df(\vec{p})[\vec{u}]$ exists along all vectors $\vec{u}$ (in particular the partial derivatives the partial derivatives $f_x(\vec{p}),\ f_y(\vec{p})$ exist) and

$$Df(\vec{p})[u_x\vec{e}_x + u_y\vec{e}_y] = f_x(\vec{p})u_x + f_y(\vec{p})u_y$$

The conditions (1) and (2) are not quite enough to guarantee differentiability (even continuity) of f.

Exercise (This is the case $n = 1$.) Let $f : (-1, 1) \to \mathbb{R}$ be the one of the following functions: (a) $f(x) = |x|$, (b)$f(x) = x$. In each case determine if the function is differentiable at the origin. If it is, determine the corresponding linear map $\mathbb{R} \to \mathbb{R}$. More generally, if f is a differentiable function in the "classical" sense on an open interval I, what is the corresponding $Df(p) : \mathbb{R} \to \mathbb{R}$ at point $p \in I$?

Exercise Prove that the function $f(x, y) = \sqrt{|xy|}$ is not differentiable at the origin.

5.2. **Continuously differentiable (C^1) functions.** A function with continuous first partial derivatives is said to be *continuously differentiable* or just C^1.

Proposition 5.2. *If the partial derivatives* $f_x,\ f_y$ *exist at all points of* U *and are continuous functions, then* f *is differentiable at all points of* U.

[11]This map is called the differential – and sometimes (confusingly) the derivative – of f at $\vec{p}$.

Proof. This is Theorem 2-8 of Spivak, specialised to $n = 2, m = 1$. We will have to use Definition 5.1 (3) above.

Let a nonzero vector $\vec{u} = u_x\vec{e}_x + u_y\vec{e}_y$ be given. Then

$$\begin{aligned} f(\vec{p}+\vec{u}) - f(\vec{p}) - f_x(\vec{p})u_x - f_y(\vec{p})u_y \\ &= f(\vec{p}+u_x\vec{e}_x+u_y\vec{e}_y) - f(\vec{p}) - f_x(\vec{p})u_x - f_y(\vec{p})u_y \\ &= f(\vec{p}+u_x\vec{e}_x+u_y\vec{e}_y) - f(\vec{p}+u_x\vec{e}_x) - f_y(\vec{p})u_y \\ &+ f(\vec{p}+u_x\vec{e}_x) - f(\vec{p}) - f_x(\vec{p})u_x \end{aligned}$$

which gives

$$\begin{aligned} \frac{|f(\vec{p}+\vec{u}) - f(\vec{p}) - f_x(\vec{p})u_x - f_y(\vec{p})u_y|}{||\vec{u}||} &\leq \frac{|f(\vec{p}+u_x\vec{e}_x+u_y\vec{e}_y) - f(\vec{p}+u_x\vec{e}_x) - f_y(\vec{p})u_y|}{||\vec{u}||} \\ &+ \frac{|f(\vec{p}+u_x\vec{e}_x) - f(\vec{p}) - -f_x(\vec{p})u_x|}{||\vec{u}||} \\ &\leq \frac{|f(\vec{p}+u_x\vec{e}_x+u_y\vec{e}_y) - f(\vec{p}+u_x\vec{e}_x) - f_y(\vec{p})u_y|}{||\vec{u}||} \\ &+ \frac{|f(\vec{p}+u_x\vec{e}_x) - f(\vec{p}) - -f_x(\vec{p})u_x|}{||\vec{u}||} \end{aligned}$$

Let us consider, in turn, the limits of each of the two terms on the right as $\vec{u} \to 0$, starting with the second term.

Consider the function

$$h(s) = f(\vec{p} + s\vec{e}_x)$$

as s varies from 0 to u_x. This is differentiable w.r.to s with derivative $h'(s) = f_x(\vec{p} + s\vec{e}_x)$. By the mean-value theorem in one variable

$$f(\vec{p}+u_x\vec{e}_x) - f(\vec{p}) = h(u_x) - h(0) = h'(u_x^*)u_x = f_x(\vec{p}+u_x^*\vec{e}_x)u_x$$

for some u_x^* (not necessarily unique, depending on u_x) between 0 and u_x. (I use this phrasing rather than "$u_x^* \in [0, u_x]$ "since u_x could be negative.) So we have

$$\frac{|f(\vec{p}+u_x\vec{e}_x) - f(\vec{p}) - -f_x(\vec{p})u_x|}{||\vec{u}||} = |f_x(\vec{p}+u_x^*\vec{e}_x) - f_x(\vec{p})|\frac{|u_x|}{||\vec{u}||} \leq |f_x(\vec{p}+u_x^*\vec{e}_x) - f_x(\vec{p})|$$

As $\vec{u} \to 0$ the expression $|f_x(\vec{p} + u_x^*\vec{e}_x) - f_x(\vec{p})|$ goes to zero by the assumed continuity of f_x. (Note that $u_x^* \to 0$ as $\vec{u} \to 0$ since u_x^* is sandwiched between 0 and u_x.

Applying the mean-value theorem to the function

$$s \mapsto f(\vec{p} + u_x\vec{e}_x + s\vec{e}_y)$$

we get, as above

$$\frac{|f(\vec{p}+u_x\vec{e}_x+u_y\vec{e}_y) - f(\vec{p}+u_x\vec{e}_x) - f_y(\vec{p})u_y|}{||\vec{u}||} \leq |f_y(\vec{p}+u_x\vec{e}_x+u_y^*\vec{e}_y)) - f_y(\vec{p})|$$

for some u_y^* (not necessarily unique, depending on u_x, u_y) between 0 and u_y. Again the expression $|f_y(\vec{p} + u_x\vec{e}_x + u_y^*\vec{e}_y)) - f_y(\vec{p})$ goes to zero as $\vec{u} \to 0$, by the assumed continuity of f_y. □

5.3. **Functions from $\mathbb{R}^n$ with values in $\mathbb{R}^m$.** There are many natural contexts in which we have to consider more than one function at a time. In Physics, for example, the electric field in a region of space has three components at each point, and should be thought of as a ($\mathbb{R}^3$) vector-valued function, possibly varying also with time. When we are differentiating or integrating functions it is often convenient to *change coordinates*. This process is best formalised as a map between (open subsets of) $\mathbb{R}^n$.

Here is the basic definition, adapted from Definition 5.1 (3) : a function $\vec{f} : \mathbb{R}^n \supset U \to \mathbb{R}^m$ is *differentiable* at a point $\vec{p} \in U$ if there is a linear map (the *derivative*) of $D\vec{f}(\vec{p}) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\vec{u}\to 0} \frac{\|\vec{f}(\vec{p}+\vec{u}) - \vec{f}(\vec{p}) - Df(\vec{p})[\vec{u}]\|}{\|\vec{u}\|} = 0$$

where the norms of vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$ are the usual ones. (In fact any norm will do.)

Let $\vec{e}_1, \ldots, \vec{e}_n$ be the standard basis of $\mathbb{R}^n$ and $\vec{\tilde{e}}_1, \ldots, \vec{\tilde{e}}_m$ be the standard basis of $\mathbb{R}^m$. For $\vec{x} \in \mathbb{R}^n$ write $\vec{x} = x^1\vec{e}_1 + \cdots + x^n\vec{e}_n$ and

$$\vec{f}(\vec{x}) = f^1(\vec{x}))\vec{\tilde{e}}_1 + \cdots + f^m(\vec{x})\vec{\tilde{e}}_m,\ \vec{x} \in U$$

so that for fixed $\vec{x}$ the components of the vector $\vec{f}(\vec{x}) \in \mathbb{R}^m$ are the real numbers $f^j(\vec{x})), j = 1, \ldots, m$. Thus a $\mathbb{R}^m$-valued function $\vec{f}$ is just a m-tuple of real-valued functions $f^j, j = 1, \ldots, m$.

Here is the basic result that links the derivative with partial derivatives:

Proposition 5.3. *Suppose the partial derivatives of the component functions*

$$\frac{\partial f^j}{\partial x^i}(\vec{p})\ i = 1, \ldots, n,\ j = 1, \ldots, m$$

exist at all points $\vec{p} \in U$ and are continuous functions on U. Then the derivative $D\vec{f}(\vec{p})$ exists at all points $\vec{p} \in U$ and is given by the partial derivatives. Explicitly:

$$D\vec{f}(\vec{p})[\sum_{i=1}^{n} u^i\vec{e}_i] = \sum_{i=1}^{n} u^i \sum_{j=1}^{m} \frac{\partial f^j}{\partial x^i}(\vec{p})\vec{\tilde{e}}_j$$

We also write this as a matrix equation

$$\begin{pmatrix} (D\vec{f}(\vec{p})[\vec{u}])^1 \\ \cdots \\ (D\vec{f}(\vec{p})[\vec{u}])^m \end{pmatrix} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f^m}{\partial x^1} & \cdots & \frac{\partial f^m}{\partial x^n} \end{pmatrix} \begin{pmatrix} u^1 \\ \cdots \\ u^n \end{pmatrix}$$

Remark 5.4. When $m = 1$ and $n = 2$, we recover (after changing the notation appropriately to match the earlier one):

$$Df(\vec{p})[u^x\vec{e}_x + u_y\vec{e}_y] = (u_x\frac{\partial f}{\partial x}(\vec{p}) + u_y\frac{\partial f}{\partial y}(\vec{p}))\vec{\tilde{e}}_1 = u_x\frac{\partial f}{\partial x}(\vec{p}) + u_y\frac{\partial f}{\partial y}(\vec{p})$$

because$= \vec{\tilde{e}}_1$ is the "standard basis" of $\mathbb{R}$ which is just the number 1.

Remark 5.5. Let us make even more explicit the statement of the previous proposition: If $\vec{f}$ is a C^1 map from an open set $U \subset \mathbb{R}^n$ to $\mathbb{R}^m$ (i.e., the component functions are continuously differentiable), then for any $\vec{p} \in U$, we have

$$\lim_{\vec{u}\to 0} \frac{\| \begin{pmatrix} f(\vec{p}+\vec{u})^1 \\ \cdots \\ f(\vec{p}+\vec{u})^m \end{pmatrix} - \begin{pmatrix} f(\vec{p})^1 \\ \cdots \\ f(\vec{p})^m \end{pmatrix} - \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(\vec{p}) & \cdots & \frac{\partial f^1}{\partial x^n}(\vec{p}) \\ \cdots & \cdots & \cdots \\ \frac{\partial f^m}{\partial x^1}(\vec{p}) & \cdots & \frac{\partial f^m}{\partial x^n}(\vec{p}) \end{pmatrix} \begin{pmatrix} u^1 \\ \cdots \\ u^n \end{pmatrix} \|}{\|\vec{u}\|} = 0$$

5.4. **The Chain Rule in one dimension.** Before proceeding, we first revise the Chain Rule of differentiation in one dimension.

If you find this section confusing, you can skip the next remark and the computation that justifies the chain rule, and just read the statement and proof of Theorem 5.5, Rudin.

We begin with a preliminary

Remark 5.6. Let f be a real-valued function defined on an open interval I, and let $x \in I$. If f is differentiable at $x_0 \in I$, define the function u_{f,x_0} on I by

$$u_{f,x_0}(x) = \begin{cases} \frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0} \ if \ x \neq x_0 \\ 0 \ if \ x = x_0 \end{cases}$$

Clearly $u_{f,x_0}(x) \to 0$ as $x \to x_0$. In other words, u_{f,x_0} is continuous at x_0. Conversely, if (for some real number λ) we have

$$f(x) - f(x_0) - \lambda(x - x_0) = (x - x_0)u_{f,x_0}(x)$$

for some function u_{f,x_0} which vanishes at x_0 and is continuous there, then f is differentiable at x_0 with derivative λ.

Now let $f : (a, b) \to \mathbb{R}$ be a function, and let $\gamma : (c, d) \to (a, b)$ be another such. We will call γ a "re-parametrisation" (usually this term implies that γ is one-one, onto, and γ and γ^{-1} are both differentiable, but this is not needed at the moment). We will let t denote the coordinate that runs from c to d, and x the coordinate that runs from a to b

Let g denote the composite function $f \circ \gamma$:

$$g(t) = f(\gamma(t))$$

Let $t_0 \in (c,d)$ and set $x_0 = \gamma(t_0)$. Suppose f is differentiable at x_0 with derivative $f'(x_0)$ and γ is differentiable at t_0 with derivative $\dot\gamma(t_0)$. Then, for $t \in (c,d)$ we have

$$\begin{aligned}
g(t) - g(t_0) - f'(x_0)\dot\gamma(t_0)(t-t_0) &= f(\gamma(t)) - f(x_0) - f'(x_0)\dot\gamma(t_0)(t-t_0)\\
&= f(\gamma(t)) - f(x_0) - f'(x_0)(\gamma(t)-x_0)\\
&\quad + f'(x_0)(\gamma(t)-x_0) - f'(x_0)\dot\gamma(t_0)(t-t_0)\\
&= u_{f,x_0}(\gamma(t))(\gamma(t)-x_0) + f'(x_0)u_{\gamma,t_0}(t)(t-t_0)\\
&= u_{f,x_0}(\gamma(t))(\gamma(t)-x_0-\dot\gamma(t_0)(t-t_0))\\
&\quad + u_{f,x_0}(\gamma(t))\dot\gamma(t_0)(t-t_0)\\
&\quad + f'(x_0)u_{\gamma,t_0}(t)(t-t_0)\\
&= u_{f,x_0}(\gamma(t))u_{\gamma,t_0}(t)(t-t_0)\\
&\quad + u_{f,x_0}(\gamma(t))\dot\gamma(t_0)(t-t_0)\\
&\quad + f'(x_0)u_{\gamma,t_0}(t)(t-t_0)\\
&= \{u_{f,x_0}(\gamma(t))u_{\gamma,t_0}(t) + u_{f,x_0}(\gamma(t))\dot\gamma(t_0) + f'(x_0)u_{\gamma,t_0}(t)\}(t-t_0)
\end{aligned}$$

which proves that g is differentiable at t_0 with derivative $f'(x_0)\dot\gamma(t_0)$.

5.5. **Chain rule in general.** Suppose that we are given open sets $V \subset \mathbb{R}^l$, $U \subset \mathbb{R}^n$ and a map $\vec\gamma : V \to U$. Suppose given some functions $f_1, \ldots, f_m$ on an open set U in $\mathbb{R}^n$, which we gather into a vector-valued function $\vec f : U \to \mathbb{R}^m$. Let $\vec g : V \to \mathbb{R}^m$ denote the composite function $\vec f \circ \vec\gamma$. So we have the chain of maps[12]

$$V \underset{\vec\gamma}{\longrightarrow} U \underset{\vec f}{\longrightarrow} \mathbb{R}^m$$

Let $\vec q \in V$ and set $\vec p = \vec\gamma(\vec q)$. Then here is the

Chain rule in abstract form: *if $\vec\gamma$ is differentiable at $\vec q$ and $\vec f$ is differentiable at $\vec p$, then $\vec g$ is differentiable at $\vec q$ and*

$$D\vec g(\vec q) = D\vec f(\vec p) \circ D\vec\gamma(\vec q)$$

In words: "the derivative of the composite function is the composite oxf the derivatives".

In particular, if $\vec f$ and $\vec\gamma$ are C^1, so is $\vec g = \vec f \circ \vec\gamma$.

It is important to be able to write out the Chain rule explicitly. As an exercise, let us consider this

[12] In practice, $\vec f$ and $\vec\gamma$ will have very different roles. This is reflected in the notation – Greek letter for the "reparametrisation" and Latin for the "dependent" variable – and my unconscious use of "map" versus "function" though they literally mean the same thing in Mathematics. Note that $\vec\gamma$ need not be injective/surjective/bijective. In many interesting cases it *is* bijective, and the term "reparametrisation" is well-merited.

5.6. **Important special case of the Chain Rule.** Let f be a real-valued C^1 function on an open set $U \in \mathbb{R}^2$ (we will use coordinates[13] x_1, x_2 rather than x, y), and let us "make a coordinate substitution":

$$x_1 = \gamma_1(t_1, t_2)$$
$$x_2 = \gamma_2(t_1, t_2)$$

We will suppose that the functions $\gamma_1,\ \gamma_2$ together define a map $\vec{\gamma} : V \to U$ for a suitable open set $V \subset \mathbb{R}^2$, with t_1, t_2 the coordinates on V. In other words, for a point $\vec{q} \in V$ with coordinates (t_1, t_2), $\vec{p} \equiv \vec{\gamma}(\vec{q})$ has coordinates $(\gamma_1(t_1, t_2), \gamma_2(t_1, t_2))$.

Let g denote the composite function $f \circ \vec{\gamma}$:

$$g(\vec{q}) = f(\vec{\gamma}(\vec{q}))\ .$$

Claim: the Chain Rule implies (is in fact equivalent to)

$$\begin{aligned} \frac{\partial g}{\partial t_1}(\vec{q}) &= \frac{\partial f}{\partial x_1}(\vec{p})\frac{\partial \gamma_1}{\partial t_1}(\vec{q}) + \frac{\partial f}{\partial x_2}(\vec{p})\frac{\partial \gamma_2}{\partial t_1}(\vec{q}) \\ \frac{\partial g}{\partial t_2}(\vec{q}) &= \frac{\partial f}{\partial x_1}(\vec{p})\frac{\partial \gamma_1}{\partial t_2}(\vec{q}) + \frac{\partial f}{\partial x_2}(\vec{p})\frac{\partial \gamma_2}{\partial t_2}(\vec{q}) \end{aligned} \tag{1}$$

Remark 5.7. The equation (1) is often written in more suggestive (but potentially ambiguous) notation in the following form, which you might find it easier to remember.

$$\begin{aligned} \frac{\partial g}{\partial t_1}(\vec{q}) &= \frac{\partial f}{\partial x_1}(\vec{p})\frac{\partial x_1}{\partial t_1}(\vec{q}) + \frac{\partial f}{\partial x_2}(\vec{p})\frac{\partial x_2}{\partial t_1}(\vec{q}) \\ \frac{\partial g}{\partial t_2}(\vec{q}) &= \frac{\partial f}{\partial x_1}(\vec{p})\frac{\partial x_1}{\partial t_2}(\vec{q}) + \frac{\partial f}{\partial x_2}(\vec{p})\frac{\partial x_2}{\partial t_2}(\vec{q}) \end{aligned}$$

Proof of the Claim: The "abstract" version of the Chain rule says:

$$Dg(\vec{q}) = Df(\vec{p}) \circ D\vec{\gamma}(\vec{q})$$

In matrix form

$$\begin{pmatrix} \frac{\partial g}{\partial t_1}(\vec{q}) & \frac{\partial g}{\partial t_2}(\vec{q}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{p}) & \frac{\partial f}{\partial x_2}(\vec{p}) \end{pmatrix} \begin{pmatrix} \frac{\partial \gamma_1}{\partial t_1}(\vec{q}) & \frac{\partial \gamma_1}{\partial t_2}(\vec{q}) \\ \frac{\partial \gamma_2}{\partial t_1}(\vec{q}) & \frac{\partial \gamma_2}{\partial t_2}(\vec{q}) \end{pmatrix} \tag{2}$$

Carrying out the matrix multiplication on the right you get an equality of vectors, which is equivalent to (1) □

[13]The notation is chosen to make the analogy with the one-dimensional Chain rule closer.

5.7. **Important special case of the special case of the Chain Rule.** An important class of "coordinate substitutions" are invertible linear transformations. Let us revisit the previous computations in this context. To keep things simple, assume that $U = \mathbb{R}^2$, (i,e., that f is defined on all of $\mathbb{R}^2$), $V = \mathbb{R}^2$ as well. Suppose the coordinate substitution of the previous subsection is of the form:

$$x_1 = a_1^1 t_1 + a_1^2 t_2 \equiv \gamma_1(t_1, t_2)$$
$$x_2 = a_2^1 t_1 + a_2^2 t_2 \equiv \gamma_2(t_1, t_2)$$

determined by an invertible matrix:

$$A_{\vec{\gamma}} \equiv \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix}$$

For the partial derivatives w.r.to the t_i we get:

$$\frac{\partial g}{\partial t_1}(\vec{q}) = a_1^1 \frac{\partial f}{\partial x_1}(\vec{p}) + a_2^1 \frac{\partial f}{\partial x_2}(\vec{p})$$
$$\frac{\partial g}{\partial t_2}(\vec{q}) = a_1^2 \frac{\partial f}{\partial x_1}(\vec{p}) + a_2^2 \frac{\partial f}{\partial x_2}(\vec{p})$$

5.8. **An abstract version of change of coordinates.** Read this subsection only if you find it helpful.

Let W be a two-dimensional real vector space. Given a point $\vec{p} \in W$, and a real-valued function f defined in a neighbourhood, the derivative $Df(\vec{p})$ (if it exists) is a linear map $W \to \mathbb{R}$ such that

$$\frac{|f(\vec{p}+\vec{u}) - f(\vec{p}) - Df(\vec{p})\vec{u}|}{\|\vec{u}\|} \underset{\|\vec{u}\|\to 0}{\to} 0$$

for some choice of norm on W.

Given a basis $(\vec{e}_1, \vec{e}_2)$, the coordinates of a vector w.r.to the basis give a pair of real numbers, and as the vector varies a pair of functions. If we denote these functions by x_1, x_2, we have

$$\vec{v} = x_1(\vec{v})\vec{e}_1 + x_2(\vec{v})\vec{e}_2$$

Given a function f defined in a neighbourhood of a point $\vec{v} \in W$, we set

$$\frac{\partial}{\partial x_1}|_{\vec{v}} f = Df(\vec{v})[\vec{e}_1]$$
$$\frac{\partial}{\partial x_1}|_{\vec{v}} f = Df(\vec{v})[\vec{e}_1]$$

("partial derivation w.r,to x_1 at $\vec{v}$ is the directional derivative at $\vec{v}$ along the basis vector $\vec{e}_1$.") As $\vec{v}$ varies, we get partial derivative operators

$$f \mapsto \frac{\partial f}{\partial x_i}, \ i = 1, 2$$

from C^1 functions to continuous functions.

Let $(\vec{f}_1, \vec{f}_2)$ be a second basis, and let t_1, t_2 denote the corresponding coordinate functions. If the bases are related by

$$\begin{aligned} \vec{f}_1 &= a_1^1\vec{e}_1 + a_2^1\vec{e}_2 \\ \vec{f}_2 &= a_1^2\vec{e}_1 + a_2^2\vec{e}_2 \end{aligned} \tag{3}$$

we have

$$\begin{aligned} t_1\vec{f}_1 + t_2\vec{f}_2 &= t_1(a_1^1\vec{e}_1 + a_2^1\vec{e}_2) + t_2(a_1^2\vec{e}_1 + a_2^2\vec{e}_2) \\ &= (a_1^1t_1 + a_1^2t_2)\vec{e}_1 + (a_2^1t_1 + a_2^2t_2)\vec{e}_2 \\ &= x_1\vec{e}_1 + x_2\vec{e}_2 \end{aligned}$$

which reproduces the relation (3). Now you can check:

$$\begin{aligned} \frac{\partial f}{\partial t_1}(\vec{v}) &= \frac{\partial f}{\partial x_1}(\vec{v})\frac{\partial x_1}{\partial t_1}(\vec{v}) + \frac{\partial f}{\partial x_2}(\vec{v})\frac{\partial x_2}{\partial t_1}(\vec{v}) \\ \frac{\partial f}{\partial t_2}(\vec{v}) &= \frac{\partial f}{\partial x_1}(\vec{v})\frac{\partial x_1}{\partial t_2}(\vec{v}) + \frac{\partial f}{\partial x_2}(\vec{v})\frac{\partial x_2}{\partial t_2}(\vec{v}) \end{aligned}$$

On dropping the argument $\vec{v}$ we get equalities of partial derivative operators on W:

$$\begin{aligned} \frac{\partial}{\partial t_1} &= \frac{\partial x_1}{\partial t_1}\frac{\partial}{\partial x_1} + \frac{\partial x_2}{\partial t_1}\frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial t_2} &= \frac{\partial x_1}{\partial t_2}\frac{\partial}{\partial x_1} + \frac{\partial x_2}{\partial t_2}\frac{\partial}{\partial x_2} \end{aligned}$$

5.9. **Polar coordinates.** We will work out an explicit and very useful case of change of coordinates.

To motivate matters, we work informally to start with. To every point (x, y) in $\mathbb{R}^2$ we associate two real numbers r, θ such that $r \geq 0$

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned}$$

Clearly $r = \sqrt{x^2 + y^2}$ so r is well-defined, but θ is only well-defined modulo 2π. On the other hand the map $(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$ is many-to-one. Nonetheless we would like to express partial derivatives in terms of x, y in terms of those w.r.to r, θ since this turns out to be useful in many contexts.

To make this precise and cast this in terms of the previous discussion, consider the open sets

(1) $V = \{(r, \theta) \in \mathbb{R}^2 | r > 0\}$

(2) $U = \{(x, y) \in \mathbb{R}^2 | (x, y) \neq (0, 0)\}$

and let $\vec{\gamma}$ be the map $(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$. This map is C^1, surjective but not injective. It does have the following important property:

Let $\vec{q} \in V$. The derivative $D\vec{\gamma}[\vec{q}]$ is an invertible linear map. There is an open neighbourhood V_1 of $\vec{q}$ such that $\vec{\gamma}_1 \equiv \vec{\gamma}|_{V_1}$ is a bijective map to its image $U_1 \subset U$, and the inverse map $\vec{\gamma_1^{-1}}$ is also a C^1 bijection. In fact, if $\vec{q} = (r_0, \theta_0)$, we can take V_1 to be the open set:

$$V_1 = \{(r, \theta) \in \mathbb{R}^2 | r > 0, \theta_0 - \pi < \theta < \theta_0 + \pi\}$$

We now proceed to compute the derivative matrix

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

the inverse map γ_1^{-1} has derivative

$$\begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{x}{r} & \frac{y}{r} \\ -\frac{y}{r^2} & \frac{x}{r^2} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} \end{pmatrix}$$

As an immediate application, let us express the Laplacian[14] operator in polar coordinates. Given a C^2 function f,

$$\Delta f \equiv -\frac{\partial}{\partial x}\frac{\partial f}{\partial x} - \frac{\partial}{\partial y}\frac{\partial f}{\partial y}$$

"Laplacian in polar coordinates" abusing notation: By the Chain rule, given any function f,

$$\begin{aligned} \frac{\partial f}{\partial x} &= \cos\theta \frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta} \\ \frac{\partial f}{\partial y} &= \sin\theta \frac{\partial f}{\partial r} + \frac{\cos\theta}{r}\frac{\partial f}{\partial \theta} \end{aligned} \tag{4}$$

This yields

$$\begin{aligned} \Delta f = &-\{\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\} \circ \{\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\} f \\ &- \{\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\} \circ \{\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\} f \\ = &-\frac{\partial}{\partial r}(\frac{\partial f}{\partial r}) - \frac{1}{r}\frac{\partial f}{\partial r} - \frac{1}{r^2}\frac{\partial}{\partial\theta}(\frac{\partial f}{\partial\theta}) \end{aligned}$$

(You are urged to check that last equality; several cancellations and simplifications happen along the way.)

[14] This is the "analyst's Laplacian"; the "geometer's Laplacian" does not have the negative sign.

"Laplacian in polar coordinates" taking care with notation: If $f = g \circ \vec{\gamma}^{-1}$, equation (4) should be written

$$\frac{\partial f}{\partial x} = \{\cos\theta \frac{\partial[f \circ \vec{\gamma}]}{\partial r} - \frac{\sin\theta}{r}\frac{\partial[f \circ \vec{\gamma}]}{\partial \theta}\} \circ \vec{\gamma}^{-1}$$
$$\frac{\partial f}{\partial y} = \{\sin\theta \frac{\partial[f \circ \vec{\gamma}]}{\partial r} + \frac{\cos\theta}{r}\frac{\partial[f \circ \vec{\gamma}]}{\partial \theta}\} \circ \vec{\gamma}^{-1}$$

and the equation for the Laplacian becomes

$$\Delta f = -\{\frac{\partial}{\partial r}(\frac{\partial[f \circ \vec{\gamma}]}{\partial r}) + \frac{1}{r}\frac{\partial[f \circ \vec{\gamma}]}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial \theta}(\frac{\partial[f \circ \vec{\gamma}]}{\partial \theta})\} \circ \vec{\gamma}^{-1}$$

5.10. **Inverse Function Theorem.** We will not prove this. Here is a statement.

Theorem 5.8. *Let $V \subset \mathbb{R}^n$ be an open set, and $\vec{\gamma} : V \to \mathbb{R}^n$ a C^1 map. Let $\vec{q} \in V$ and set $\vec{p} = \vec{\gamma}(\vec{q})$. Suppose that the differential $D\vec{\gamma}[\vec{q}]$ is invertible as a linear map $\mathbb{R}^n \to \mathbb{R}^n$. Equivalently, the matrix of partial derivatives of $\vec{\gamma}$ at $\vec{q}$,*

$$\begin{pmatrix} \frac{\partial \gamma^1}{\partial t^i}(\vec{q}) & \cdots & \frac{\partial \gamma^1}{\partial t^n}(\vec{q}) \\ \cdots & \cdots & \cdots \\ \frac{\partial \gamma^n}{\partial t^1}(\vec{q}) & \cdots & \frac{\partial \gamma^n}{\partial t^n}(\vec{q}) \end{pmatrix}$$

has nonzero determinant.

Then there is an open V_1 such that

(1) $\vec{q} \in V_1 \subset V$,

(2) $\vec{\gamma}_1 \equiv \vec{\gamma}|_{V_1}$ (which means $\vec{\gamma}$ restricted to V_1) is injective,

(3) $U_1 \equiv \vec{\gamma}(V_1)$ is open, and

(4) the inverse map $\vec{\gamma}_1^{-1} : U_1 \to V_1$ is C^1.

Remark 5.9. A C^1 *diffeomorphism* between open subsets[15] $U, V \subset \mathbb{R}^n$ is a bijection $\vec{\gamma} : V \to U$ such that both $\vec{\gamma}$ and its inverse $\vec{\gamma}^{-1}$ are C^1 maps. Given such a $\vec{\gamma}$ the differential $D\vec{\gamma}(\vec{q})$ is an invertible linear map at every $\vec{q} \in V$. (Why? Use the Chain Rule.) So the hypothesis of the inverse function theorem is a necessary condition for the conclusion to hold.

Remark 5.10. With the notation of the preceding theorem, assume that the differential $D\vec{\gamma}(\vec{q})$ is invertible at *every* point $\vec{q} \in V$. Then the Inverse Function Theorem guarantees that the image $\vec{\gamma}(V) \subset \mathbb{R}^n$ is open. But the map $\vec{\gamma} : V \to \vec{\gamma}(V)$ need not be injective and so need not possess an inverse. (The point of the Theorem is that *"locally" an inverse exists.*) You have already encountered an example of this phenomenon.

[15] In differential topology this definition extends to manifolds.

5.11. **Interlude on norms.** We consider finite-dmensional real ceror spaces. I assume that you know the definition of a norm on such a vector space W. Here are important examples. *Any* basis $\vec{e}_i$ defines an L_∞ norm:

$$||\vec{w}||_\infty = \sup_i |w_i|$$

where $\vec{w} = \sum_i w_i$. This is sometimes called a "box norm" because of the shape of the "unit ball" $\{\vec{w}|||\vec{w}||_\infty \leq 1\}$. The L^1 norm corresponding to the basis is

$$||\vec{w}||_i = \sum_i |w_i|$$

If W has an inner-product $<,>$, this defines a norm $||.||_2$:

$$||\vec{w}||_2^2 =< \vec{w}, \vec{w} >= \sum_i |w_i'|^2$$

where $\{\vec{u}_i\}$ is any orthonomal basis and $\vec{w} = \sum_i w_i'\vec{u}_i$.

Here is a proof that any two norms on a finite-dimensional vector space are equivalent. Let $||.||$ be an arbitrary norm, and let $||.||_\infty$ correspond to a basis as above.

Step 1: Note that for any $\vec{w}$

$$||\vec{w}|| = ||\sum_i w_i\vec{e}_i|| \underset{\text{triangle inequality}}{\leq} \left\{\sup_i ||\vec{e}_i||\right\} \times dim\ W \times ||\vec{w}||_\infty$$

Step 2: Conclude from Step 1 that the function $\vec{w} \mapsto ||\vec{w}||$ is continuous with the respect to metric topology given by $||.||_\infty$. By the Bolzano-Weiertrass theorem, $||.||$ attains its minimum on the "unit sphere" $\{\vec{w}|||\vec{w}||_\infty = 1\}$. This minimum value cannot be zero (why?); denote this by C. We have then

$$||\vec{w}|| \geq C||\vec{w}||_\infty$$

Given norms $||.||_1$ and $||.||_2$ on vector spaces W_1, W_2 respectively, the *operator norm* on the vector space $L(W_1, W_2)$ of linear maps $T : W \to W'$ is defined by

$$||T|| = \sup\{||T\vec{w}||_2\big|||\vec{w}||_1 = 1\}$$

We have the obvious inequality:

$$||T\vec{w}||_2 \leq ||T|| \times ||\vec{w}||_1$$

Recall that the dual vector space W' to W is the space of linear maps $\lambda : W \to \mathbb{R}$. This induces an operator norm $||.||'$ on W' (we take the norm on $\mathbb{R}$ to be the standard one), so that

$$|\lambda(\vec{w})| \leq ||\lambda||' \times ||\vec{w}||$$

Check that $(W')'$ can be naturally identified with W, and

$$||\vec{w}||'' = ||\vec{w}||$$

5.12. **Interlude on integrals of vector-valued functions.** Consider functions $\vec{f} : I \to W$ where $I = [a, b]$ and W is a finite-dimensional real vector space.

Let us agree that $\vec{f}$ is integrable if $\lambda \circ \vec{f} : I \to \mathbb{R}$ is integrable for every linear $\lambda : W \to \mathbb{R}$. Exercise: Choose any basis $\vec{e}_1, \ldots, \vec{e}_n$ of W, and write $\vec{f}(t) = f_1(t)\vec{e}_1 + \cdots + f_n(t)\vec{e}_n$. Show that $\vec{f}$ is integrable iff each f_i is integrable.

For any integrable $\vec{f}$ set

$$\int_I \vec{f} = \sum_i \{\int_I f_i\}\vec{e}_i$$

Exercise: Show that this is well-defined independent of the choice of basis. Show that the integral of $\vec{f}$ can be defined to be the unique vector such that for any linear $\lambda : W \to \mathbb{R}$,

$$\lambda\Big(\int_0^1 \vec{f}(t)dt\Big) = \lambda(\int_I \vec{f}) = \int_I \lambda \circ \vec{f} = \int_0^1 \lambda\Big(\vec{f}(t)\Big)dt$$

We will use the following inequality:

Proposition 5.11. *For any norm $\|.\|$ on W, we have*

$$\| \int_I \vec{f}\| \leq \int_I \|\vec{f}\|$$

Proof. Let us first consider the case[16] when the norm is induced by an inner-product $<,>$. For any $\vec{w} \in W$, we have

$$| < \vec{w}, \int_0^1 \vec{f}(t)dt > | = | \int_0^1 < \vec{w}, \vec{f}(t) > dt| \leq \int_0^1 | < \vec{w}, \vec{f}(t) > |dt$$
$$\underset{Cauchy-Scwartz}{\leq} \int_0^1 \|\vec{w}\| \times \|\vec{f}(t)\|dt = \|\vec{w}\| \int_0^1 \|\vec{f}(t)\|dt$$

where we have written out $\int_I \vec{f} = \int_0^1 \vec{f}(t)dt$, etc. Taking $\vec{w} = \int_I \vec{f}$, we get

$$\| \int_I \vec{f}\|^2 \leq \| \int_I \vec{f}\| \times \int_I \|\vec{f}\|$$

Since the inequality of the Proposition is trivially true if $\int_I \vec{f} = 0_W$, we can assume $\| \int_I \vec{f}\| \neq 0$ and conclude that

$$\| \int_I \vec{f}\| \leq \int_I \|\vec{f}\|$$

In the case when the norm is not given by an inner product, we argue as follows. Recall that given any norm $\|.\|$ on W, the dual space of linear maps $\lambda : W \to \mathbb{R}$ has a norm:

$$\|\lambda\| = \sup\{|\lambda(\vec{w})| \;\big|\; \|\vec{w}\| = 1\}$$

[16]This ingenious proof is taken from `https://www2.math.upenn.edu/ kazdan/508F08/MVT.pdf`

and in turn

$$\|\vec{w}\| = \sup\{|\lambda(\vec{w})| \;\big|\; \|\lambda\| = 1\}$$

Let us now repeat the above argument with an arbitrary norm. For any λ, we have

$$|\lambda(\int_0^1 \vec{f}(t)dt)| = |\int_0^1 \lambda(\vec{f}(t))dt| \leq \int_0^1 |\lambda(\vec{f}(t))|dt$$
$$\underset{definition\ of\ \|\lambda\|}{\leq} \int_0^1 \|\lambda\| \times \|\vec{f}(t)\|dt = \|\lambda\| \int_0^1 \|\vec{f}(t)\|dt$$

This gives

$$\|\int_I \vec{f}\| = \sup\{\lambda(\int_S \vec{f}) \;\big|\; \|\lambda\| = 1\} \leq \int_I \|\vec{f}\|$$

□

Remark: If we had defined integrals in terms of tagged partitions (i.e., worked with Riemann integrals rather than Darboux integrals), the inequality of the proposition would be an easy consequence of the triangle inequality applied to the Riemann sums. See mathoverflow for proofs of the equivalence (and equality) of Riemann and Darboux integrals.

5.13. **Mean-value theorem.** Consider as in §5.3 a C^1 function $\vec{f} : \mathbb{R}^n \supset U \to \mathbb{R}^m$.

Theorem 5.12. *If $\vec{p} \in U$, $\vec{u} \in \mathbb{R}^n$, and the line segment $[\vec{p}, \vec{p}+\vec{u}] \subset U$, then*

$$\|\vec{f}(\vec{p}+\vec{u}) - \vec{f}(\vec{p})\| \leq \sup_{s\in[0,1]} \|D\vec{f}(\vec{p}+s\vec{u})\| \times \|\vec{u}\|$$

Proof. We have

$$\begin{aligned}\vec{f}(\vec{p}+\vec{u}) - \vec{f}(\vec{p}) &= \int_0^1 \frac{d\vec{f}}{ds}(\vec{p}+s\vec{u})ds \\ &= \int_0^1 D\vec{f}(\vec{p}+s\vec{u})[\vec{u}]ds \\ &= \big[\int_0^1 D\vec{f}(\vec{p}+s\vec{u})ds\big][\vec{u}]\end{aligned}$$

Taking norms, we get

$$\begin{aligned}\|\vec{f}(\vec{p}+\vec{u}) - \vec{f}(\vec{p})\| &= \|\big[\int_0^1 D\vec{f}(\vec{p}+s\vec{u})ds\big][\vec{u}]\| \\ &\leq \|\int_0^1 D\vec{f}(\vec{p}+s\vec{u})ds\| \times \|\vec{u}\| \\ &\leq \int_0^1 \|D\vec{f}(\vec{p}+s\vec{u})\|ds \times \|\vec{u}\| \\ &\leq \sup_{s\in[0,1]} \|D\vec{f}(\vec{p}+s\vec{u})\| \times \|\vec{u}\|.\end{aligned}$$

□

Corollary 5.13. *A C^1 map is Lipschitz on compact sets. That is, given a C^1 map $\vec{f} : U \to \mathbb{R}^m$ as above and a compact set $S \subset U$, there exists a positive constant L_S such that*

$$\|\vec{f}(\vec{x}') - \vec{f}(\vec{x})\| \leq L_S \|\vec{x}' - \vec{x}\|,\ \vec{x}, \vec{x}'\ \in S$$

Proof. Clear if S is convex. I learned the following proof for the general case from https://math.stackexchange.com/questions/2338799/c1-function-on-compact-set-is-lipschitz.

Suppose there exists a sequence of pairs of points $\vec{x}_n, \vec{x}'_n$ in S such that

$$\|\vec{f}(\vec{x}'_n) - \vec{f}(\vec{x}_n)\| \geq n \|\vec{x}'_n - \vec{x}_n\|,\ \vec{x}, \vec{x}'\ \in S$$

By compactness of S, one can find a convergent subsequence, which we relabel, so that $\vec{x}_n \to \vec{x}_0$, $\vec{x}'_n \to \vec{x}'_0$. Since $\vec{f}(S)$ is compact, it follows that $\vec{x}_0 = \vec{x}'_0$. But then for large enough n, the pairs of points $\vec{x}_n, \vec{x}'_n$ lie in a (convex) ball around this limit point, which is a contradiction. □

6. Change of variables in integration

6.1. **Change of variables in one dimension.** You have already encountered this. Here is one version.

Proposition 6.1. *Let $\gamma : (c, d) \to (a, b)$ be a C^1 bijection with C^1 inverse, and suppose that $\dot{\gamma}(t)$ is everywhere positive. Then given any continuous real-valued f on (a, b) and a compact $[a_1, b_1] \subset (a, b)$*

$$\int_{\gamma^{-1}(a_1)}^{\gamma^{-1}(b_1)} f(\gamma(t))\dot{\gamma}(t)dt = \int_{a_1}^{b_1} f(x)dx$$

Proof. (This is a good time to revise the circle of ideas around the fundamental theorem of calculus.) There exists a C^1 function F such that $F' = f$. By the Chain rule

$$f(\gamma(t))\dot{\gamma}(t) = \dot{G}(t)$$

where $G(t) = F(\gamma(t))$. Then

$$\int_{\gamma^{-1}(a_1)}^{\gamma^{-1}(b_1)} f(\gamma(t))\dot{\gamma}(t)dt = \int_{\gamma^{-1}(a_1)}^{\gamma^{-1}(b_1)} \dot{G}(t)dt = G(\gamma^{-1}(b_1)) - G(\gamma^{-1}(a_1))$$
$$= F(b_1) - F(a_1) = \int_{a_1}^{b_1} f(x)dx$$

□

What if $\dot{\gamma}(t)$ is everywhere *negative*? (Note that since γ is a C^1 map with a C^1 inverse $\dot{\gamma}(t)$ cannot vanish anywhere; hence it cannot change sign.) Then

$$\gamma^{-1}(b_1) < \gamma^{-1}(a_1)$$

and we have

$$-\int_{\gamma^{-1}(b_1)}^{\gamma^{-1}(a_1)} f(\gamma(t))\dot{\gamma}(t)dt = -\int_{\gamma^{-1}(b_1)}^{\gamma^{-1}(a_1)} \dot{G}(t)dt = -G(\gamma^{-1}(a_1)) + G(\gamma^{-1}(b_1))$$
$$= -F(a_1) + F(b_1) = \int_{a_1}^{b_1} f(x)dx$$

So we have two cases to consider:

(1) $\dot{\gamma}(t) > 0$ in which case $\gamma^{-}(a_1) < \gamma^{-}(b_1)$ and

$$\int_{\gamma^{-1}(a_1)}^{\gamma^{-1}(b_1)} f(\gamma(t))\dot{\gamma}(t)dt = \int_{a_1}^{b_1} f(x)dx$$

(2) $\dot{\gamma}(t) < 0$ in which case $\gamma^{-1}(b_1) < \gamma^{-1}(a_1)$ and

$$-\int_{\gamma^{-1}(b_1)}^{\gamma^{-1}(a_1)} f(\gamma(t))\dot{\gamma}(t)dt = \int_{a_1}^{b_1} f(x)dx$$

which can be rewritten

$$\int_{\gamma^{-1}(b_1)}^{\gamma^{-1}(a_1)} f(\gamma(t))|\dot{\gamma}(t)|dt = \int_{a_1}^{b_1} f(x)dx$$

since $|\dot{\gamma}(t)| = -\dot{\gamma}(t)$.

Here is a restatement of the previous proposition that covers both possibilities of the sign of $\dot{\gamma}$:

Proposition 6.2. *Let $\gamma : (c,d) \to (a,b)$ be a C^1 bijection with C^1 inverse, and suppose that $\dot{\gamma}(t)$ is nowhere-vanishing (i.e., either positive everywhere or negative everywhere). Then given any continuous real-valued f on (a,b) and a bounded closed interval $I \subset [a,b]$*

$$\int_{\gamma^{-1}(I)} f(\gamma(t))|\dot{\gamma}(t)|dt = \int_I f(x)dx$$

Remark 6.3. Another way to get a statement that covers the two cases is the following. Note that *so far* we have used the notation

$$\int_a^b g(x)dx$$

to mean the Riemann integral of a function g over a closed interval $[a,b]$. In particular, there is no meaning attached to

$$\int_a^b g(x)dx$$

if $a > b$. If we set, when $a > b$,

$$\int_a^b g(x)dx \underset{def}{=} -\int_b^a g(x)dx$$

where on the right hand side we mean the Riemann integral over $[b, a]$, we have

$$\int_{\gamma^{-1}(a_1)}^{\gamma^{-1}(b_1)} f(\gamma(t))\dot{\gamma}(t)dt = \int_{a_1}^{b_1} f(x)dx$$

holding regardless of the sign of $\dot{\gamma}$. This will be relevant when we discuss integration of one-forms along a one-cycle. For the moment, we do not go down this road.

6.2. **Change of variables: statement of the main result.** *Objects on the left are flagged with a tilde.*

Theorem 6.4. *Let $\vec{\gamma} : \tilde{U} \to U$ be a C^1 diffeomorphism between open subsets $\tilde{U}, U$ in $\mathbb{R}^n$.*

If S is a Jordan-measurable subset of $\mathbb{R}^n$ such that its closure $\overline{S}$ (in $\mathbb{R}^n$) is contained in U, then $\tilde{S} \equiv \vec{\gamma}^{-1}(S)$ is Jordan-measurable in $\mathbb{R}^n$ and $\overline{\tilde{S}} \subset \tilde{U}$.

Further, if f is integrable on S, $\tilde{f} \equiv f \circ \vec{\gamma}$ is integrable on $\tilde{S}$, and,

$$\int_{\tilde{S}} \tilde{f}|J_{\vec{\gamma}}| = \int_S f$$

where $J_{\vec{\gamma}}(\vec{t})$ is determinant of the derivative (the linear map $\mathbb{R}^n \to \mathbb{R}^n$ represented by the matrix of partial derivatives of $\vec{\gamma}$) of $\vec{\gamma}$ at $\vec{t} \in \tilde{U}$.

Remark 6.5. Here are some other situations where you can apply the above result:

(1) If f is a continuous function on U, then $\tilde{f} = f \circ \vec{\gamma}$ is continuous on $\tilde{U}$, and the above formula holds for the restrictions (which we have seen earlier are integrable) of f and $\tilde{f}$ on S and $\tilde{S}$.

(2) If f is a continuous function on U with compact support, we have defined its integral

$$\int_U f = \int_{\mathbb{R}^n} f_e$$

where f_e is the extension by zero to $\mathbb{R}^n$. (We have changed the notation for extension by zero, because $\tilde{f}$ now has another meaning!) In this case, $\tilde{f}$ has compact support on $\tilde{U}$, and we have

$$\int_{\tilde{U}} \tilde{f}|J_{\vec{\gamma}}| = \int_U f$$

Note that $|J_{\vec{\gamma}}|$ will not have compact support on $\tilde{U}$, but the product $\tilde{f}|J_{\vec{\gamma}}|$ will be continuous with compact support on $\tilde{U}$.

(3) We have not defined the notion of an integrable function on an arbitrary open subset of $\mathbb{R}^n$. This can be done in several equivalent ways, and then

the following holds: if f is integrable on U, then $\tilde{f}|J_{\tilde{\gamma}}|$ is integrable on $\tilde{U}$ and

$$\int_{\tilde{U}} \tilde{f}|J_{\tilde{\gamma}}| = \int_U f$$

We will give a proof of Theorem 8.1 following C.C. Pugh: *Real Mathematical Analysis*. This proof was pointed out by Soumya Das Gupta, Class of 2023.

6.3. **Preliminaries: Composition of diffeomorphisms.** . *If the statement of the Theorem holds for $\vec{\gamma} : \tilde{U} \to U$ and $\vec{\gamma}_1 : \tilde{U}_1 \to \tilde{U}$, then it holds for the composite map $\vec{\gamma} \circ \vec{\gamma}_1 : \tilde{U}_1 \to U$.*

Proof. Apply the theorem first to f and $\vec{\gamma}$ and then to $(f \circ \vec{\gamma}) \times |J_{\vec{\gamma}}|$ and $\vec{\gamma}_1$ to conclude that

$$\begin{aligned}\int_{\mathbb{R}^n} f(\vec{x})d\vec{x} &= \int_{\mathbb{R}^n} f(\vec{\gamma}(\vec{t}))|J_{\vec{\gamma}}(\vec{t})|d\vec{t} \\ &= \int_{\mathbb{R}^n} f(\vec{\gamma}(\vec{\gamma}_1(\vec{t_1})))|J_{\vec{\gamma}}(\vec{\gamma}_1(\vec{t_1}))||J_{\vec{\gamma}_1}(\vec{t_1})|d\vec{t_1} \\ &= \int_{\mathbb{R}^n} f(\vec{\gamma} \circ \vec{\gamma}_1(\vec{t_1}))|J_{\vec{\gamma}\circ\vec{\gamma}_1}(\vec{t_1})|d\vec{t_1}\end{aligned}$$

In the last step we have used the Chain rule to argue that

$$J_{\vec{\gamma}}(\vec{\gamma}_1(\vec{t_1}))J_{\vec{\gamma}_1}(\vec{t_1}) = J_{\vec{\gamma}\circ\vec{\gamma}_1}(\vec{t_1}) \ .$$

□

6.4. **Preliminaries: determinants and volumes.** . We first deal with the case when $\tilde{U} = U = \mathbb{R}^n$ and $\tilde{\gamma}$ is a linear map given by multiplication by an invertible matrix A^{-1}. Check that in this case, the statement of the Theorem 8.1 reduces to:

Proposition 6.6. *Let $A : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then $A(S)$ is Jordan-measurable for any Jordan-measurable S, and*

$$vol(A(S)) = |\det A| vol(S) \tag{5}$$

Proof. We divide this into steps. We use "for all S" as short-hand for "all Jordan-measurable S".

Step 1: *Fix A. Suppose that for any rectangle R, $A(R)$ is Jordan-measurable and $vol(A(R)) = |\det A| vol(R)$. Then if S is Jordan-measurable so is $A(S)$.* We have $A(\partial S) = \partial A(S)$, so it suffices[17] to check: if T is a compact set with content zero, so is $A(T)$. Let $\epsilon > 0$ be given and let $R_1, \ldots, R_n$ be rectangles such that

- $T \subset R_1 \cup \cdots \cup R_n$, and $\sum_i vol(R_i) < \frac{\epsilon}{2|\det\ A|}$.

Then

[17]Closure, interior, boundary etc. are preserved under homeomorphisms.

- $A(T) \subset A(R_1) \cup \cdots \cup A(R_n)$, and $\sum_i vol(A(R_i)) < \frac{\epsilon}{2}$.

We are nearly done, except that the sets $A(R_i)$ are in general not rectangles, but parallelograms (and in dimensions > 2, "parallelopipeds"). But by assumption, each is the image of a rectangle by A and hence Jordan-measurable. For each i, we can find finitely many rectangles R_i^j such that $A(R_i) \subset \cup_j R_i^j$ and $\sum_j vol(R_j^i) < vol(A(R)_i) + \frac{\epsilon}{2n}$. Then $A(T) \subset \cup_i A(R_i) \subset \cup_{i,j} R_i^j$x and $\sum_{i,j} vol(R_i^j) < \epsilon$.

Step 2: *Fix A. Suppose that for any rectangle R, $A(R)$ is Jordan-measurable and $vol(A(R)) = |\det A| vol(R)$. Then $vol(A(S)) = |\det A| vol(S)$ holds for every S.* Fix a rectangle containing S, and let P be any partition of it. Then

$$L(P, \chi_S) = \sum_{R \in P, R \subset S} area(R) \leq area(S) \leq \sum_{R \in P, R \cap S \neq \emptyset} area(R) = U(P, \chi_S)$$

and since S is Jordan-measurable,

$$\sup_P \sum_{R \in P, R \subset S} area(R) = \inf_P \sum_{R \in P, R \cap S \neq \emptyset} area(R) = area(S) \ .$$

On the other hand, since

$$\bigcup_{R \in P, R \subset S} A(R) \subset A(S) \subset \bigcup_{R \in P, R \cap S \neq \emptyset} A(R)$$

we have

$$\sum_{R \in P, R \subset S} area(A(R)) \leq area(A(S)) \leq \sum_{R \in P, R \cap S \neq \emptyset} area(A(R))$$

(The sets $A(R)$ are not quite disjoint, but $A(R) \cap A(R') = A(R \cap R')$ has Jordan-measure zero if $R \neq R'$.) By assumption, we can rewrite this as follows:

$$|det\ A| \sum_{R \in P, R \subset S} area(R) \leq area(A(S)) \leq |det\ A| \sum_{R \in P, R \cap S \neq \emptyset} area(R)$$

From this it follows that $area(A(S)) = |det\ A| area(S)$.

Step 3: *If the statement of the Proposition (and in particular the equality 5) holds for invertible A, A' and all S, it holds for the product AA' and all S.* Use multiplicativity of the determinant.

Step 4: Using Proposition 2.18 of Artin's *Algebra*[18], Prentice-Hall (1991), we can write any A as a product of elementary matrices. By Step 3, it suffices to check that the statement of the Proposition holds for any elementary matrix and any S. By Steps 1 and 2, it suffices to check this for any rectangle R. The cases when the effect of A is to permute two coordinates or to scale one coordinate are straightforward. If the effect of A is to add a multiple of one coordinate to another, use Fubini to reduce to (these) two dimensions. □

[18]You may have a different edition.

6.5. **The case when $\tilde{S} = \tilde{R}$ a cube.** We first prove a special case of the theorem on change of variables. The general case will be addressed in §8 after we introduce partitions of unity.

Theorem 6.7. *Let $\vec{\gamma} : \tilde{U} \to U$ be a C^1 diffeomorphism between open subsets $\tilde{U}, U$ in $\mathbb{R}^n$.*

If $\tilde{R}$ is a cube (i.e., a rectangle with all sides equal) such that $\tilde{R} \subset \tilde{U}$, then $R \equiv \vec{\gamma}(\tilde{R})$ is Jordan-measurable in $\mathbb{R}^n$.

Further, if f is integrable on R, $\tilde{f} \equiv f \circ \vec{\gamma}$ is integrable on $\tilde{R}$, and further,

$$\int_{\tilde{R}} \tilde{f} |J_{\vec{\gamma}}| = \int_R f$$

where $J_{\vec{\gamma}}(\vec{t})$ is determinant of the derivative (the linear map $\mathbb{R}^n \to \mathbb{R}^n$ represented by the matrix of partial derivatives of $\vec{\gamma}$) of $\vec{\gamma}$ at $\vec{t} \in \tilde{U}$.

Proof. Since any integrable function can be written as difference of non-negative integrable functions, we can assume that $f(\vec{x}) \geq 0 \ \forall \vec{x} \in R$.

Step 1: *We will consider only partitions of $\tilde{R}$ into sub-rectangles that are hypercubes with sides of equal length a.* Fix such a partition $P = \{\tilde{R}_i\}$. Recall that $R = \vec{\gamma}(\tilde{R})$; let $R_i = \vec{\gamma}(\tilde{R}_i)$ and set

$$\begin{aligned}
m_i &\equiv \inf\{\tilde{f}(\vec{t}) || \vec{t} \in \tilde{R}_i\} =_b \inf\{f(\vec{x}) || \vec{x} \in R_i\} \\
M_i &\equiv \sup\{\tilde{f}(\vec{t}) || \vec{t} \in \tilde{R}_i\} =_b \sup\{f(\vec{x}) || \vec{x} \in R_i\} \\
m_i' &\equiv \inf\{\tilde{f}(\vec{t}) |J_{\vec{\gamma}}(\vec{t})| || \vec{t} \in \tilde{R}_i\} \\
M_i' &\equiv \sup\{\tilde{f}(\vec{t}) |J_{\vec{\gamma}}(\vec{t})| || \vec{t} \in \tilde{R}_i\}
\end{aligned}$$

The equalities labelled b hold because $\vec{\gamma}$ is a bijection.

We have the following inequalities of functions on R:

$$\sum_i m_i \chi_{R_i^0}(\vec{x}) \leq f(\vec{x}) \sum_i \chi_{R_i^0}(\vec{x}) \leq \sum_i M_i \chi_{R_i^0}(\vec{x})$$

Integrating over R, we get

$$\sum_i m_i vol(R_i^0) \leq \int_{\cup_i R_i^0} f \leq \sum_i M_i vol(R_i^0)$$

Since $vol(R_i) = vol(R_i^0)$ and $\int_{\cup_i R_i^0} f = \int_R f$, this yields:

$$\sum_i m_i vol(R_i) \leq \int_R f \leq \sum_i M_i vol(R_i)$$

Step 2: Fix i. We will estimate $m_i vol(R_i)$ from *below* in terms of a. We introduce some more notation:

(1) $J_i \equiv \sup\{|J_{\vec{\gamma}}(\vec{t})| \big| \vec{t} \in \tilde{R}_i\}$, and

(2) $J_i^* \equiv |J_{\vec{\gamma}}(\vec{t}_i^*)|$ where t_i^* is the centre of the rectangle $\tilde{R}_i$.

For $\vec{t} \in \tilde{R}_i$ we have

$$\tilde{f}(\vec{t}) \geq \tilde{f}(\vec{t})\frac{|J_{\vec{\gamma}}(\vec{t})|}{J_i} \geq m_i'\frac{1}{J_i}$$

Since the inequality $\tilde{f}(\vec{t}) \geq m_i'\frac{1}{J_i}$ holds at every $\vec{t} \in \tilde{R}_i$, we get $m_i \geq m_i'\frac{1}{J_i}$, which in turn implies:

$$m_i vol(R_i) \geq m_i'\frac{vol(R_i)}{J_i} = m_i'\frac{vol(R_i)}{J_i^*}\frac{J_i^*}{J_i} = m_i' vol(\tilde{R}_i)\frac{vol(R_i)}{J_i^* vol(\tilde{R}_i)}\frac{J_i^*}{J_i}$$

We will now estimate *with errors vanishing as* $a \to 0$:

(1) $\frac{J_i^*}{J_i}$ from *below* by 1, and

(2) $\frac{vol(R_i)}{J_i^* vol(\tilde{R}_i)}$ from *below* by 1.

Step 3: By uniform continuity of $|J_{\vec{\gamma}}|$ on the compact set $\tilde{R}$, we have – *independently of* i – $0 \leq J_i - J_i^* \leq D_J(a)$, with $D_J(a) \to 0$ as $a \to 0$. This yields, for every i:

$$\frac{J_i^*}{J_i} = 1 - \frac{J_i - J_i^*}{J_i} \geq 1 - C_J(\vec{\gamma}, \tilde{R}, a)$$

where

$$C_J(\vec{\gamma}, \tilde{R}, a) = \frac{D_J(a)}{\inf_{\vec{t}\in\tilde{R}}\{|J_{\vec{\gamma}}(\vec{t})|\}}$$

(That my earlier argument requires $|J_{\vec{\gamma}}|$ to be differentiable was pointed out by Rajarshi Ghosh, Class of 2023.)

Step 4: Consider next $\frac{vol(R_i)}{J_i^*}$. If the derivative of $\vec{\gamma}$ were a constant (in other words, if $\vec{\gamma}$ were of the form $\vec{\gamma}(\vec{t}) = A\vec{t} + \vec{x}_0$) we would have

$$\frac{vol(R_i)}{|J_i^*|} = \frac{vol(R_i)}{|det\ A|} = vol(\tilde{R}_i)$$

We will show that in general $\frac{vol(R_i)}{J_i^* vol(\tilde{R}_i)} \geq 1 - \ small\ error.$

Consider the derivative $D\vec{\gamma}(\vec{t}_i^*)$ of $\vec{\gamma}$ at $\vec{t}_i^*$. This is given by a $n \times n$ matrix A, with entries the partial derivatives of $\vec{\gamma}$ at $\vec{t}_i^*$, and $det\ A = J_i^*$. Write

$$\vec{\gamma} = A \circ \vec{\gamma}'\ where\ \vec{\gamma}' = A^{-1}\vec{\gamma}$$

By the Chain rule,

$$D\vec{\gamma}(\vec{t}_i^*) = AD\vec{\gamma}'(\vec{t}_1^*) \ \text{(matrix multiplication)}$$

from which we conclude that $D\vec{\gamma}'(\vec{t}_i^*) = I_n$, the identity matrix. Set $R_i' = \vec{\gamma}'(\tilde{R}_i)$. Note that

$$\frac{vol(R_i)}{|J_i^*|} = \frac{vol\big(A(R_i')\big)}{|J_i^*|} = \frac{vol\big(A(R_i')\big)}{|det\ A|} = vol(R_i')$$

Step 5: We will now estimate $vol(R_i')$ in terms of $vol(\tilde{R}_i)$ with small error. We are able to do this because $D\vec{\gamma}'(\vec{t}_i^*)$ (the differential of $\vec{\gamma}'$ at the centre of the square $\tilde{R}_i$) is the identity and $\vec{\gamma}_i'$ is well-approximated over the small square $\tilde{R}_i$ by its derivative at its centre $\vec{t}_i^*$.

First, composing with two translations, we can assume that $\vec{t}_i^* = \vec{\gamma}'(\vec{t}_i^*) = 0_{\mathbb{R}^n}$.

We have, for $\vec{t} \in \tilde{R}_i$

$$\begin{aligned}\vec{\gamma}'(\vec{t}) &= \int_0^1 ds \frac{d\vec{\gamma}'}{ds}(s\vec{t}) \\ &= \Big[\int_0^1 ds\ D\vec{\gamma}'(s\vec{t})\Big](\vec{t})\end{aligned}$$

Since $D\vec{\gamma}'(0_{\mathbb{R}^n}) = I_n$, we can write:

$$\vec{\gamma}'(\vec{t}) - \vec{t} = \int_0^1 ds\ [D\vec{\gamma}'(s\vec{t}) - I_n]\vec{t}$$

Taking norms, we get:

$$\begin{aligned}||\vec{\gamma}'(\vec{t}) - \vec{t}|| &\leq \sup_{s\in[0,1]} ||D\vec{\gamma}'(s\vec{t}) - I_n|| \times ||\vec{t}|| \\ &\leq sup_{\vec{t}'\in\tilde{R}_i} ||D\vec{\gamma}'\big(\vec{t}'\big) - I_n|| \times ||\vec{t}||\end{aligned}$$

Using compactness of $\tilde{R}$ and uniform continuity $\vec{\gamma}$ and its derivative, we have

$$||\vec{\gamma}'(\vec{t}) - \vec{t}|| \leq C(\vec{\gamma}, \tilde{R}, a) \times ||\vec{t}||$$

where $C(\vec{\gamma}, \tilde{R}, a)$ is non-negative, and goes to zero as $a \to 0$.

If $\vec{x} \in R_i'$, then $\vec{x} = \vec{\gamma}'(\vec{t})$ for some $\vec{t} \in \tilde{R}_i$. Writing $\vec{x} = \vec{t} + \gamma'(\vec{t}) - \vec{t}$ gives

$$||\vec{x}|| \leq (1 + C(\vec{\gamma}, \tilde{R}, a))||\vec{t}|| \ .$$

We also have

$$||\vec{t}|| \leq ||\vec{x}|| + C(\vec{\gamma}, \tilde{R}, a)||\vec{t}||$$

which can be iterated to give

$$||\vec{t}|| \leq \frac{1}{1 - C(\vec{\gamma}, \tilde{R}, a)}||\vec{x}||$$

We conclude:

$$(1 - C(\vec{\gamma}, \tilde{R}, a))||\vec{t}|| \leq ||\vec{x}|| \leq (1 + C(\vec{\gamma}, \tilde{R}, a))||\vec{t}|| \ .$$

Note that the above inequalities are true regardless of the norm chosen on $\mathbb{R}^n$. Take as norm

$$||(x_1, \ldots, x_n)|| = \sup\{|x_1|, \ldots, |x_n|\}$$

(the "L^∞"-norm) to conclude:

$$(1 - C(\vec{\gamma}, \tilde{R}, a))\tilde{R}_i \subset R_i' \subset (1 + C(\vec{\gamma}, \tilde{R}, a))\tilde{R}_i$$

This gives the inequalities: to conclude:

$$(1 - C(\vec{\gamma}, \tilde{R}, a))^n vol(\tilde{R}_i) \leq vol(R_i') \leq (1 + C(\vec{\gamma}, \tilde{R}, a))^n vol(\tilde{R}_i)$$

Step 6: The preceding justifies the first of the inequalities below; a similar reasoning yields the other as well:

$$\sum_i m_i' vol(\tilde{R}_i)(1 - C_J(\vec{\gamma}, \tilde{R}, a))(1 - C(\vec{\gamma}, \tilde{R}, a))^n \leq \sum_i m_i vol(R_i)$$
$$\leq \sum_i M_i' vol(\tilde{R}_i)(1 + C_J(\vec{\gamma}, \tilde{R}, a))(1 + C(\vec{\gamma}, \tilde{R}, a))^n$$

As $a \to 0$

$$\sum_i m_i' vol(\tilde{R}_i)(1 - C_J(\vec{\gamma}, \tilde{R}, a))(1 - C(\vec{\gamma}, \tilde{R}, a))^n \nearrow \int_{\tilde{R}} \tilde{f}|J_{\tilde{\gamma}}|$$

$$\sum_i M_i' vol(\tilde{R}_i)(1 + C_J(\vec{\gamma}, \tilde{R}, a))(1 + C(\vec{\gamma}, \tilde{R}, a))^n \searrow \int_{\tilde{R}} \tilde{f}|J_{\tilde{\gamma}}|$$

This forces

$$\sum_i m_i vol(R_i) \nearrow \int_R f$$

$$\sum_i M_i vol(R_i) \searrow \int_R f$$

and finally yields

$$\int_{\tilde{R}} f|J_{\tilde{\gamma}}| = \int_R f$$

□

Exercise: Check the above proof. Why does it not matter that we are considering partitions by equal-sized squares and taking the limit $a \to 0$?

We have to still prove the claim in the Theorem about transforms of Jordan-measurable sets.

Proposition 6.8. *Suppose given a C^1 diffeomorphism $\tilde{\gamma} : U \to \tilde{U}$ of open sets in $\mathbb{R}^n$ and a Jordan measurable subset S of $\mathbb{R}^n$ such that $\overline{S} \subset U$. Then $\tilde{S} \equiv \tilde{\gamma}(S)$ is a Jordan-measurable subset of $\mathbb{R}^n$.*

Proof. ∂S is a compact and and of content zero. So given $\epsilon > 0$, there exist finitely many hypercubes $R_1, \ldots, R_N(\epsilon)$ such that $\partial S \subset \bigcup_i R_i$ and

$$\sum_i vol(R_i) < \epsilon$$

By shrinking the R_i if necessary, we can assume that each $R_i \subset U$. Clearly

$$\partial(\tilde{S}) = \tilde{\gamma}(\partial S) \subset \bigcup_i \tilde{\gamma}(R_i)$$

By Corollary 5.13,

$$\tilde{\gamma}(R_i) \subset \tilde{R}_i$$

where $\tilde{R}_i$ is a hypercube each of whose sides is $L_{\partial S}$ times the side of R_i. So

$$\partial(\tilde{S}) \subset \bigcup_i \tilde{\gamma}(tR_i)$$

and

$$\sum_i vol(\tilde{R}_i) < L_{\partial S} \times \epsilon^n$$

This proves that $\partial(\tilde{S})$ has content zero and so $\tilde{S}$ is Jordan-measurable. □

7. Partitions of unity

These are used extensive in analysis, especially in geometric contexts. Here is the basic result:

Theorem 7.1. *Let U be an open subset of $\mathbb{R}^n$ and $\{U_\alpha\}_{\alpha\in J}$ be an open cover of U. Then there exists a family $\{\psi_i\}_{i\in I}$ of continuous (even C^∞) functions with compact support such that*

(1) $0 \le \psi_i(x) \le 1, x \in U$

(2) for every i, $supp(\psi_i)$ is contained in some $U_{\alpha(i)}$,

(3) given any point $x \in U$, there exists a neighbourhood U_x such that all but finitely many ψ_i vanish on U_x,

(4) $\sum_i \psi_i(x) = 1$ *(the sum makes sense because of (3)).*

The functions $\{\psi_i\}_{i\in I}$ are said to constitute a *partition of unity subordinate to the cover* $\{U_\alpha\}_{\alpha\in J}$. In fact, the functions ψ_i can be chosen to be infinitely differentiable, and are then said to constitute a *smooth partition of unity.*

Remark 7.2. Note that the indexing sets I and J can be different. Can we take $I = J$ and arrange that $supp(\psi_\alpha) \subset U_\alpha$? This is the case if (a) we drop the requirement that the ψ_α have compact support, or (b) the open sets U_α are *relatively compact*, that is, if each $\overline{U}_\alpha$ (closure in U) is compact.

We will need the following special case. First, an

Exercise: Any open subset U of $\mathbb{R}^n$, we can find a countable family of (hyper)cubes (i.e, rectangles of the form $I_1 \times \ldots I_n$ with all the intervals I_j of equal length) R_j whose interiors R_j^0 form an open cover of U

Theorem 7.3. *Let $\{R_i\}_{i\in\mathbb{N}}$ be a countable family of hypercubes as above. Then there exists a family of be continuous functions with compact support $\{\psi_i\}_{i\in\mathbb{N}}$ such that*

(1) $0 \le \psi_i(x) \le 1, x \in U$,

(2) *for every* i, $supp(\psi_i)$ *is contained in* R_i^0,

(3) *given any point* $x \in U$, *there exists a neighbourhood* U_x *such that all but finitely many* ψ_i *vanish on* U_x,

(4) $\sum_i \psi_i(x) = 1,\ x \in U$ *(the sum makes sense because of (3)).*

We will not prove the existence of partitions of unity. We *will* now work out a series of examples illustrating their construction.

Example 1: Let U be the interval $(0,1)$. Consider the open cover consisting of the subintervals $U_1 = (0, 3/4)$ and $U_2 = (1/4, 1)$. Let $\tilde{\psi}_1$ be defined by

$$\tilde{\psi}_1(t) = \begin{cases} 2/3 - t & if\ 0 < t < 2/3 \\ 0 & if\ 2/3 \le t < 1 \end{cases}$$

and $\tilde{\psi}_2$ be defined by

$$\tilde{\psi}_2(t) = \begin{cases} t - 1/3 & if\ 1/3 < t < 1 \\ 0 & if\ 0 < t \le 1/3 \end{cases}$$

Note that $\tilde{\psi}_1 + \tilde{\psi}_2$ is nowhere vanishing, and set

$$\psi_1 = \frac{\tilde{\psi}_1}{\tilde{\psi}_1 + \tilde{\psi}_2}$$

$$\psi_2 = \frac{\tilde{\psi}_2}{\tilde{\psi}_1 + \tilde{\psi}_2}$$

We have $supp(\psi_1) = (0, 2/3] \subset U_1$, $supp(\psi_2) = [1/3, 1) \subset U_2$ and $\psi_1 + \psi_2 = 1$ so this is *almost* a partition of unity subordinate to the cover $\{U_1, U_2\}$. Unfortunately the supports of the ψ_i are *not* compact.

The trouble is that we have tried to manage with too few $\psi_i's$. Here is the remedy. First some notation. Given a real numbers $a < b$, let $T_{[a,b]}$ denote the continuous function (with graph an isosceles triangle of height $\frac{b-a}{2}$ and base $b - a$):

$$T_{[a,b]}(t) = \begin{cases} 0 & if\ t < a \\ 0 & if\ t > b \\ \frac{b-a}{2} - |t - \frac{b+a}{2}| & if\ a \le t \le b \end{cases}$$

Note that $supp(T_{[a,b]}) = [a, b] \subset (0, 1)$ provided $0 < a < b < 1$.

(1) For $n \ge 1$, define $\tilde{\psi}_n$ by

$$\tilde{\psi}_n = T_{[\frac{n}{n+1}, \frac{n+2}{n+3}]}$$

The supports of $\tilde{\psi}_1, \tilde{\psi}_2, \ldots, \tilde{\psi}_n, \tilde{\psi}_{n+1}, \ldots$ are

$$[\frac{1}{2}, \frac{3}{4}],\ [\frac{2}{3}, \frac{4}{5}],\ \ldots,\ [\frac{n}{n+1}, \frac{n+2}{n+3}],\ [\frac{n+1}{n+2}, \frac{n+3}{n+4}],\ \ldots$$

from which it follows (by examination) that if $t > 1/2$. at least one of the functions $\tilde{\psi}_n$'s is non-zero, and there is a neighbourhood on which all but finitely many are nonzero.

(2) (Reflect the above functions around the point $t = 1/2$.) For $n \geq 1$, define the functions $\tilde{\psi}_{-n}$ by

$$\tilde{\psi}_{-n} = T_{[\frac{1}{n+3}, \frac{1}{n+1}]}$$

The supports of $\tilde{\psi}_{-1}, \tilde{\psi}_{-2}, \ldots, \tilde{\psi}_{-n}, \tilde{\psi}_{-(n+1)}, \ldots$ are

$$[\frac{1}{4}, \frac{1}{2}],\ [\frac{1}{5}, \frac{1}{3}],\ \ldots,\ [\frac{1}{n+3}, \frac{1}{n+1}],\ [\frac{1}{n+4}, \frac{1}{n+2}],\ \ldots$$

from which it follows (by examination) that if $t < 1/2$. at least one of the functions $\tilde{\psi}_{-n}$'s is non-zero, and there is a neighbourhood on which all but finitely many are nonzero.

(3) Note that the functions listed so far all vanish at $t = \frac{1}{2}$. So we add a function $\tilde{\psi}_0$ defined by

$$\tilde{\psi}_0 = T_{[\frac{1}{3}, \frac{2}{3}]}$$

The (continuous) functions $\tilde{\psi}_n, n = \ldots, -2, -1, 0, 1, 2, \ldots$ have the following properties:

(1) $\tilde{\psi}_n(t) \geq 0, t \in (0, 1)$, and

(2) given any given $t \in (0, 1)$, there exists a neighbourhood $U_t \subset (0, 1)$ such that all but finitely many $\tilde{\psi}_n$ vanish on U_t, and at least one $\tilde{\psi}_n(t) \neq 0$.

(3) $supp(\tilde{\psi}_n) \subset U_2$ if $n \geq 0$ and $supp(\tilde{\psi}_n) \subset U_1$ if $n \leq 0$. All supports are compact.

From this it follows that the functions ψ_n, defined by

$$\psi_n(t) = \frac{\tilde{\psi}_n(t)}{\sum_m \tilde{\psi}_m(t)}$$

gives a partition of unity subordinate to the cover $\{U_1, U_2\}$. With a little more work, we could construct C^∞ partitions of unity.

Example 2: Let U be the real line $\mathbb{R}$. Consider the open cover consisting of the intervals $U_n = (n-1, n+1),\ n \in \mathbb{Z}$. Let $\tilde{\psi}_n, n \in \mathbb{Z}$ be defined by

$$\tilde{\psi}_n = T_{[n-\frac{3}{4}, n+\frac{3}{4}]}$$

Note that $supp(\tilde{\psi}_n)$ is compact and contained in U_n. As in the previous example, these functions can be used to generate a partition of unity subordinate to the open cover. Note that in this case we are able to manage with the same indexing set because the U_n are relatively compact.

8. Change of variables in integration: moving beyond rectangles using a partition of unity.

We will illustrate the use of a partition of unity by proving the general theorem.

Theorem 8.1. *Let $\tilde{\gamma} : \tilde{U} \to U$ be a C^1 diffeomorphism between open subsets $\tilde{U}, U$ in $\mathbb{R}^n$.*

If f is a continuous function with compact support on U,

$$\int_{\tilde{U}} \tilde{f}|J_{\tilde{\gamma}}| = \int_U f$$

Proof. Since f has compact support, the function $\tilde{f}$, and hence the function $\tilde{f}|J_{\tilde{\gamma}}|$ have compact support as well.

Choose a family of cubes $\tilde{R}_i \subset \tilde{U}$ as in the Theorem 7.3. (Note, however, that we are dealing with $\tilde{U}$ rather than U.) Let $\tilde{\psi}_i$ be a corresponding partition of unity. By compactness of *supp* $\tilde{f}$, only finitely many of the ψ_i are non-vanishing there. Let $I \subset \mathbb{N}$ denote the corresponding set of indices. We have

$$\begin{aligned}
\int \tilde{f}|J_{\tilde{\gamma}}| &= \int \sum_{i\in I} \tilde{\psi}_i \tilde{f}|J_{\tilde{\gamma}}| \\
&= \sum_{i\in I} \int \tilde{\psi}_i \tilde{f}|J_{\tilde{\gamma}}| \\
&= \sum_{i\in I} \int \psi_i f \\
&= \int \sum_{i\in I} \psi_i f \\
&= \int f
\end{aligned}$$

□

In the above, we have set $\psi_i \equiv \tilde{\psi}_i \circ \tilde{\gamma}^{-1}$. Note that $\sum_{i\in I} \psi = 1$ on *supp* f.

9. Integrable non-negative functions on open sets; absolutely integrable functions on open sets (to be covered later)

So far we have defined the Riemann integrability (and integrals when this holds) of bounded functions on closed rectangles in $\mathbb{R}^n$, and then extended this notion to bounded functions on Jordan-measurable subsets of $\mathbb{R}^n$. Bounded continuous functions on a Jordan-measurable subset are integrable.

Not every open set is Jordan-measurable. This could be because the set is not bounded – recall that a Jordan-measurable set is bounded by definition – but there exist examples of bounded open sets which are not Jordan-measurable. Here is an example[19].

Consider the rationals in the interval $(0,1)$, and enumerate them in some way. For $n \geq 1$, let A_n be an open interval centred on the nth rational such that $A_n \subset (0,1)$ and $\sum_{n=1}^{\infty} length\ (A_n) < \frac{1}{3}$. Let

$$A = \underset{n}{\cup} A_n$$

Note that A is open and its closure is the closed interval $I \equiv [0,1]$. Let $P = \{I_1, \ldots, I_k\}$ be any partition of I. Let us estimate the lower and upper Darboux sums of χ_A w.r.to P. We have

$$U(P, \chi_A) = \sum_{i=1}^{k} M_{I_i}(\chi_A)\ell(I_i) = \sum_i \ell(I_i) = 1$$

(where we use the fact that $\overline{A} = I$) and

$$L(P, \chi_A) = \sum_{i=1}^{k} m_{I_i}(\chi_A)\ell(I_i) = \sum_{i \text{ such that } I_i \subset A} \ell(I_i)$$

Suppose now (after relabelling, if necessary) that $I_1, \ldots, I_{k'}$ (for $k' \leq k$) are the intervals that are contained in A, and the rest are not. Then the union $I' = I_1 \cup \ldots I_{k'}$ is compact and covered by the family of open sets $\{A_n\}_n$, Then a finite cover exists, say by the open sets $A_{i_1}, \ldots, A_{i_N}$. Now it is a simple matter to check that

$$L(P, \chi_A) = \ell(I') \leq \sum_{l=1}^{N} length(A_{i_l}) < \frac{1}{3}$$

We can now conclude that χ_A is not Riemann integrable.

Since the open set A in the above example is bounded and χ_A is not Jordan-measurable, Lebesgue's theorem can be used to conclude that the boundary ∂A has nonzero content. It could also happen that an open set is Jordan-measurable for the simple reason that it is not bounded – an obvious example is $\mathbb{R}^n$.

[19] This is adapted from `https://www2.math.upenn.edu/ kazdan/509S07/jordan-meas.pdf`.

We now define the class of absolutely Riemann integrable functions on an arbitrary open set $U \subset \mathbb{R}^n$. In case U is bounded and Jordan-measurable, this will agree with the class of Riemann integrable functions defined earlier.

Definition 9.1. Let $U \subset \mathbb{R}^n$ be open.

(1) A real-valued function f on U is *locally integrable* if it is bounded and Riemann integrable on every compact and Jordan-measurable set $K \subset U$.

(2) A function $f : U \to [0, \infty]$ is *integrable* if it is locally integrable and

$$\sup\{ \int_K f(x)dx \Big| K \subset U,\ K \text{ compact and Jordan-measurable}\} < \infty$$

in which case we set

$$\int_U f(x)dx \underset{def}{=} \sup\{ \int_K f(x)dx \Big| K \subset U,\ K \text{ compact and Jordan-measurable}\}$$

(3) A real-valued function f on U is *absolutely integrable* if it is locally integrable and

$$\sup\{ \int_K |f(x)|dx \Big| K \subset U,\ K \text{ compact and Jordan-measurable}\} < \infty$$

in which case we set

$$\int_U f(x)dx \underset{def}{=} \int_U f_+(x)dx - \int_U f_-(x)dx$$

(Exercise: why does this make sense?)

9.1. **Bounded open sets.** We will not concern ourselves much with integration on bounded but non Jordan-measurable open sets. But here are two results to keep in mind. We omit the proofs.

Proposition 9.2. *If U is open, bounded and Jordan-measurable, a bounded Riemann-integrable function g on U is absolutely integrable, and the two notions of integral of g on U coincide. In other words,*

$$\int_R \tilde{g}(x)dx = \sup_K \int_K g_+(x)dx - \sup_K \int_K g_-(x_d x$$

On the left R is any closed rectangle containing U, and $\tilde{g}$ is the extension by zero of g to R. On the right $g = g_+ - g_-$ is the decomposition of g into its positive and negative parts, and the suprema over K run over compact, Jordan-measurable sets $K \subset U$.

Proposition 9.3. *If U is open and bounded, any bounded locally integrable function g on U is absolutely integrable. In particular this holds if g is continuous and bounded on U.*

9.2. **Exhaustions.** It is often useful to consider "exhaustions" by bounded sets. Keep in mind the following proposition, whose proof we again omit.

Proposition 9.4. *Suppose U is open (which need not be bounded) and $U_1 \subset U_2 \subset \ldots U_l \subset \ldots$ a nested sequence of open subsets of U such that*

(1) $\bigcup_l U_l = U$,

(2) each U_l is bounded (equivalently, the closure $\overline{U}_l$ is compact[20])

(3) each $\overline{U}_l \subset U_{l+1}$

(4) each $\overline{U}_l$ is Jordan-measurable.

If f is a locally integrable function on U, then it is absolutely integrable iff the increasing sequence of nonnegative numbers

$$\cdots \leq \int_{\overline{U}_l} |f(x)|dx \leq \int_{\overline{U}_{l+1}} |f(x)|dx \leq \ldots$$

is bounded above, in which case

$$\lim_{l\to\infty} \int_{\overline{U}_l} f(x)dx = \int_U f(x)dx$$

The following are adapted from Rudin's exercises 7, 8 and 9 of Chapter 6.

Exercise: Let $0 < \epsilon << 1$.

(1) If $f : (0,1) \to \mathbb{R}$ is absolutely integrable on $(0,1)$, show that

$$\int_{(0,1)} |f(x)|dx = \lim_{\epsilon\to 0} \int_{[\epsilon,1-\epsilon]} |f(x)|dx$$

and

$$\int_{(0,1)} f(x)dx = \lim_{\epsilon\to 0} \int_{[\epsilon,1-\epsilon]} f(x)dx$$

(2) If f is the restriction of an integrable function on $[0,1]$, show that f is absolutely integrable on $(0,1)$, and

$$\int_{(0,1)} f(x)dx = \int_{[0,1]} f(x)dx$$

(3) Give an example of an unbounded function $f : (0,1) \to [0,\infty)$ which is integrable on $(0,1)$.

Exercise: Let $f : [1,\infty) \to [0,\infty)$ be monotonically decreasing. Prove that f is locally integrable. Prove that f is integrable iff

$$\sum_{l=1}^{\infty} f(\alpha n) < \infty \text{for some } \alpha > 0 \quad .$$

[20]In this case we say that U_l is *relatively compact.*

Exercise: Consider the function f defined on $(0,\infty)$:

$$f(x) = \frac{\cos x}{1+x}$$

Use the last exercise to show that this is *not* absolutely integrable. Use integration by parts to show that on the other hand

$$\lim_{\epsilon\to 0}\int_{\epsilon}^{1/\epsilon} f(x)dx$$

exists. Note that

$$\frac{\cos x}{1+x} - \frac{\sin x}{(1+x)^2} = (\frac{\sin x}{1+x})'$$

9.3. **Equivalent definition using partition of unity.** Let U be an open subset of $\mathbb{R}^n$. By Theorem 7.3, there exist a countable family of non-negative continuous functions ψ_i with compact support such that locally finitely many are non-zero, and such that $\sum_i \psi_i(x) = 1,\ x \in U$. Following Spivak, we say that a real-valued function on U is absolutely integrable, if

(1) $\psi_i f$ is integrable for each i, and

(2) $\sum_i \int \psi_i |f| < \infty$

In which case, both f_+ and f_- are absolutely integrable and we define

$$\int f = \sum_i \int \psi_i f_+ - \sum_i \int \psi_i f_-$$

One needs to show that the definition is (a) notion is independent of the partition of unity chosen, and (b) agrees with the earlier definition.

Note that Spivak strangely omits the requirement that the ψ_i have compact support. This is a well-known error in the book. See: `https://www.jirka.org/spivak-errata.html`.

10. Integration over all of $\mathbb{R}^n$: an example (to be covered later)

As far as non-negative functions are concerned, the theory is straightforward. Given such a function f, we say that f is (Riemann) integrable if it is Riemann integrable on each rectangle $R \subset \mathbb{R}^n$ and

$$\sup_{R\subset\mathbb{R}^n}\int_R f(\vec{x})d\vec{x} < \infty$$

in which case we set

$$\int_{\mathbb{R}^n} f(\vec{x})d\vec{x} = \sup_{R\subset\mathbb{R}^n}\int_R f(\vec{x})d\vec{x}$$

Let R_L denote the rectangle $[-L,L]^n$, and B_t the closed ball of radius t centred at the origin.

Exercise: Prove that a non-negative function f is integrable iff it is integrable on all R_L and

$$\lim_{L\to\infty}\int_{R_L} f < \infty$$

iff it is integrable on all B_t and

$$\lim_{t\to\infty}\int_{B_t} f < \infty$$

in which case either of these limits equals $\int_{\mathbb{R}^n} f$.

Let us prove that the function $G(x) = \exp -x^2$ is integrable on $\mathbb{R}$. (Recall that in one dimension we use the alternate notations $\int_a^b = \int_{[a,b]}$.) We start with the basic estimate: if $1 \le a \le b$

$$\int_{[a,b]} G(x)dx \le \int_{[a,b]} xG(x)dx = \frac{1}{2}\int_{[a,b]}(-G)'(x)dx = \frac{1}{2}(e^{-a^2} - e^{-b^2})$$

For $L \ge 0$, set $I_1(L) = \int_{[-L,L]} G(x)dx$. If $1 \le L \le M$ we have

$$I_1(M) \le I_1(L) + e^{-L^2} - e^{-M^2} \le I_1(L) + e^{-L^2}$$

In particular, taking $L = 1$, we see that for $M \ge 1$ (and therefore for all M since $I_1(M)$ is an increasing function of M),

$$I_1(M) \le I_1(1) + e^{-1}$$

from which it follows (fill in the details) that G is integrable and

$$I_1 \equiv \int_{\mathbb{R}} e^{-x^2}dx = \lim_{M\to\infty} I_1(M)$$

We will now evaluate I_1.

Consider the function $G_2(x, y) = \exp -x^2 - y^2$ on $\mathbb{R}^2$, For $L \ge 0$, set

$$I_2(L) = \int_{R_L} G_2(x, y)dxdy$$

where R_L is the square $[-L, L] \times [L, L] \subset \mathbb{R}^2$. Applying Fubini, we see that

$$I_2(L) = I_1(L)^2$$

If $1 \le L \le M$ we have

$$I_2(M) - I_2(L) = (I_1(M) - I_1(L))(I_1(M) + I_1(L)) \le 2I_1 e^{-L^2}$$

from which it follows that G_2 is integrable on $\mathbb{R}^2$ and

$$I_2 = \lim_{L\to\infty} I_2(L) = lim_{L\to\infty} I_1(L)^2 = I_1^2$$

Let D_T denote the closed disc of radius T centred on the origin. We have

$$\mathcal{I}_T \equiv \int_{D_T} e^{-x^2-y^2}dxdy \underset{T\to\infty}{\nearrow} I_2$$

Consider the map $\mathbb{R}_{\ge 0} \times \mathbb{R} \to \mathbb{R}^2$ given by

$$(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$$

This is continuous and onto, and the preimage of D_T is $\tilde{D}_T \equiv \{0 \le r \le T\} \times \mathbb{R}$. Given a function f on $\mathbb{R}^2$, denote by $\tilde{f}$ the function defined by

$$\tilde{f}(r,\theta) = f(r\cos\theta, r\sin\theta)$$

Note that if f is continuous, so is $\tilde{f}$

Claim: given a continuous function f on $\mathbb{R}^2$,

$$\int_{D_T} f(x,y)dxdy = \int_{\tilde{R}_T} \tilde{f}(r,\theta) r d\theta dr$$

where $\tilde{R}_T$ is the rectangle $\{(r,\theta) | 0 \le r \le T,\ 0 \le \theta \le 2\pi\}$.

Assuming the Claim, we can evaluate $\mathcal{I}_T$:

$$\begin{aligned}
\mathcal{I}_T &\equiv \int_{D_T} e^{-x^2-y^2} dxdy \\
&\underset{Claim}{=} \int_{\tilde{R}_T} e^{-r^2} r d\theta dr \\
&\underset{Fubini}{=} \int_0^T \left(\int_0^{2\pi} e^{-r^2} r d\theta \right) dr \\
&= 2\pi \int_0^T e^{-r^2} r dr \\
&= 2\pi \int_0^T (-e^{-r^2})' dr \\
&= \pi \{1 - e^{-R^2}\}
\end{aligned}$$

It follows that $I_2 = \pi$ and $I_1 = \sqrt{\pi}$.

It remains to justify Claim, using the change of variables theorem. The trouble is that our theorem refers to C^1 diffeomorphisms of open sets, and we have a map of closed sets.

Consider first the map of *open*[21] sets $\mathbb{R}_{>0} \times \mathbb{R} \to \mathbb{R}^2 \setminus \{(0,0)\}$ given by

$$(r,\theta) \mapsto (x(r,\theta) = r\cos\theta, y(r,\theta) = r\sin\theta)$$

The derivative at (r,θ) is given by the matrix

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

so that the Jacobian determinant at the point (r,θ) is r. The map is onto, and a local diffeomorphism but not injective since the points $(r, \theta + 2\pi n)$, with $n \in \mathbb{Z}$ all map to the same point $(r\cos\theta, r\sin\theta)$.

Let $U \subset D_T$ be defined by

$$U = \{(x,y) | 0 < x^2 + y^2 < T^2\} \setminus \{(x,0) | 0 < x < T\}$$

[21]Note that we are now restricting to strictly positive values of r.

Let $V = \{(r.\theta)|0 < r < T,\ 0 < \theta < 2\pi\} \subset \tilde{R}_T$. (Thus V is the interior of the rectangle $\tilde{R}_T$.) The map $\vec{\gamma} : V \to U$ given by

$$\vec{\gamma}(r, \theta) = (r\cos\theta, r\sin\theta)$$

is a C^1 diffeomorphism and so the Change of Variables formula holds, and yields: *provided f has compact support and $supp(f) \subset U$*,

$$\int_{D_T} f(x, y)dxdy = \int_{\tilde{R}_T} \tilde{f}(r, \theta)rd\theta dr$$

Suppose given a family of continuous functions ϕ_ϵ on $\mathbb{R}^2$, parametrised by $0 < \epsilon << 1$ such that

(1) ϕ_ϵ has compact support and $supp(\phi_\epsilon) \subset U$,

(2) $0 \leq \phi_\epsilon(x, y) \leq 1,\ (x, y) \in \mathbb{R}^2$,

(3) $\phi_\epsilon(x, y) = 1$ if $\epsilon < \sqrt{x^2 + y^2} < T - \epsilon$ and $\frac{x}{\sqrt{x^2+y^2}} < 1 - \epsilon$.

Then, given any f that is continuous on D_T, we have

$$\int_{D_T} \phi_\epsilon(x, y)f(x, y)dxdy = \int_{\tilde{R}_T} \tilde{\phi}_\epsilon(r, \theta)\tilde{f}(r, \theta)rd\theta dr$$

Exercise: (a) Construct such a family of functions ϕ_ϵ. (b) Prove that as $\epsilon \to 0$, we recover the equality of the Claim.

11. Exterior algebra: supplementary remarks to Spivak

Let V be a real vector space of dimension n. We let V^* denote the space of linear maps $\phi : V \to \mathbb{R}$. (These are sometimes called *linear forms "on" V*, and V^* is called the dual vector space of V.).

Exercise If $\{e_1, \ldots, e_n\}$ is a basis of V, the linear forms $\phi_i,\ i = 1, \ldots, n$, defined by

$$\phi_i(\sum_j a_j e_j) = a_i$$

(where $a_j \in \mathbb{R}$) form a basis (the dual basis) of V^*,

11.1. **The tensor algebra.** We set[22] $\mathcal{T}^0(V) = \mathbb{R}$, $\mathcal{T}^1(V) = V^*$, and, for $k = 2, \ldots,$,

$$\mathcal{T}^k(V) = \text{set of multilinear maps } T : \underbrace{V \times \cdots \times V}_{k\ factors} \to \mathbb{R}$$

This is a vector space. If $e_1, \ldots, e_n$ is a basis of V, it is clear by multilinearity that a k-form T is determined by its values

$$T(e_{i_1}, \ldots, e_{i_k})$$

[22]This is Spivak's notational original sin, because usually what we denote $\mathcal{T}^k(V)$ is denoted $\mathcal{T}^k(V^*)$. With the usual convention $\mathcal{T}^1(V^*) = V^*$, etc. See Spivak's Addendum 4.

for $1 \le i_1 \le n, \ \dots, \ 1 \le i_k \le n$, and these values can be independently specified. We conclude that $dim\ \mathcal{T}^k(V) = n^k$.

We set

$$\mathcal{T}^*(V) = \oplus_k \mathcal{T}^k(V) \quad (\text{"algebraic direct sum of vector spaces"})$$

An element of the above direct sum is a sequence $(T_0, T_1, \dots, T_k \dots,)$ with each T_k a k-tensor, and all but finitely many zero. We next make this into a ring with the product defined as follows. First, elements of $\mathcal{T}^0(V) = \mathbb{R}$ multiply k-tensors as scalars. Second, given $S \in T^k(V)$ and $T \in T^l(V)$, we set

$$S \otimes T(v_1, \dots, v_k, v_{k+1}, \dots, v_{k+l}) = S(v_1, \dots, v_k)T(v_{k+1}, \dots, v_{k+l})$$

This is a bilinear map $\mathcal{T}^k(V) \times \mathcal{T}^l(V) \to \mathcal{T}^{k+l}(V)$, and extends in a natural way to give a noncommutative, but associative product that makes $\mathcal{T}^*(V)$ a $\mathbb{R}$-algebra. (That is to say, $\mathbb{R} = \mathcal{T}^0(V)$ is a subring.)

Because of associativity, we need not distinguish between $S \otimes (T \otimes U)$ and $(S \otimes T) \otimes U$ and will denote this simply $S \otimes T \otimes U$, and deal similarly with all products.

Theorem 4-1 of Spivak says: The k-tensors $\phi_{i_1} \otimes \cdots \otimes \phi_{i_k}$, with each index i_j running over $1, 2, \dots, n$, form a basis for $\mathcal{T}^k(V)$, which therefore has dimension n^k, as we already saw.

Given a linear map of vector spaces $f : V \to W$, you know already the *transpose* $f^* : W^* \to V^*$, defined by

$$f^*\rho(v) = \rho(f(v))$$

This extends to a linear map

$$f^* : \mathcal{T}^*W \to \mathcal{T}^*V$$

preserving degrees and tensor products. That is, $f^*T \in \mathcal{T}^k(V)$ if $T \in \mathcal{T}^k(W)$, and

$$f^*S \otimes f^*T = f^*(S \otimes T)$$

11.2. **The exterior algebra.** We define a subspace $\Lambda^k V$ of $\mathcal{T}^k(V)$ (the space of *alternating/"skew-symmetric" k-forms*) as follows: $\Lambda^0 V = \mathbb{R}$, $\Lambda^1 V = \mathcal{T}^1(V) = V^*$, and for $k \ge 2$:

$$\Lambda^k V = \{T \in \mathcal{T}^k(V) | T(v_1, \dots, v_l, \dots, v_m, \dots, v_k) = -T(v_1, \dots, v_m, \dots, v_l, \dots, v_k)\}$$

For example, $\Lambda^2 V$ is the space of skew-symmetric bilinear forms:

$$\Lambda^2 V = \{T \in \mathcal{T}^2(V) | T(v_1, v_2) = -T(v_2, v_1)\}$$

If $e_1, \dots, e_n$ is a basis of V, it is clear by multilinearity and skew-symmetry that an alternating k-form T is determined by its values

$$T(e_{i_1}, \dots, e_{i_k})$$

for $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, and these values can be independently specified. We conclude that

(1) $\Lambda^k V = \{0\}$ if $k > n$, and

(2) $dim\ \Lambda^k V = \binom{n}{k}$.

We next define linear maps

$$Alt^k : \mathcal{T}^k(V) \to \mathcal{T}^k(V)$$

with the properties:

(1) $Alt^k(T) \in \Lambda^k V$ for T any k-tensor, i.e., $Alt^k(\mathcal{T}^k(V)) \subset \Lambda^k V \subset \mathcal{T}^k(V)$, and

(2) $Alt^k(\omega) = \omega$ if $\omega \in \Lambda^k V$; in particular $Alt^k(\Lambda^k(V)) = \Lambda^k V$.

Here is the definition of Alt^k, check that this works: Alt^0 and Alt^1 are identity maps, and for $k \geq 2$:

$$Alt^k(T)(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} sgn(\sigma) T(v_{\sigma[1]}, \ldots, v_{\sigma[k]})$$

For example,

$$Alt^2(T)(v_1, v_2) = \frac{1}{2}(T(v_1, v_2) - T(v_2, v_1))$$

For future use, note that the kernel of the map Alt^2 is the space of symmetric bilinear forms

$$S^2 V \equiv \{T \in \mathcal{T}^2(V) | T(v_1, v_2) = T(v_2, v_1)\}$$

Putting together the subspaces $\Lambda^k V$ we get the *exterior algebra*

$$\oplus_k \Lambda^k(V) \equiv \Lambda^* V \subset \mathcal{T}^*(V) \equiv \oplus_k \mathcal{T}^k(V)$$

Note that the direct sum on the left can be taken over the range $0 \leq k \leq n$.

The term *exterior algebra* is premature, because we have not defined a product on $\Lambda^* V$. We cannot simply induce a product from the tensor product. For example, the tensor product of two linear forms ϕ and ψ, which belong to $\Lambda^1 V = \mathcal{T}^1(V) = V^*$ is not an alternating bilinear form since

$$\phi \otimes \psi(v_1, v_2) = \phi(v_1)\psi(v_2) \neq -\phi(v_2)\psi(v_1) = -\phi \otimes \psi(v_2, v_1)$$

in general. Instead, we define a product (the *"exterior/wedge product"*) of alternating forms $\omega \in \Lambda^k V$ and $\eta \in \Lambda^l V$ by

$$\omega \wedge \eta = \frac{(k+l)!}{k!l!} Alt^{k+l}(\omega \otimes \eta)$$

(In particular, for linear forms ϕ, ψ this gives $\phi \wedge \psi = 2Alt^2(\phi \otimes \psi) = \phi \otimes \psi - \psi \otimes \phi$.) Note that

$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$$

That the exterior product gives a bilinear map $\Lambda^k V \times \Lambda^l V \to \Lambda^{k+l} V$ is clear. It is also clear that given a linear map $f : V \to W$ of vector spaces, the induced maps $\Lambda^k W \to \Lambda^k V$ are compatible with exterior products just as with tensor products. What is less clear is that the exterior product is associative. This is the content of Theorem 4-4(3) of Spivak.

11.3. **The exterior algebra as quotient of the tensor algebra: optional, "not in syllabus".** Consider the (clearly linear and surjective) map $Q : \mathcal{T}^*(V) \to \Lambda^* V$, given by

$$Alt(S_0, S_1, \dots, S_k, \dots) = (S_0, S_1, \dots, Alt^k(S_k), \dots)$$

(I omit Alt in the first two components since in degree 0 and 1, $\Lambda^k V = \mathcal{T}^k(V)$.) Theorem 4-4(1) says that the kernel of Q is a two-sided ideal in $\mathcal{T}^*(V)$. (In fact this is the two-sided ideal generated by $S^2 V \subset \mathcal{T}^2(V)$.) From this it follows that $\Lambda^* V$, with the product

$$\omega \wedge' \eta = Alt(\omega \otimes \eta)$$

is an (associative) algebra. The wedge product as defined earlier defines an isomorphic algebra, with the isomorphism being given by

$$(\Lambda^* V, \wedge') \longrightarrow (\Lambda^* V, \wedge)$$
$$\Lambda^k V \ni \omega \longmapsto k!\omega \in \Lambda^k V$$

so that

$$\omega \wedge' \eta = Alt(\omega \otimes \eta) \longmapsto (k+l)! Alt(\omega \otimes \eta) = \frac{(k+l)!}{k!l!} Alt(k!\omega \otimes l!\eta) = k!\omega \wedge l!\eta$$

All this is the content of Theorem 4-4 of Spivak. This reformulation shows that the key step is the first. the wedge product $\wedge'$ is more natural from the algebraic point of view, but $\wedge$ is used in geometry.

11.4. **Determinants.** Recall the definition of the determinant of a linear transformation $A : V \to V$, with V an n-dimensional vector space:

$det\ A$ = determinant of the matrix representing A w.r.to any basis of V

Define the linear map $\mathcal{D}(A) : \Lambda^n V \to \Lambda^n V$ by

$$\mathcal{D}(A)[\omega](v_1, \dots, v_n) = \omega(Av_1, \dots, Av_n)$$

Proposition 11.1. $\mathcal{D}(A)[\omega] = (det\ A)\omega$

Proof. Note that $\mathcal{D}$ (like det) is multiplicative: $\mathcal{D}(AB) = \mathcal{D}(A)\mathcal{D}(B)$, so it suffices to check that the two functions agree on a (multiplicative) generating set for the space of linear maps $V \to V$. Choose a basis and use Proposition 2.18 of Artin's *Algebra*, Prentice-Hall (1991). □

11.5. **Orientation/Oriented basis.** Let V be an n-dimensional real vector space. Since $\Lambda^n V$ is a one-dimensional real vector space, removing the zero vector leaves two connected components, and an orientation on V is the choice of one of these components.

On the other hand, consider the set of bases of V. Any two bases are related by a unique invertible $n \times n$ real matrix, and we can divide the set of bases into two equivalence classes, declaring that two bases are equivalent if they are related by an invertible matrix with positive determinant. Choosing one of the equivalence classes amounts to choosing an orientation, the map being given by

$$\text{equivalence class of } \{e_1, \ldots, e_n\} \mapsto \text{equivalence class of } \phi_1 \wedge \cdots \wedge \phi_n$$

12. Interlude: an example of Stokes' Theorem in action

Let $U \subset \mathbb{R}^2$ be a non-empty, connected open set. A *vector field on U is a map* $\vec{v} : U \to \mathbb{R}^2$. Think of a fluid flowing around in U, and $\vec{v}$ as its "velocity field". We let (x, y) denote the coordinates in $\mathbb{R}^2$ and write $\vec{v}(x, y) = (v_x(x, y), v_y(x, y))$, so that v_x, v_y are real-valued functions on U. We will suppose that $\vec{v}$ is C^1; equivalently v_x, v_y are differentiable, and their partial derivatives are continuous.

The *curl* of $\vec{v}$ is the function[23] defined by

$$curl(\vec{v}) = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}$$

(There is *another* function we can define, the *divergence* of $\vec{v}$, defined by

$$div(\vec{v}) = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}$$

This is not relevant as yet.)

By "a (parametrised) path in U from A to B", we mean a map $\vec{\gamma} : [a, b] \to U$ for some closed interval $[a, b]$ such that $\vec{\gamma}(a) = A$ and $\vec{\gamma}(b) = B$. We will consider C^1 paths, i.e., we will assume that $\vec{\gamma}$ extends to an open interval $(a - \epsilon, b + \epsilon) \supset [a, b]$ as a C^1-map to U. By the line-integral of $\vec{v}$ along $\vec{\gamma}$ we mean the Riemann integral

$$\int_{\vec{\gamma}} \vec{v} \cdot d\vec{\gamma} \underset{def}{\equiv} \int_a^b \{v_x(\vec{\gamma}(s)) \frac{d\vec{\gamma}_x}{ds}(s) + v_y(\vec{\gamma}(s)) \frac{d\vec{\gamma}_y}{ds}(s)\} ds$$

Note the important fact that the integral is "invariant under orientation-preserving reparametrisation". That is, if $r : [c, d] \to [a, b]$ is a C^1-diffeomorphism

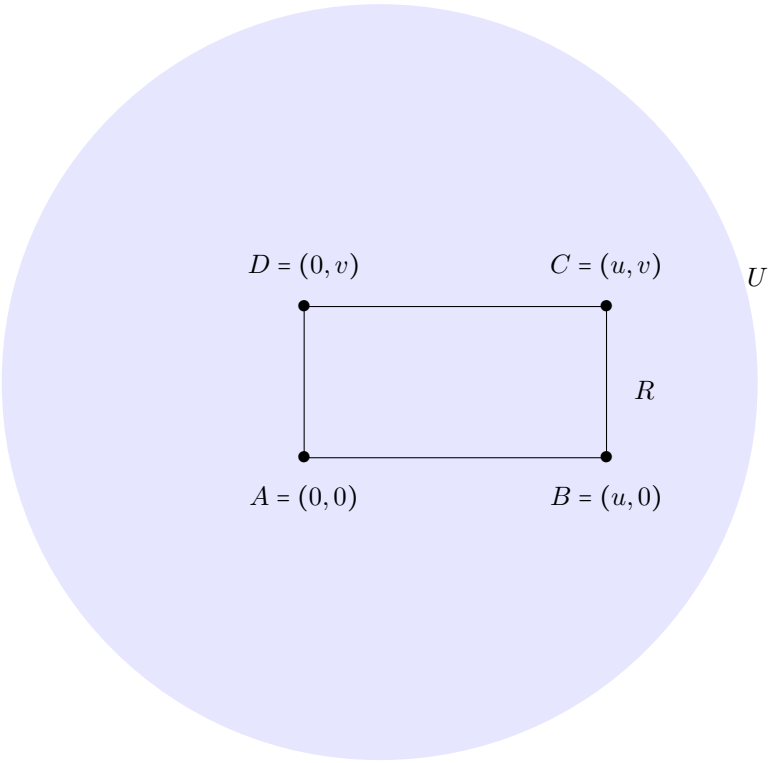
[23] Warning: in three dimensions analogue of this would be a vector field.

such that $r(c) = a,\ r(d) = b$ and $\vec{\gamma}' = \vec{\gamma} \circ r$,

$$\begin{aligned}\int_{\vec{\gamma}'} \vec{v} \cdot d\vec{\gamma}' &= \int_c^d \{v_x(\vec{\gamma}'(t))\frac{d\vec{\gamma}'_x}{dt}(t) + v_y(\vec{\gamma}'(t))\frac{d\vec{\gamma}'_y}{dt}(t)\}dt \\ &= \int_c^d \{v_x(\vec{\gamma}(r(t))\frac{d\vec{\gamma}_x}{ds}(r(t)) + v_y(\vec{\gamma}(r(t)))\frac{d\vec{\gamma}_y}{ds}(r(t))\}\frac{dr(t)}{dt}dt \\ &\underset{\#}{=} \int_a^b \{v_x(\vec{\gamma}(s))\frac{d\vec{\gamma}_x}{ds}(s) + v_y(\vec{\gamma}(s))\frac{d\vec{\gamma}_y}{ds}(s)\}ds \\ &= \int_{\vec{\gamma}} \vec{v} \cdot d\vec{\gamma}\end{aligned}$$

where the equality # is a 1-dimensional version of change-of-variables.

12.1. **A rectangle.** Let $R \subset U$ be the rectangle with "corners" A, B, C, D as follows:



Theorem 12.1. (Special case of Stokes' Theorem) *Let $\vec{v}$ be a C^1 vector field on U. Then*

$$\begin{aligned}\int_R curl(\vec{v}) = &\int_{\vec{\gamma}_{AB}} \vec{v} \cdot d\vec{\gamma}_{AB} + \int_{\vec{\gamma}_{BC}} \vec{v} \cdot d\vec{\gamma}_{BC} \\ &+ \int_{\vec{\gamma}_{CD}} \vec{v} \cdot d\vec{\gamma}_{CD} + \int_{\vec{\gamma}_{DA}} \vec{v} \cdot d\vec{\gamma}_{DA}\end{aligned}$$

where $\vec{\gamma}_{AB}$ is any C^1 parametrisation of the line segment from A to B, etc.

Proof. The proof is a simple calculation using Fubini. First,

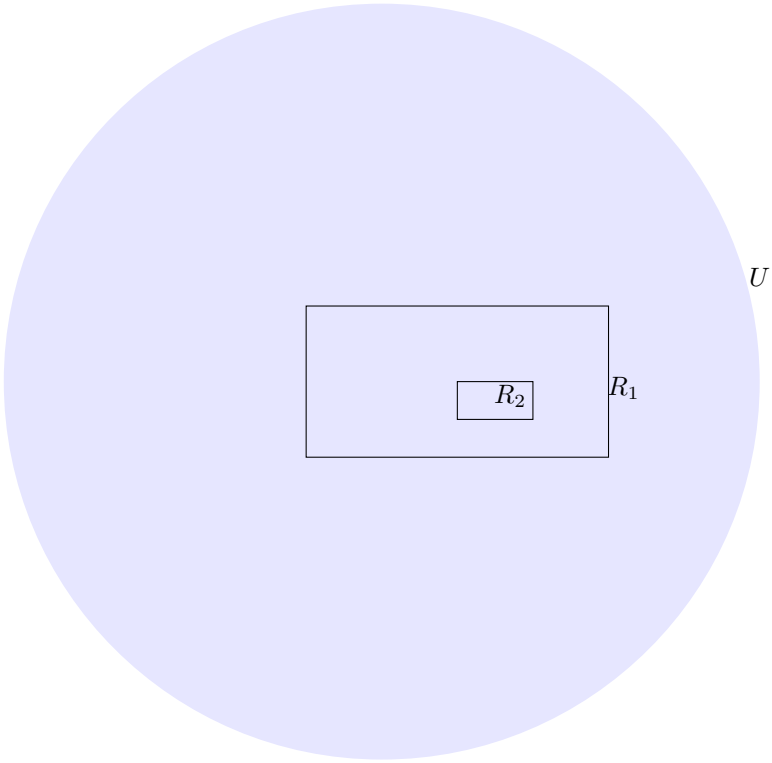
$$\begin{aligned}\int_R curl(\vec{v}) &= \int_0^v \{\int_0^u \frac{\partial v_y}{\partial x}(x,y)dx\}dy - \int_0^u \{\int_0^v \frac{\partial v_x}{\partial y}(x,y)dy\}dx \\ &= \int_0^v \{v_y(u,y) - v_y(0,y)\}dy - \int_0^u \{\vec{v}_x(x,v) - \vec{v}_x(x,0)\}dx\end{aligned}$$

Now it is a simple matter to identify each the four terms on the right with a corresponding line integral. For example, let $\vec{\gamma}_{AB}(s) = (s,0)$, so that $(\vec{\gamma}_{AB})_x(s) = s$ and $(\vec{\gamma}_{AB})_y(s) = 0$. Then

$$\int_{\vec{\gamma}_{AB}} \vec{v} \cdot d\vec{\gamma}_{AB} = \int_0^u \{\vec{v}_x(s,0) \times 1 + \vec{v}_y(x,0) \times 0\}ds = \int_0^u \vec{v}_x(x,0)dx$$

□

12.2. **A rectangle minus a smaller rectangle.** Let $R_1 \subset U$ be a rectangle as before, and R_2 the smaller rectangle:

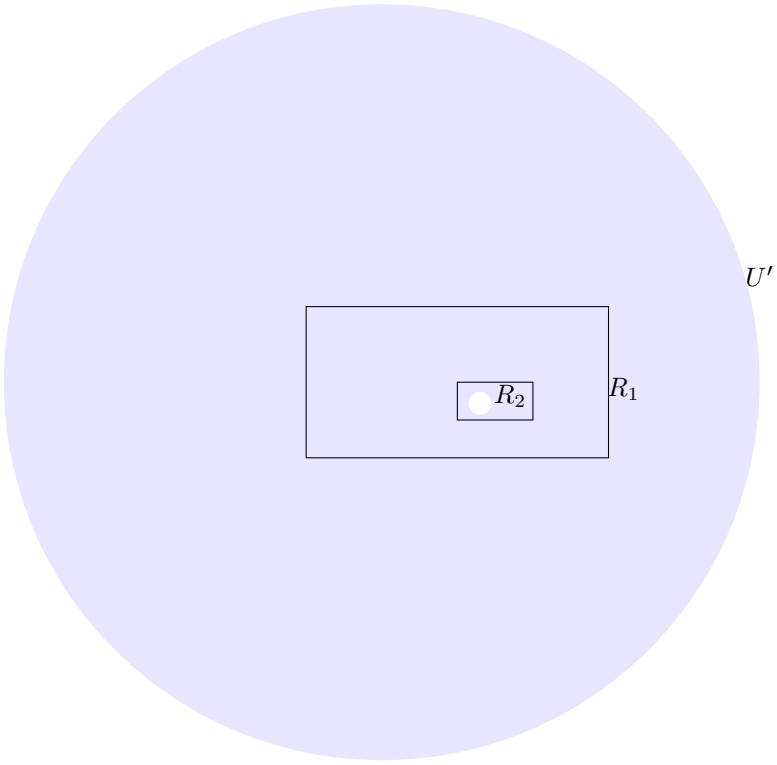


Let $\vec{v}$ be a C^1 vector field on U as before. Let $S = R_1 \setminus R_2^0$, where R_2^0 is the interior of R_2. The set S is closed, and Jordan-measurable. Since $R_1 = R_2 \cup S$ and $R_2 \cap S$ has zero content (recall our football field),

$$\int_S curl(\vec{v}) = \int_{R_1} curl(\vec{v}) - \int_{R_2} curl(\vec{v}) = L_1 - L_2$$

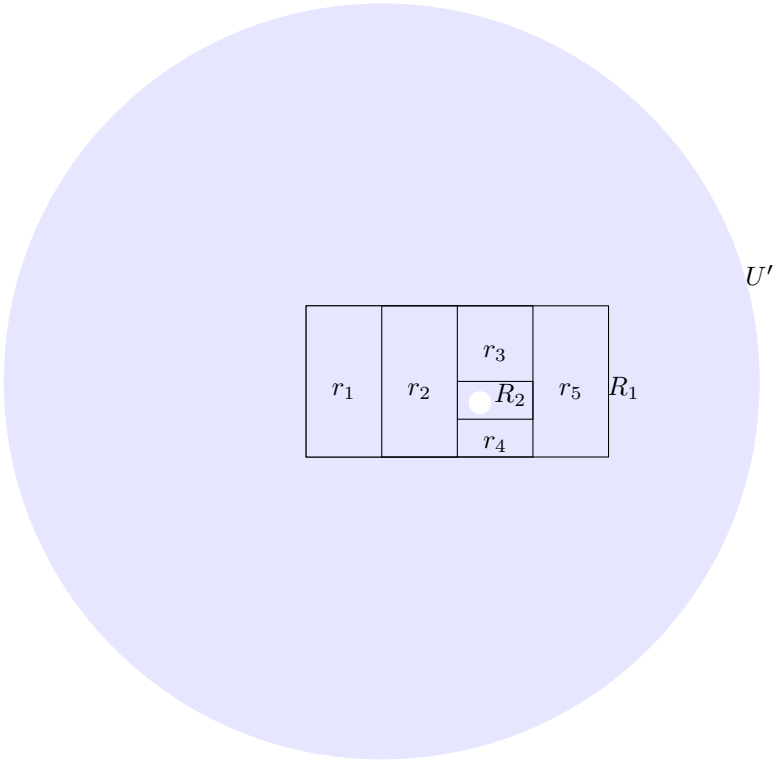
where L_1 and L_2 are the "anticlockwise line integrals of $\vec{v}$ around the boundaries of R_1 and R_2 respectively".

12.3. **As above, but $\vec{v}$ is not defined everywhere inside R_2.** Suppose $U' = U \setminus$ a closed disc inside R_2:



Let $\vec{v}$ be a C^1 vector field on U' as before. Let $S = R_1 \setminus R_2^0$, where R_2^0 is the interior of R_2.

Consider rectangles r_1, r_2, r_3, r_4, r_5 (we could do with fewer) as in the figure:



Then, using the football-ground argument,

$$\begin{aligned}\int_S curl(\vec{v}) &= \sum_i \int_{r_i} curl(\vec{v}) \ (football - ground\ argument) \\ &= \sum_i l_i \ \ (Stokes\ on\ r_i)\end{aligned}$$

where l_i is the line-integral around the boundary of r_i.

Important Exercise: Check that the line-integrals on the common boundary segments cancel out so that again,

$$\int_S curl(\vec{v}) = L_1 - L_2$$

So the "hole" in U does not affect our answer. Note that, however, we have no expression now for the integral of curl $\vec{v}$ over R_1.

12.4. **The gradient of a function is curl-free.** Let $U \subset \mathbb{R}^2$ be a non-empty, connected open set. Given a differentiable function f, we can define its gradient, which is the vector field

$$grad\ f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$$

For example, consider on $\mathbb{R}^2$ the function $2f(x,y) = x^2 + y^2$. Then $grad\ f$ is the "radial vector field":

$$grad\ f = (x, y)$$

If $U = \mathbb{R}^2 \setminus \{(0,0)\}$ and $f(x,y) = \log\sqrt{x^2+y^2}$,

$$grad\ f = (\frac{x}{r^2}, \frac{y}{r^2})$$

"A gradient vector-field is curl-free":

$$curl\ grad\ f = \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} = 0$$

On the other hand, consider the vector field on $U = \mathbb{R}^2 \setminus \{(0,0)\}$:

$$\vec{v}(x,y) = (-\frac{y}{r^2}, \frac{x}{r^2})$$

This is curl-free since

$$curl\ \vec{v} = \frac{1}{r^2} - \frac{2x^2}{r^4} + \frac{1}{r^2} - \frac{2y^2}{r^4} = 0$$

To see that it is not the gradient of a vector field, consider the closed path

$$\vec{\gamma} : [0,1] \to U,\ s \mapsto \vec{\gamma}(s) = (\cos 2\pi s, \sin 2\pi s)$$

By "closed" we mean that $\vec{\gamma}(0) = \vec{\gamma}(1) = (1,0)$. Computing the line-integral of $\vec{v}$ "along" (or "around") the path, we find

$$\int_{\vec{\gamma}} \vec{v}.d\vec{\gamma} \underset{check}{=} \int_0^1 2\pi ds = 2\pi$$

That this is not a "gradient vector field" follows from:

Lemma 12.2. *If $\vec{v} = grad\ f$, then for any path $\vec{\gamma} : [a,b] \to U$ from point $A = \vec{\gamma}(a)$ to point $B = \vec{\gamma}(b)$, we have*

$$\int_{\vec{\gamma}} \vec{v}.d\vec{\gamma} = f(B) - f(A)$$

In particular, if $\vec{\gamma}$ is closed (i.e. $A = B$), we have

$$\int_{\vec{\gamma}} \vec{v}.d\vec{\gamma} = 0$$

Proof. We use the definition of line integral.

$$\begin{aligned}
\int_{\vec{\gamma}} \vec{v}.d\vec{\gamma} \underset{check}{=} \int_a^b ds &= \int_a^b \{v_x(\vec{\gamma}(s))\frac{d\vec{\gamma}_x}{ds}(s) + v_y(\vec{\gamma}(s))\frac{d\vec{\gamma}_y}{ds}(s)\}ds \\
&= \int_a^b \{\frac{\partial f}{\partial x}(\vec{\gamma}(s))\frac{d\vec{\gamma}_x}{ds}(s) + \frac{\partial f}{\partial y}(\vec{\gamma}(s))\frac{d\vec{\gamma}_y}{ds}(s)\}ds \\
&\underset{chain\ rule}{=} \int_a^b \frac{d\{f \circ \vec{\gamma}\}}{ds}(s)ds \\
&= f(\vec{\gamma}(b)) - f(\vec{\gamma}(a)) = f(B) - f(A)
\end{aligned}$$

□

12.5. **"Line integrals along two paths are equal if one can be deformed to the other".** More precisely,

Theorem 12.3. *Let $U \subset \mathbb{R}^2$ be a connected domain, and A, B two (not necessarily distinct) points in U and $\vec{\gamma}_0$ and $\vec{\gamma}_1$ two C^∞ paths (parametrised by $[0,1]$) in U from A to B. Suppose there is a C^∞ map $\Gamma : R \equiv [0,1] \times [0,1] \to U$ such that*

(1) $\Gamma(s,0) = \vec{\gamma}_0(s)$,

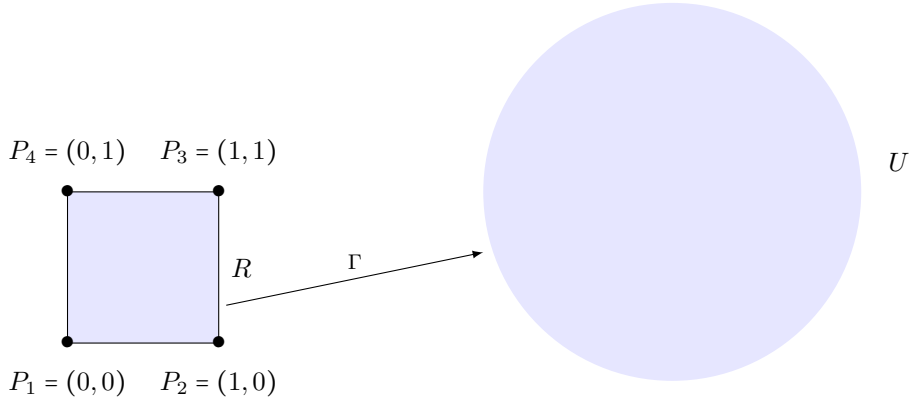
(2) $\Gamma(s,1) = \vec{\gamma}_1(s)$,

(3) $\Gamma(0,t) = A$,

(4) $\Gamma(1,t) = B$

(Here (s,t) are the obvious coordinate variables on the square R.) Then, for any curl-free vector field $\vec{v} = (v_x, v_y)$ on U, we have

$$\int_{\vec{\gamma}_0} \vec{v} \cdot d\vec{\gamma}_0 = \int_{\vec{\gamma}_1} \vec{v} \cdot d\vec{\gamma}_1$$



Proof. Consider the following vector field $\vec{v}'$ on R:

$$\vec{v}' \equiv (v'_s, v'_t) = (v_x \frac{\partial x}{\partial s} + v_y \frac{\partial y}{\partial s}, v_x \frac{\partial x}{\partial t} + v_y \frac{\partial y}{\partial t})$$

where on the right we have abused notation so that

$$\frac{\partial x}{\partial s}(s,t) \text{ means } \frac{\partial \Gamma_x}{\partial s}(s,t), \quad v_x(s,t) \text{ means } v_x(\Gamma(s,t)), \text{ etc.}$$

Let us compute $curl\ \vec{v}'$ without assuming that $curl\vec{v} = 0$. This is an exercise in the chain rule:

$$\begin{aligned}
curl\ \vec{v}' &= \frac{\partial v'_t}{\partial s} - \frac{\partial v'_s}{\partial t} \\
&= \frac{\partial v_x}{\partial s}\frac{\partial x}{\partial t} + v_x\frac{\partial^2 x}{\partial s \partial t} + \frac{\partial v_y}{\partial s}\frac{\partial y}{\partial t} + v_y\frac{\partial^2 y}{\partial s \partial t} - (s \leftrightarrow t) \\
&=_a \frac{\partial v_x}{\partial x}\frac{\partial x}{\partial s}\frac{\partial x}{\partial t} + \frac{\partial v_x}{\partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial t} + \frac{\partial v_y}{\partial x}\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} + \frac{\partial v_y}{\partial y}\frac{\partial y}{\partial s}\frac{\partial y}{\partial t} - (s \leftrightarrow t) \\
&=_b + \frac{\partial v_x}{\partial y}\frac{\partial y}{\partial s}\frac{\partial x}{\partial t} + \frac{\partial v_y}{\partial x}\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - (s \leftrightarrow t) \\
&= \Big(\frac{\partial v_y}{\partial x} - \frac{\partial v_y}{\partial x}\Big)\Big(\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial y}{\partial s}\frac{\partial x}{\partial t}\Big) \\
&= \big(curl\ \vec{v}\big)\Big(\frac{\partial x}{\partial s}\frac{\partial y}{\partial t} - \frac{\partial y}{\partial s}\frac{\partial x}{\partial t}\Big) \\
&= \big(curl\ \vec{v}\big)J_\Gamma(s,t)
\end{aligned}$$

where you should recognise the second factor as the determinant of the matrix of partial derivatives of the map Γ

This shows that if $\vec{v}$ is curl-free, so is $\vec{v}'$.

Hence, by Theorem 12.1, the line integral "around" R vanishes. In other words, using the notation of that theorem (with obvious replacements $(A,B,C,D) \rightsquigarrow (P_1,P_2,P_3,P_4)$ and $\vec{\gamma} \rightsquigarrow \tilde{\gamma}$) we have the equality in the **first** line:

$$\underbrace{\int_{\tilde{\gamma}_{P_1P_2}} \vec{v}' \centerdot d\tilde{\gamma}_{P_1P_2}}_{=\int_{\tilde{\gamma}_0}\vec{v}.d\tilde{\gamma}_0} + \underbrace{\int_{\tilde{\gamma}_{P_2P_3}} \vec{v}' \centerdot d\tilde{\gamma}_{P_2P_3}}_{=0} = \underbrace{\int_{\tilde{\gamma}_{P_1P_4}} \vec{v}' \centerdot d\tilde{\gamma}_{P_1P_4}}_{=0} + \underbrace{\int_{\tilde{\gamma}_{P_4P_3}} \vec{v}' \centerdot d\tilde{\gamma}_{P_4P_3}}_{=\int_{\tilde{\gamma}_1}\vec{v}.d\tilde{\gamma}_1}$$

To finish the proof, we need to prove the inequalities in the second line. We will do the two on the left hand-side; the others follow similarly.

Consider first the line integral of $\vec{v}'$ along the line segment P_1P_2. We parametrise this as follows: $\tilde{\gamma}_{P_1P_2}(s) = (s,0)$. Then

$$\begin{aligned}
\int_{\tilde{\gamma}_{P_1P_2}} \vec{v}' \centerdot d\tilde{\gamma}_{P_1P_2} &\underset{check}{=} \int_0^1 v'_s(s,0)ds \\
&= \int_0^1 \Big(v_x(\Gamma(s,0))\frac{\partial \Gamma_x}{\partial s}(s,0) + v_y(\Gamma(s,0))\frac{\partial \Gamma_y}{\partial s}(s,0)\Big)ds \\
&= \int_0^1 \Big(v_x(\vec{\gamma}_0(s))\frac{d\{\vec{\gamma}_0\}x}{ds}(s) + v_y(\vec{\gamma}(s))\frac{d\{\vec{\gamma}_0\}_y}{ds}(s)\Big)ds \\
&\underset{def}{=} \int_{\vec{\gamma}_0} \vec{v} \centerdot d\vec{\gamma}_0
\end{aligned}$$

Next, consider the integral of $\vec{v}'$ along P_2P_3. We parametrise the path as follows: $\tilde{\gamma}_{P_2P_3}(t) = (1,t)$. Then

$$\int_{\tilde{\gamma}_{P_2P_3}} \vec{v}' \cdot d\tilde{\gamma}_{P_2P_3} \underset{check}{=} \int_0^1 v_t'(1,t)dt$$

But

$$v_t'(1,t) = v_x(\Gamma(1,t))\frac{\partial \Gamma_x}{\partial t}(1,t) + v_y(\Gamma(1,t))\frac{\partial \Gamma_y}{\partial t}(1,t) = 0$$

since $\Gamma(1,t) = B$ for all t. □

12.6. **Any curl-free vector field on a simply connected open set is a gradient vector field.** Consider a non-empty open set $U \subset \mathbb{R}^2$. Let us agree to call U *path-connected* if any two points can be joined by a smoothly (i.e., C^∞) parametrised $\vec{\gamma}$. Let agree that U is *simply connected* if it is path-connected, and given any two smooth paths $\vec{\gamma}_0$ and $\vec{\gamma}_1$ with the same end-points, they can be smoothly interpolated by a smooth Γ as in the Theorem. Clearly any convex open U is simply connected, but convexity is not necessary.

Theorem 12.4. *If U is simply connected and $\vec{v}$ is a curl-free vector field, it is the gradient of a function f.*

Proof. Fix a point $A \in U$, and define f by

$$f(B) = \int_{\vec{\gamma}} \vec{v} \cdot d\vec{\gamma}$$

where $\vec{\gamma}$ is *any* smooth path from A to B. Exercise: Check that f is differentiable and $\vec{v} = grad\ f$. □

You will find another proof for the case when U is all of $\mathbb{R}^2$ or more generally, a "star-shaped region" in Exercise 14.1 below. The proof is in the language of 1-forms.

13. Vector fields, 1-forms, differential forms

For simplicity, all maps will be assumed infinitely differentiable in this section and the next.

Let V be a real vector space with dimension n, and U an open subset. A *vector field*, as defined in Spivak, is a (C^∞) map $F : U \to V$. A 0-*form* on U is a map from U to $\Lambda^0 V = \mathbb{R}$, or in simple terms, a real-valued function f. A k-*form* is map[24] $\omega : U \to \Lambda^k V$; the *degree* of such an ω is k. One-forms are particularly important.

[24]Notation alert! In the section on exterior algebra we used Greek lower case letters ω etc. to denote elements of $\Lambda^k V$, etc. Now we are using them to denote *functions into* $\Lambda^k V$ etc. In other words, in the present context, given $p \in U$, and a k-form ω, we get $\omega(p) \in \Lambda^k V$.

Let $\{e_1, \ldots, e_n\}$ be a basis for V, and $\{\phi_1, \ldots, \phi_n\}$ the dual basis of V^*. The coordinate functions $x^1, \ldots, x^n$ are defined by

$$V \ni p \mapsto \sum_i x^i(p) e_i$$

Equivalently, $x^i(p) = \phi_i(p)$ (i.e., evaluate $\phi_i \in V^*$ on $p \in V$ to get the value of the function x^i at the vector p.)

Given a k form ω and a l-form η, they can be multiplied to give a $k+l$-form $\omega \wedge \eta$:

$$[\omega \wedge \eta](p) \underset{def}{=} \omega(p) \wedge \eta(p)$$

In particular, a k-forms can be multiplied by a 0-form, i.e., a function.

Given a function f on U, we define its *exterior derivative* df to be the 1-form defined by

$df(p)[u] = Df(p)[u]$ = the directional derivative of f at p along the vector u.

Verify that

$$dx^i(p) = \phi_i$$

That is, dx^i is the "constant" 1-form that takes the value $\phi_i \in V^*$ at all points p. (Although we have been talking of forms on an open set U, the functions x^i and their exterior derivatives are defined on all of V.)

I claim that

$$df = \sum_i \frac{\partial f}{\partial x^i} dx^i$$

The proof itself is trivial once you understand the formula, which says that at every p, and for every $u \in V$,

$$Df(p)[u] = \sum_i \frac{\partial f}{\partial x^i}(p) dx^i(p)[u] = \sum_i \frac{\partial f}{\partial x^i}(p) \phi_i(u)$$

It is enough to check this for u running over the basis vectors $e_1, \ldots, e_n$. Taking $u = e_j$, we need to verify:

$$Df(p)[e_j] = \sum_i \frac{\partial f}{\partial x^i}(p) \phi_i(e_j) = \frac{\partial f}{\partial x^j}(p)$$

Repeating a calculation from long ago,

$$\frac{\partial f}{\partial x^j}(p) = \lim_{t \to 0} \frac{f(p + te_j) - f(p)}{t} = Df(p)[e_j]$$

Incidentally, this also gives a proof that if f is C^∞, so is df, since it is a sum

$$\sum_i \frac{\partial f}{\partial x^i} \phi_i$$

of terms of the type "a C^∞ function times a constant one-form".

The exterior derivative map can be uniquely extended to linear maps

$$k\text{-forms on } U \ni \omega \mapsto d\omega \in (k+1)\text{-forms on } U$$

such that

(1) $d(df) = 0$ for any function f, and

(2) given a k form ω and a l-form η,

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$$

The most straightforward way to see this is as follows. If ω is a k-form, it can be written uniquely in the form

$$\omega = \sum_{1 \leq i_1 < \cdots < i_k \leq n} w_{i_1,\ldots,i_k} dx^{i_1} \wedge \ldots dx^{i_k}$$

where the $w_{i_1,\ldots,i_k}$ are functions. The conditions (1) and (2) force the formula:

$$\begin{aligned} d\omega &= \sum_{1 \leq i_1 < \cdots < i_k \leq n} dw_{i_1,\ldots,i_k} \wedge dx^{i_1} \wedge \ldots dx^{i_k} \\ &= \sum_j \sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{\partial w_{i_1,\ldots,i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \ldots dx^{i_k} \end{aligned}$$

Define $d\omega$ as above, and check that (2) is satisfied in general. (Important fact (check): $d \circ d\ \omega = 0$ holds in arbitrary degree.)

To summarise, associated to an open set $U \subset \mathbb{R}^n$ we have the following machinery:

(1) The space of 0-forms on U, i.e., the space of real-valued C^∞ functions. These can be added and multiplied and so form a commutative ring. Since $\mathbb{R}$ can be regarded as a subring (by identifying a real number with the constant function that takes that number as its value everywhere), this is in fact an $\mathbb{R}$-algebra.

(2) The space of exterior k-forms, with k running over $k = 0, 1, \ldots, n$. Each of these is a module over the ring of functions. More generally, we can multiply a k-form ω with an l-form η to get a $k+l$-form $\omega \wedge \eta$. This multiplication is "graded commutative", i.e.,

$$\omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$$

(3) Given a k-form ω, one can define its exterior derivative, a $k+1$-form $d\omega$, with the properties listed above.

14. Pull-backs

All forms (including 0*-forms, i.e., functions) will be assumed* C^∞*.*

Suppose given a (C^∞) map $\Phi : \tilde{U} \to U$ of open sets, with $U \subset V$ and $\tilde{U} \subset \tilde{V}$, with $V, \tilde{V}$ finite-dimensional vector spaces. Given a point $\tilde{p} \in \tilde{U}$, the differential $D\Phi(\tilde{p})$ gives a linear map $\tilde{V} \to V$. However, given a vector field $\tilde{F}$ on $\tilde{U}$ this cannot be used to define one on U except in special circumstances. For example, if Φ is a diffeomorphism, we can define

$$(\Phi_* F)(p) = D\Phi(\tilde{p})[F(\tilde{p})] \quad \textit{where } \tilde{p} = \Phi^{-1}(p)$$

On the other hand, we already know how to "move" functions from U to $\tilde{U}$: simply compose with Φ. (There is *no* requirement that Φ is a diffeomorphism.) Let us formalise this by defining a map Φ^* as follows.

$$\Phi_0^* : \{C^\infty \text{ real-valued functions on } U\} \longrightarrow \{C^\infty \text{ real-valued functions on } \tilde{U}\}$$
$$f \longmapsto \Phi_0^* f = f \circ \Phi$$

This can be extended to forms; there is a unique set of linear maps

$$\Phi_k^* : \{C^\infty \ k\text{-forms on } U\} \longrightarrow \{C^\infty \ k\text{-forms on } \tilde{U}\}$$
$$\eta \longmapsto \Phi^* \eta$$

such that

$$\Phi_{k+1}^* d\omega = d\Phi_k^* \omega, \ \textit{and}$$
$$\Phi_{k+l}^* (\omega \wedge \eta) = \Phi_k^* \omega \wedge \Phi_l^* \eta \ .$$

(As before given ω is a k-form and η an l-form.)

We will henceforth omit the subscript, and write simply $\Phi^* = \Phi_k^*$.

Exercise: Let $\Phi : \tilde{U} \to U$ be as above.

(1) Choose bases on V and $\tilde{V}$, and let $x^1, \ldots, x^n$ and $t^1, \ldots, t^m$ be the respective sets of coordinate functions. Verify that

$$\Phi^* dx^i = \sum_{k=1}^{m} \frac{\partial x^i}{\partial t^j} dt^j$$

Solution: Note that we have proved for any function f on U,

$$df = \sum_i \frac{\partial f}{\partial x^i} dx^i$$

Therefore (or similarly) for any function $\tilde{f}$ on $\tilde{U}$,

$$d\tilde{f} \underset{A}{=} \sum_j \frac{\partial \tilde{f}}{\partial t^j} dt^j$$

Take $\tilde{f} = \Phi^* x^i = x^i \circ \Phi$, which by abuse of notation we write simply as x^i. Then

$$\Phi^* dx^i = d\Phi^* x^j = \sum_j \frac{\partial x^i}{\partial t^j} dt^j$$

Alternatively, one can prove this directly, essentially reproving (A) for the case $\tilde{f} = \Phi^* dx^i$. Let $\tilde{u} = \sum_j$ be vector in $\tilde{V}$, $\tilde{p} \in \tilde{U}$. Then

$$\Phi^* dx^i(\tilde{p})[\tilde{u}] = dx^i(D\Phi(\tilde{p})[\tilde{u}]) =$$

(2) Let η be a one-form on U. Let $\tilde{p} \in \tilde{U}$, and let $p = \Phi(p) \in U$ denote the image of $\tilde{p}$. The following equation relates $\Phi^*\eta(\tilde{p})$ and $\eta(p)$:

$$\Phi^*\eta(\tilde{p})[u] = \eta(p)[D\Phi(\tilde{p})[u]]$$

Hint: it is enough to check this for $\eta = dx^i$ and $u = \frac{\partial}{\partial t^j}$.

(3) Let $m = n$, and choose coordinates $x^1, \ldots, x^n$ and $t^1, \ldots, t^n$. Prove that

$$\Phi^* dx^1 \wedge \cdots \wedge dx^n = J_\Phi dt^1 \wedge \cdots \wedge dt^n$$

Here is an explicit calculation, with $n = 2$:

$$\begin{aligned}
\Phi^* dx^1 \wedge dx^2 &= \Phi^* dx^1 \wedge \Phi^* dx^2 \\
&= \{\frac{\partial x^1}{\partial t^1} dt^1 + \frac{\partial x^1}{\partial t^2} dt^2\} \wedge \{\frac{\partial x^2}{\partial t^1} dt^1 + \frac{\partial x^2}{\partial t^2} dt^2\} \\
&= \{\frac{\partial x^1}{\partial t^1}\frac{\partial x^2}{\partial t^2} - \frac{\partial x^1}{\partial t^2}\frac{\partial x^2}{\partial t^1}\} dt^1 \wedge dt^2
\end{aligned}$$

We have used $dt^1 \wedge dt^1 = 0$, $dt^2 \wedge dt^2 = 0$, and $dt^1 \wedge dt^2 = -dt^2 \wedge dt^1$.

Exercise: (Exterior calculus in $\mathbb{R}^3$) Let x, y, z (rather than x^1, x^2, x^3) denote the standard coordinate functions in $\mathbb{R}^3$. The exterior derivative maps are as follows:

(1) Given a function f,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

(2) Given a 1-form $\omega = \omega_x dx + \omega_y dy + \omega_z dz$,

$$d\omega = (\frac{\partial \omega_y}{\partial x} - \frac{\partial \omega_x}{\partial y}) dx \wedge dy + (\frac{\partial \omega_z}{\partial y} - \frac{\partial \omega_y}{\partial z}) dy \wedge dz + (\frac{\partial \omega_x}{\partial z} - \frac{\partial \omega_z}{\partial x}) dz \wedge dx$$

(3) Given a two-form $\eta = \eta_{xy} dx \wedge dy + \eta_{yz} dy \wedge dz + \eta_{zx} dz \wedge dx$,

$$d\eta = (\frac{\partial \eta_{yz}}{\partial x} + \frac{\partial \eta_{zx}}{\partial y} + \frac{\partial \eta_{xy}}{\partial z}) dx \wedge dy \wedge dz$$

(4) Given a three-form $\rho = \rho_{xyz} dx \wedge dy \wedge dz$, $d\rho = 0$.

Check the above formulae, and check that $d^2 = 0$.

Exercise:(Exterior calculus in $\mathbb{R}^2$) Let x, y (rather than x^1, x^2) denote the standard coordinate functions in $\mathbb{R}^2$. The exterior derivative maps are as follows:

(1) Given a function f,
$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

(2) Given a 1-form $\omega = \omega_x dx + \omega_y dy$,
$$d\omega = (\frac{\partial \omega_y}{\partial x} - \frac{\partial \omega_x}{\partial y})dx \wedge dy$$

(3) Given a two-form $\eta = \eta_{xy} dx \wedge dy$, $d\eta = 0$.

Remark 14.1. Here is another proof that given any 1-form ω on $\mathbb{R}^n$ such that $d\omega = 0$ there is a function f such that $df = \omega$. The following f works:
$$f(x,y) = \int_0^1 \{x\omega_x(xt,yt) + y\omega_y(xt,yt)\}dt$$
For
$$\begin{aligned}\frac{\partial f}{\partial x} &= \int_0^1 \{\omega_x(xt,yt) + xt\frac{\partial \omega_x(xt,yt)}{\partial x} + yt\frac{\partial \omega_y(xt,yt)}{\partial x}\}\\ &= \int_0^1 \{\omega_x(xt,yt) + xt\frac{\partial \omega_x(xt,yt)}{\partial x} + yt\frac{\partial \omega_x(xt,yt)}{\partial y}\}\\ &= \int_0^1 \frac{d}{dt}[t\omega_x(xt,yt)]\\ &= t\omega(xt,yt)|_0^1 = \omega_x(x,y)\end{aligned}$$
The "differentiation under the integral sign" has to be justified.

(This is a special case of the Poincaré Lemma, which holds in greater generality. Note, for example, that exactly the same proof would work if ω was only defined on an open set U with the property[25]: there is a base-point $(x_0, y_0) \in U$ such that for every point $(x, y) \in U$, the straight line
$$[0,1] \ni t \mapsto t(x,y) + (1-t)(x_0,y_0)$$
is contained in U. See Spivak.)

Check the above formulae, and check that $d^2 = 0$.

[25]of being "star-shaped".

15. Singular cubical chains

I begin with a very useful construction. This is exercise 4-22 of Spivak, put in a general context. Let $\mathcal{S}$ be an arbitrary set. The free $\mathbb{Z}$-module generated by $\mathcal{S}$ is the space of maps $n : \mathcal{S} \to \mathbb{Z},\ c \mapsto n(c)$ such that $n(c) = 0$ for all but finitely many $c \in S$. Two such functions can be added, and any function can be multiplied by an integer:

$$(n + n')(c) \underset{def}{=} n(c) + n'(c)$$
$$(an)(c) \underset{def}{=} an(c),\ a \in \mathbb{Z}$$

We will represent a function $c \mapsto n(c)$ as a $\sum_c n(c)c$, where the summation runs over all $c \in S$, but only finitely many $n(c)$ are nonzero.

Let A be a topological space. (In the present context we will take A to be a subset of $\mathbb{R}^m$, usually open.) A *singular n-cube in A* is a map $c : [0,1]^n \to A$. For the present c is taken to be continuous, eventually we will require c to be C^∞, which by definition means that it extends as a C^∞ map from an open set containing $[0,1]^n$.

Warning! (a) Spivak writes $A \subset \mathbb{R}^n$; this should read $A \subset \mathbb{R}^m$, since the letter n [has already been]/[will be] used for the dimension of the cube. (b) Spivak does not specify – except in one place – when his maps are smooth. While topology works with continuous maps, calculus requires differentiability.

A singular (cubical[26]) n-chain (in A) is an element of the free $\mathbb{Z}$-module generated by n-cubes (in A). In other words, it is a sum

$$\sum_i n_i c_i$$

where n_i are integers (such that all but finitely many are zero) and the c_i are n-cubes in A.

A 0-cube is a map $\{0\} \to A$, and can be identified with its image. So a 0-chain is an element of the free $\mathbb{Z}$-module generated by A, and can be represented by a formal sum $\sum_i n_i p_i$, with $p_i \in A$ and integers n_i, with finitely many nonzero.

Given an n-cube c, we can define its *boundary* ∂c, which is a $(n-1)$-chain. Once this is done, the boundary of an n-chain is defined by linearity:

$$\partial(\sum_i n_i c_i) \underset{def}{=} \sum_i n_i \partial c_i$$

[26]This is opposed to simplicial chains, which are more common in topology. The theory can be developed with simplicial chains as well, but we use cubical chains since Spivak does.

You can find the definition of boundary of a cube in Spivak. Here I deal with the special cases of 1-cubes and 2-cubes. (The boundary of a 0-cube is defined to be zero.)

(1) A 1-cube is a map $\gamma : [0,1] \to A$, and

$$\partial\gamma = \gamma(1) - \gamma(0)$$

(2) A 2-cube is a map $c : [0,1]^2 \to A$. Consider first the case when the 2-cube in question is the identity map $I^2 : [0,1]^2 \to [0,1]^2$. We first define

$$\partial I^2 = -I^2_{(1,0)} + I^2_{(1,1)} + I^2_{(2,0)} - I^2_{(2,1)}$$

where each of the terms on the right is a 1-cube in $[0,1]^2$, each mapping the interval $[0,1]$ onto one of the edges of the square $[0,1]^2$:

$$I^2_{(1,0)}(t) = (0,t)$$
$$I^2_{(1,1)}(t) = (1,t)$$
$$I^2_{(2,0)}(t) = (t,0)$$
$$I^2_{(2,1)}(t) = (t,1)$$

Note that

$$\begin{aligned}\partial(\partial I^2) &= -I^2_{(1,0)}(1) + I^2_{(1,0)}(0) + I^2_{(1,1)}(1) - I^2_{(1,1)}(0) + I^2_{(2,0)}(1) - I^2_{(2,0)}(0) - I^2_{(2,1)}(1) + I^2_{(2,1)}(0)\\ &= -(0,1) + (0,0) + (1,1) - (1,0) + (1,0) - (0,0) - (1,1) + (0,1) = 0\end{aligned}$$

We define the boundary of an arbitrary 2-cube by

$$\partial c = c \circ \partial I^2 \equiv -c \circ I^2_{(1,0)} + c \circ I^2_{(1,1)} + c \circ I^2_{(2,0)} - c \circ I^2_{(2,1)}$$

and the boundary of a 2-chain $\sum_i n_i c_i$ by linearity.

Exercise $\partial(\partial c) = 0$.

In general, given an singular n-cube, ∂c is defined analogously:

$$\partial c \underset{def}{=} c \circ \partial I^n$$

We refer to Spivak for the definition of ∂I^n; also proved there is the fact that $\partial(\partial I^n) = \partial I^n \circ \partial I^{n-1} = 0$. As a consequence

$$\partial(\partial c) = \partial(c \circ \partial I^n) = (c \circ \partial I^n) \circ \partial I^{n-1} = c \circ (\partial I^n \circ \partial I^{n-1}) = c \circ \partial^2 I^n = 0$$

16. Stokes' Theorem in one dimension

Let $\tilde{\omega}$ be a 1-form defined on an open interval $I \supset [0,1]$. If t denotes the natural coordinate on I, $\tilde{\omega} = \tilde{\omega}_t dt$, with $\tilde{\omega}_t$ a 0-form. i.e., a real-valued function. We define

$$\int_{[0,1]} \tilde{\omega} = \int_0^1 \tilde{\omega}_t(t) dt$$

where on the right we mean the Riemann integral of the function $\tilde{\omega}_t$. *Note that the dt on the right harks back to our notation for the Riemann integral,*

whereas when we write $\tilde{\omega} = \tilde{\omega}_t dt$, dt is a one-form on I, i.e., a map from I to $\Lambda^1\mathbb{R} = \mathbb{R}^* = \mathbb{R}$.

Let ω be a C^∞ one-form an open set U in $\mathbb{R}$. Consider a singular 1-cube, i.e., a C^∞ map $\gamma : [0,1] \to U$. Recall that to say γ is C^∞ means that it extends as a C^∞ map (for which we retain the same notation) from an open interval $\tilde{I} \supset [0,1]$. We will let x denote the natural coordinate function on U and t the coordinate function on $\tilde{I}$. Then

(1) $\omega = \omega_x dx$ for a (unique) function ω_x,

(2) $\gamma^*\omega = (\omega_x \circ \gamma)(dx/dt)dt \equiv \tilde{\omega}_t dt$

We set

$$\int_\gamma \omega \underset{def}{=} \int_{[0,1]} \gamma^*\omega = \int_0^1 \tilde{\omega}_t(t)dt = \int_0^1 \omega_x(\gamma(t))\frac{dx}{dt}(t)dt$$

Suppose now that $\omega = df$ for some function f on U. Then

$$\int_\gamma df = \int_0^1 \frac{df}{dx}(\gamma(t))\frac{dx}{dt}(t) = \int_0^1 \frac{d}{dt}(f \circ \gamma)(t)dt = f(\gamma(1)) - f(\gamma(0))$$

Now $\partial\gamma$ is the singular 0-chain $\gamma(1) - \gamma(0)$, so the above equation can be rewritten

$$\int_\gamma df = \int_{\partial\gamma} f$$

There is no reason to restrict ourselves to a singular 1-cube on the LHS since on the right we have admitted an integral linear combination of 0-cubes. Let then $\sum_i n_i\gamma_i$ be an integral linear combination of 1-cubes in U, i.e., a 1-chain in U. Define, for any 1-form on ω on U,

$$\int_{\sum_i n_i\gamma_i} \omega = \sum_i n_i \int_{\gamma_i} \omega = \sum_i n_i \int_{[0,1]} \gamma_i^*\omega$$

Then we have

$$\int_{\sum_i n_i\gamma_i} df \underset{def}{=} \sum_i n_i \int_{\gamma_i} df = \sum_i n_i \int_{\partial\gamma_i} f \underset{def}{=} \int_{\partial\{\sum_i n_i\gamma_i\}} f$$

17. Stokes' Theorem, and proof in two dimesions

Here is the statement in general, which should come as no surprise:

Theorem 17.1. *Let η be a $k-1$-form on an open set $U \subset V$, with V an n-dimensional real vector space. Let $\sum_i n_i c_i$ be an k-chain in U. Then*

$$\int_{\sum_i n_i c_i} d\eta = \int_{\partial\{\sum_i n_i c_i\}} \eta$$

Proof. (In the case $k = 2$.) It is enough to do this for one 2-cube c; the case of a 2-chain follows by linearity. So let η be a 1-form on U, and let $c : [0,1] \to U$ be the restriction of a C^∞ map $\tilde{c} : \tilde{U} \to U$ where $\tilde{U}$ is an open

set in $\mathbb{R}^2$ containing $[0,1]^2$. Set $\tilde{c}^*\eta \equiv \tilde{\eta}$. Write $\tilde{\eta} = \tilde{\eta}_x dx + \tilde{\eta}_y dy$ where $\tilde{\eta}_x$ and $\tilde{\eta}_y$ are C^∞ functions on $\tilde{U}$; then

$$d\tilde{\eta} = (\frac{\partial \tilde{\eta}_y}{\partial dx} - \frac{\partial \tilde{\eta}_x}{\partial y}) dx \wedge dy$$

and

$$\int_c d\eta \underset{def}{=} \int_{[0,1]^2} c^* d\eta = \int_{[0,1]^2} dc^*\eta = \int_{[0,1]^2} d\tilde{\eta} \underset{def}{=} \int_{[0,1]^2} (\frac{\partial \tilde{\eta}_y}{\partial x} - \frac{\partial \tilde{\eta}_x}{\partial y}) dx dy$$

where in the last expression we mean the Riemann integral.

Let us evaluate the integral using Fubini. Note that although η lives on U, the integral is being done on $[0,1]^2$. We have

$$\begin{aligned}
\int_{[0,1]^2} (\frac{\partial \tilde{\eta}_y}{\partial x}) dx dy &= \int_0^1 \{ \int_0^1 (\frac{\partial \tilde{\eta}_y}{\partial x}) dx \} dy \\
&= \int_0^1 \tilde{\eta}_y(1,y) dy - \int_0^1 \tilde{\eta}_y(0,y) dy \\
&= \int_{c \circ I^2_{(1,1)}} \eta - \int_{c \circ I^2_{(1,0)}} \eta
\end{aligned}$$

where in all but the last line we mean the Riemann integral. Applying Fubini to the second term, but this time reversing the order of integration, we get

$$\begin{aligned}
\int_{[0,1]^2} (\frac{\partial \tilde{\eta}_x}{\partial y}) dx dy &= \int_0^1 \{ \int_0^1 (\frac{\partial \tilde{\eta}_x}{\partial y}) dy \} dx \\
&= \int_0^1 \tilde{\eta}_x(x,1) dx - \int_0^1 \tilde{\eta}_x(x,0) dx \\
&= \int_{c \circ I^2_{(2,1)}} \eta - \int_{c \circ I^2_{(2,0)}} \eta
\end{aligned}$$

Gathering terms, we get

$$\begin{aligned}
\int_c d\eta &= \int_{c \circ I^2_{(1,1)}} \eta - \int_{c \circ I^2_{(1,0)}} \eta - \int_{c \circ I^2_{(2,1)}} \eta + \int_{c \circ I^2_{(2,0)}} \eta \\
&= \int_{c \circ \{I^2_{(1,1)} - I^2_{(1,0)} - I^2_{(2,1)} + I^2_{(2,0)}\}} \eta = \int_{c \circ \partial I^2} \eta = \int_{\partial c} \eta
\end{aligned}$$

□

Exercise

(1) Consider the following 1-form on $\mathbb{R}^2 \setminus \{(0,0)\}$:

$$\eta \equiv -\frac{y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$$

(a) Show that $d\eta = 0$.

(b) Let $\phi : \{(r,\theta) | r > 0\} \subset \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{(0,0)\}$ be the map $\phi(r,\theta) = (r\cos\theta, r\sin\theta)$. Show that $\phi^*\eta = d\theta$.

(c) For $n \in \mathbb{Z}$, let γ_n be the 1-cube in $\mathbb{R}^2 \setminus \{(0,0)\}$:

$$\gamma_n(t) = (\cos 2\pi nt, \sin 2\pi nt)$$

Prove that

$$\int_{\gamma_n} \eta = 2\pi n$$

(d) Conclude that there does not exist a 2-chain c in $\mathbb{R}^2 \setminus \{(0,0)\}$ such that $\partial c = \gamma_1$.

(e) Conclude that there does not exist a function θ such that $d\theta = \eta$.

For a very clear – and more down-to-earth – account of Stokes and forms, see

`https://www.math.purdue.edu/~arapura/preprints/diffforms.pdf`.

An old classic is *Differential Forms with Applications to the Physical Sciences* (Dover Books on Mathematics) by H. Flanders.

